The Covariant Galileon Model of Dark Energy

Koekenbier, L.

The Covariant Galileon Model of Dark Energy

THESIS

submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF SCIENCE
in
PHYSICS & MATHEMATICS

Author :	L.R.J. Koekenbier
Student ID :	1530224
Supervisor for Physics :	A. Silvestri
Supervisor for Mathematics :	H.J. Hupkes

Leiden, The Netherlands, 29 August 2018

The Covariant Galileon Model of Dark Energy

L.R.J. Koekenbier

Huygens-Kamerlingh Onnes Laboratory, Leiden University &
Mathematisch Instituut, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

29 August 2018

Abstract

The Covariant Galileon Model is an extension of General Relativity constructed by adding an extra scalar degree of freedom to it. The mathematical background of the model is therefore also found in differential geometry. The equations of motion of the model can be derived from its Lagrangian. Using the ADM formalism and tools from differential geometry the EFT functions of the model are then found. Numerical solutions to the model are given for two different sets of parameters and for variations of the the present day matter density and the Hubble constant at the start of the simulation. From these it is concluded that a more thorough Monte Carlo simulation of the model is a useful tool for further analysis of the model. Furthermore more research is needed for a better interpretation of the found solutions to the model.

Contents

Chapter 1

Introduction

Since the dawn of man, mankind has always looked up to the sky and wondered about the contents of the cosmos, trying to unveil the mysteries of the universe. Our current understanding of the universe on these grand scales is given by Einstein's Theory of General Relativity. It posits that space and time form one geometrical structure and that gravity is the curvature of that structure. The inception of this theory meant a huge leap forward in our understanding of the universe and it brought forth many big technological advancements. Currently however challenges to the theory have arisen. In particular the accelerated expansion of the universe can not be satisfactorily explained with standard General Relativity. The cosmological constant can help solve these issues, but for theoretical reasons this approach is not favoured by everyone. Therefore new models of gravity have been sought to explain among others this accelerated expansion. One such model is the Covariant Galileon Model, which will be discussed here. The goal of this thesis is to give an understanding of this model and to show how this model can be falsified.

The Covariant Galileon Model is an extension of General Relativity and thus cannot be understood without knowledge of General Relativity. As such, this thesis will first dedicate itself to give a quick overview of the mathematics involved in these theories, before moving to reviewing the basics of General Relativity and then giving the details of the Covariant Galileon Model. The thesis will end with a numerical investigation of the Covariant Galileon Model. This numerical analysis will show how measurements can be used to falsify this model.

Now a little bit of technical details about this thesis. Firstly, if not otherwise specified, differentiable shall mean $\mathcal{C}^\infty$. Secondly, Einstein's summation convention will be used throughout this thesis. This means that, unless stated otherwise, whenever in a term of an equation two indices are repeated, one being a lower index and the other an upper index, then there is a sum implied over all values of the indices. Greek indices will always take values in $\{0, 1, 2, 3\}$, while from chapter 3 onward Latin indices take values in $\{1, 2, 3\}$. Note that the i in $\frac{\partial}{\partial u^i}$ is considered to be a lower index.

Chapter 2

Geometry

In this chapter the mathematical foundation will be laid for the physical theories that will be discussed in the next two chapters. The mathematics discussed will involve differential geometry. This branch of mathematics concerns itself with the properties of objects that exhibit a certain type of smoothness. In particular differential manifolds will be discussed. These objects are mathematical spaces that model certain types of smooth shapes. They provide a rich mathematical structure that in the next chapters will turn out to be fundamental to the description of Nature of the physical theories discussed. This chapter will give an overview of the properties of differential manifolds.

2.1 Manifolds

A differential manifold is a topological space that locally looks like $\mathbb{R}^n$. This notion of 'local sameness' will be made precise in the following definitions, but the general idea will be sketched here first. A differentiable manifold will be locally like $\mathbb{R}^n$ in the sense that around any point in the differentiable manifold there exists a region that is homeomorphic to an open subset of $\mathbb{R}^n$. Such a local homeomorphism will be called a chart and the set of all charts will be called an atlas. In this way there is a translation from the possibly complicated differentiable manifold to the simpler space $\mathbb{R}^n$ in the same way a chart in a 'normal' atlas, the one you can buy in a store, gives a translation between the complicated structure of a part of the surface of the Earth and the simple piece of paper on which the chart is written. This translation can then be exploited to lift structure from $\mathbb{R}^n$ to the differentiable manifold. In particular a notion of differentiability can be defined this way, hence the name differentiable manifold.

Now in order to define a differentiable manifold, let M be a topological space. First the charts and atlas need to be defined:

Definition 2.1.1. An n-dimensional differentiable atlas for M is a set

$$\mathcal{A} = \{(U_i, h_i, V_i) \colon i \in I\}$$

where I is an index set and:

1. $\{U_i \colon i \in I\}$ is an open cover of M
2. $V_i \subset \mathbb{R}^n$ is open for all $i \in I$
3. $h_i \colon U_i \to V_i$ is a homeomorphism for all $i \in I$
4. the gluing maps $(h_j \circ h_i^{-1})|_{h_i(U_i \cap U_j)} \colon h_i(U_i \cap U_j) \to h_j(U_i \cap U_j)$ are differentiable for all $i, j \in I$

An element of an atlas is called a chart. Points 1. to 3. thus define a chart. The fourth point will make it possible to lift the differentiability from $\mathbb{R}^n$ to M. Since there are multiple possible atlases available for a given M, this can lead to multiple possible notions of differentiability.

To clear this ambiguity an equivalence relation can now be defined on the set of n-dimensional differentiable atlases for M. Let $\mathcal{A} = \{(U_i, h_i, V_i) \colon i \in I\}$ and $\mathcal{A}' = \{(U'_j, h'_j, V'_j) \colon j \in J\}$ be two such atlases. They are called equivalent, notated as $\mathcal{A} \sim \mathcal{A}'$, if their union $\mathcal{A} \cup \mathcal{A}'$ is again an n-dimensional differentiable atlas for M. This is the case if and only if for all $i \in I$ and $j \in J$ the maps

$$(h'_j \circ h_i^{-1})|_{h_i(U_i \cap U'_j)} \colon h_i(U_i \cap U'_j) \to h'_j(U_i \cap U'_j)$$
$$(h_i \circ h_j'^{-1})|_{h'_j(U_i \cap U'_j)} \colon h'_j(U_i \cap U'_j) \to h_i(U_i \cap U'_j)$$

are differentiable.

Now to single out a specific notion of differentiability arising from the atlases, the following definition is made:

Definition 2.1.2. An n-dimensional differentiable atlas $\mathcal{A}$ for M is called an n-dimensional differentiable structure in M if it satisfies:

$$\mathcal{A}' \sim \mathcal{A} \implies \mathcal{A}' \subset \mathcal{A}$$

The equivalence class of an atlas $\mathcal{A}$ thus contains a unique n-dimensional differentiable structure in M

$$\mathcal{S}_{\mathcal{A}} = \bigcup_{\mathcal{A}' \sim \mathcal{A}} \mathcal{A}'$$

and a differentiable structure can thus be specified by giving an atlas of the equivalence class of the differentiable structure.

The definition of a differentiable manifold is now as follows:

Definition 2.1.3. An n-dimensional differentiable manifold is a pair $(M, \mathcal{S})$ where M is a second countable Hausdorff topological space and $\mathcal{S}$ is an n-dimensional differentiable structure in M.

Now also substructures can be defined:

Definition 2.1.4. Let X be an n-dimensional differentiable manifold. A subset $Y \subset X$ is called a k-dimensional differentiable submanifold of X if for every $p \in Y$ there exists a chart (U, h, V) for X such that $p \in U$ and

$$h(U \cap Y) = \{x = (x_1, \dots, x_n) \in V \colon x_{k+1} = \cdots = x_n = 0\}$$

Submanifold substructures turn out to be manifolds in their own right:

Proposition 2.1.5. *Let X be an n-dimensional differentiable manifold and Y be a k-dimensional differentiable submanifold of X, then Y is a k-dimensional differentiable manifold.*

Proof. Take the subspace topology on Y, then Y is also a second countable Hausdorff topological space. For every $p \in Y$ take the chart (U_p, h_p, V_p) for X as given in Definition 2.1.4 and let the set $\mathcal{A} = \bigcup_{p \in Y} \{(U_p^Y, h_p^Y, V_p^Y)\}$ be given by:

$$\begin{aligned} U_p^Y &= U_p \cap Y \\ h_p^Y &= \pi \circ (h_p|_{U_p^Y}) \\ V_p^Y &= \pi(V_p) \end{aligned}$$

where $\pi : \mathbb{R}^n \to \mathbb{R}^k$ is the projection on the first k coordinates. In this way $\{U_p^Y \colon p \in Y\}$ is an open cover of Y. Since the projection is an open map and the V_p are open, the V_p^Y are also open.

Now let $\iota \colon \mathbb{R}^k \to \mathbb{R}^n$ be the inclusion map of $\mathbb{R}^k$ into the first k coordinates of $\mathbb{R}^n$. Since the topology on $\mathbb{R}^k$ and the subspace topology on

$\iota(\mathbb{R}^k)$ agree, ι is a homeomorphism onto its image, with inverse $\pi|_{\iota(\mathbb{R}^k)}$. It holds $h_p^Y(U_p^Y) = \pi \circ (h_p|_{U_p^Y})(U_p \cap Y) = \pi(V_p \cap \iota(\mathbb{R}^k)) = V_p^Y$ by definition of h_p, so $h_p^Y\colon U_p^Y \to V_p^Y$ is a bijection. Since the projection and inclusion are continuous and h_p is a homeomorphism, $h_p^Y = \pi \circ (h_p|_{U_p^Y})$ and $(h_p^Y)^{-1} = (h_p)^{-1} \circ (\pi^{-1}|_{V_p^Y}) = (h_p)^{-1} \circ \iota|_{V_p^Y}$ are continuous. Hence h_p^Y is a homeomorphism.

Now for two $p, q \in Y$ the gluing map is:

$$h_p^Y \circ (h_q^Y)^{-1}|_{h_q^Y(U_p^Y \cap U_q^Y)} = (\pi \circ h_p \circ (h_q)^{-1} \circ \iota)|_{h_q^Y(U_p^Y \cap U_q^Y)}$$

Since π and ι are differentiable and $h_p \circ (h_q)^{-1}$ is differentiable by definition, the gluing map $h_p^Y \circ (h_q^Y)^{-1}|_{h_q^Y(U_p^Y \cap U_q^Y)}$ is differentiable for all $p, q \in Y$. This shows that $\mathcal{A}$ is a k-dimensional differentiable atlas for Y. Taking the associated differentiable structure $\mathcal{S}_\mathcal{A}$ in Y (from the equivalence class of $\mathcal{A}$) makes $(Y, \mathcal{S}_\mathcal{A})$ into a k-dimensional differentiable manifold. □

Now an example:

Example 2.1.6. Consider a second countable Hausdorff n-dimensional topological vector space V. Take a basis $(e_i)_i$ and let $\phi\colon V \to \mathbb{R}^n$ be the map given by $\phi(v) = \phi(v^i e_i) = (v^1, \ldots, v^n)$. Then it is clear that $\{(V, \phi, \mathbb{R}^n)\}$ is an n-dimensional differentiable atlas for V. By taking the associated differentiable structure of this atlas, V becomes an n-dimensional differentiable manifold.

Using the atlas, coordinates can be defined on the differentiable manifold:

Definition 2.1.7. Let $(M, \mathcal{A})$ be an n-dimensional differentiable manifold and let x^1 to x^n be the coordinate functions of $\mathbb{R}^n$, so for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ it holds $x^i(x) = x_i$ for all i. Choose a chart $(U, h, V) \in \mathcal{A}$. Then the functions $x^1 \circ h$ to $x^n \circ h$ are called a local coordinate system in U. Any $p \in U$ can now be written as $p = (p_1, \ldots, p_n)$, where $p_i = (x^i \circ h)(p)$ is called the i-th coordinate of p.

This local coordinate system transfers the coordinate structure of $\mathbb{R}^n$ to the differentiable manifold. Since the U cover M, a local coordinate system can be found around any point of M.

There is more structure that can be carried over to the differentiable manifold. The differentiable structure allows namely for the lift of differentiability on $\mathbb{R}^n$ to differentiability on manifolds. For this we have the following two definitions:

Definition 2.1.8. Let X be an n-dimensional differentiable manifold with n-dimensional differentiable structure $\mathcal{M} = \{(U_i, h_i, V_i) : i \in I\}$. A function $f : X \to \mathbb{R}$ is called differentiable if for every $i \in I$ the function $f \circ h_i^{-1} : V_i \to \mathbb{R}$ is differentiable.

With respect to a given coordinate system u^1 to u^n defined by a chart (U, h, V) the partial derivative of f with respect to u^i at $p \in X$ is defined as $\frac{\partial f}{\partial u^i}(p) = \frac{\partial f \circ h^{-1}}{\partial x^i}(h(p))$.

Definition 2.1.9. Let $(X, \mathcal{M})$ and $(X', \mathcal{M}')$ be n- and m-dimensional differentiable manifolds respectively and $f : X \to Y$ a map. Then f is called differentiable if it is continuous and the function

$$(h' \circ f \circ h^{-1})|_{h(U \cap f^{-1}(U'))} : h(U \cap f^{-1}(U')) \to V'$$

is differentiable for all $(U, h, V) \in \mathcal{M}$ and $(U', h', V') \in \mathcal{M}'$.

With respect to a given coordinate system u^1 to u^n defined by a chart (U, h, V) on X and coordinate system u'^1 to u'^m defined by a chart (U', h', V') on X' the i-th coordinate of the partial derivative of f with respect to u^j at $p \in X$ is defined as $(u'^i \circ \frac{\partial f}{\partial u^j})(p) = \frac{\partial u'^i \circ f}{\partial u^j}(p) = \frac{\partial x'^i \circ h' \circ f \circ h^{-1}}{\partial x^j}(h(p))$, where the x^i denote the coordinate functions on V and x'^i denote the coordinate functions on V'. From this definition it is clear that the chain rule holds for differentiable functions on manifolds.

Fortunately one doesn't need to check differentiability on the whole differentiable structure:

Lemma 2.1.10. *Let $\mathcal{A} \subset \mathcal{M}$ be an atlas for X and $f : X \to \mathbb{R}$ a function. If $f \circ h_i^{-1}$ is differentiable for all $(U_i, h_i, V_i) \in \mathcal{A}$, then f is differentiable.*

Lemma 2.1.11. *Let $\mathcal{A} \subset \mathcal{M}$ and $\mathcal{B} \subset \mathcal{N}$ be an atlas for X respectively Y and $f : X \to Y$ a map. If $h_j \circ f \circ h_i^{-1}$ is differentiable for all $(U_i, h_i, V_i) \in \mathcal{A}$ and $(U_j, h_j, V_j) \in \mathcal{B}$, then f is differentiable.*

Proof. These Lemmas both follow from the fact that the gluing maps are differentiable. □

In the rest of this thesis every time the term manifold is used the adjective differentiable is implied and the choice of a suitable differentiable structure is assumed.

2.2 Tangent spaces

Since a manifold in general is a topological space, it has no natural way of comparing its elements, in contrast to for example a vector space. The differential structure however gives a way to be able to make this comparison locally. The concept needed for this comparison is that of the tangent space. This section is dedicated to its construction.

The tangent space shall rely on the idea of the velocity of curves on the manifold. To define this, the following definition is given:

Definition 2.2.1. Let M be a manifold. A differentiable curve on M is a differentiable map $\gamma\colon(-\epsilon,\epsilon)\to M$ for some $0<\epsilon$. A differentiable curve through a point $p\in M$ is a differentiable curve γ on M such that $\gamma(0)=p$. The space of all curves through p is denoted by K_p.

Let $p\in M$ and let (U,h,V) be a chart around p. Now an equivalence relation on K_p can be defined by:

$$\gamma\sim\gamma' \iff \frac{dh\circ\gamma}{dt}(0)=\frac{dh\circ\gamma'}{dt}(0)$$

It is easy to see that this definition is independent from the chosen chart and thus well defined. The equivalence class of an element $\gamma\in K_p$ will be noted as $[\gamma]$. This equivalence relation can be thought of as grouping together differentiable curves on M with a common velocity vector at p.

Denote by F_p the set of differentiable functions to $\mathbb{R}$ defined on a neighbourhood around p. F_p forms a real algebra under pointwise operations. Given an equivalence class $[\gamma]$ with representative $\gamma\in K_p$ a map function can now be defined by:

$$X_{[\gamma]}\colon F_p\to\mathbb{R},\ f\mapsto\frac{df\circ\gamma}{dt}(0)$$

This map is called the tangent vector to $[\gamma]$ at p. By the chain rule this definition is independent from the chosen representative, so the function is well defined. Note that the functions $X_{[\gamma]}$ are linear due to the fact that the derivative is linear and F_p is a real algebra.

Having this the tangent space can now be defined:

Definition 2.2.2. Let M be a manifold and let $p\in M$. The tangent space T_pM of M at p is the set $T_pM=\{X_{[\gamma]}\colon\gamma\in K_p\}$.

An element $X_{[\gamma]}$ of the tangent space can be thought of as a derivative in the direction of γ, tangent to the manifold at p. The tangent space itself can thus be seen as the set of all directional derivatives at p in directions tangent to the manifold at p. This becomes more clear in the following theorem:

Theorem 2.2.3. *Let M be an n-dimensional manifold and let $p \in M$, then the tangent space T_pM of M at p is an n-dimensional real vector space.*

Proof. (Adapted from Kobayashi and Nomizu [1]) Since the functions $X_{[\gamma]}$ are linear, T_pM is a subset of the vector space $Hom(F_p, \mathbb{R})$. Choose a local coordinate system u^1 to u^n around p. The functions $\frac{\partial}{\partial u^i}|_p \colon F_p \to \mathbb{R}$ are linear for all i, so they are elements of $Hom(F_p, \mathbb{R})$. It is going to be proved that these functions are a basis for T_pM. Let $\gamma \in K_p$ and write $\gamma_i(t)$ for the coordinates of $\gamma(t)$ in the chosen coordinate system. For $f \in F_p$ it holds:

$$\begin{aligned} X_{[\gamma]}f &= \frac{df \circ \gamma}{dt}(0) = \frac{df \circ h^{-1} \circ h \circ \gamma}{dt}(0) \\ &= \sum_i \frac{\partial f \circ h^{-1}}{\partial x^i}(h(\gamma(0)))\frac{dx^i \circ h \circ \gamma}{dt}(0) \\ &= \sum_i \frac{\partial f \circ h^{-1}}{\partial x^i}(h(p))\frac{d\gamma_i}{dt}(0) = \sum_i \frac{\partial f}{\partial u^i}(p)\frac{d\gamma_i}{dt}(0) \end{aligned}$$

Hence every $X_{[\gamma]}$ is a linear combination of the $\frac{\partial}{\partial u^i}|_p$. Now let $\sum_i a_i \frac{\partial}{\partial u^i}|_p$ be a linear combination of the $\frac{\partial}{\partial u^i}|_p$. Define a curve γ by setting its coordinates to $\gamma_i(t) = u^i(p) + a_i t$ for all i. The associated tangent vector is then given by, for $f \in F_p$:

$$X_{[\gamma]}f = \frac{df \circ \gamma}{dt}(0) = \sum_i \frac{\partial f}{\partial u^i}(p)\frac{d\gamma_i}{dt}(0) = \sum_i a_i \frac{\partial f}{\partial u^i}(p)$$

where the result of the previous equation was used. So $\sum_i a_i \frac{\partial}{\partial u^i}|_p$ is indeed an element of T_pM. This means that all elements of T_pM are linear combinations of the functions $\frac{\partial}{\partial u^i}|_p$. Since the $\frac{\partial}{\partial u^i}|_p$ are part of the vector space $Hom(F_p, \mathbb{R})$, the $\frac{\partial}{\partial u^i}|_p$ induce a linear structure on T_pM and T_pM becomes an n-dimensional real vector space with basis $B = (\frac{\partial}{\partial u^i}|_p \colon 1 \leq i \leq n)$. □

The explicit vector space structure in terms of curves is now found as follows. Let $\gamma, \eta \in K_p$ and $\lambda, \mu \in \mathbb{R}$, then, because T_pM is a vector space, there exists a $\psi \in K_p$ such that $X_{[\psi]} = \lambda X_{[\gamma]} + \mu X_{[\eta]}$. Let $(u^i)_i$ be a local coordinate system around p. For an $f \in F_p$ it thus holds:

$$\begin{aligned}\frac{\partial f}{\partial u^i}(p)\frac{d\psi_i}{dt}(0) &= X_{[\psi]}f = \lambda X_{[\gamma]} + \mu X_{[\eta]} \\ &= \lambda \frac{\partial f}{\partial u^i}(p)\frac{d\gamma_i}{dt}(0) + \mu \frac{\partial f}{\partial u^i}(p)\frac{d\eta_i}{dt}(0) \\ &= \frac{\partial f}{\partial u^i}(p)\frac{d\lambda\gamma_i + \mu\eta_i}{dt}(0)\end{aligned}$$

Thus ψ is a curve whose velocity vector is the sum, weighted appropriately by λ and μ, of the velocity vectors of γ and η. This will also be denoted as $[\psi] = \lambda[\gamma] + \mu[\eta]$.

Since $B = (\frac{\partial}{\partial u^i}|_p \colon 1 \leq i \leq n)$ is a basis of T_pM, the explicit dependence of the tangent vectors on the differentiable curve equivalence class $[\gamma]$ will often be suppressed. Furthermore, note that for all $X \in T_pM$ it holds for $f, g \in F_p$ that $X(fg) = f(p)X(g) + g(p)X(f)$.

Now an example:

Example 2.2.4. Let V be a second countable Hausdorff n-dimensional topological vector space. It was shown in Example 2.1.6 that V is a manifold with atlas $\{(V, \phi, \mathbb{R}^n)\}$, for basis $(e_i)_i$. Then ϕ gives a coordinate system $(u^i)_i$ on whole V. Let $p \in V$, then T_pV is an n-dimensional vector space with basis $(\frac{\partial}{\partial u^i}|_p)_i$. Hence V and T_pV are isomorphic with isomorphism given by $\psi(v) = \psi(v^i e_i) = v^i \frac{\partial}{\partial u^i}|_p$.

Since the tangent space is a vector space, it has a dual:

Definition 2.2.5. Let M be a manifold and let $p \in M$. The cotangent space T_p^*M of M at p is the dual to T_pM. An element of T_p^*M is called a covector. The basis of T_p^*M dual to the basis $B = (\frac{\partial}{\partial u^i}|_p \colon 1 \leq i \leq n)$ of T_pM is denoted by $B^* = ((du^i)_p \colon 1 \leq i \leq n)$, so $(du^i)_p(\frac{\partial}{\partial u^j}|_p) = 1$ if $i = j$ and zero otherwise.

If there is a map between two manifolds, this induces a map between the tangent spaces of those manifolds:

Definition 2.2.6. Let $f \colon M \to N$ be a differentiable map between manifolds. The push forward of f at $p \in M$ is the map:

$$Df(p) \colon T_pM \to T_{f(p)}N,\ X_{[\gamma]} \mapsto X_{[f \circ \gamma]}$$

By the chain rule it is clear that this map is independent of the representative of γ and thus is well defined.

The push forward is linear:

Proposition 2.2.7. *Let $f\colon M \to M'$ be a differentiable map between manifolds and $p \in M$. Let (U, h, V) be a chart of M around p and (U', h', V') be a chart of M' around $f(p)$. The push forward of f at $p \in M$ is linear and its matrix is given by $J(h' \circ f \circ h^{-1})(h(p))$.*

Proof. Let $(u^i)_i$ and $(u'^i)_i$ be the local coordinate systems associated to the respectively unprimed and primed charts in the proposition and let $X_{[\gamma]}, X_\eta \in T_pM$ and $\lambda, \mu \in \mathbb{R}$. There exists a$\psi \in K_p$ such that $[\psi] = \lambda[\gamma] + \mu[\eta]$. For a function $g \in F'_{f(p)}$ it holds:

$$\begin{aligned} Df(p)(\lambda X_{[\gamma]} + \mu X_{[\eta]})(g) &= Df(p)(X_{[\psi]})(g) = X_{[f\circ\psi]}g = \frac{dg \circ f \circ \psi}{dt}(0) \\ &= \frac{\partial g}{\partial u'^j}(f(p))u'^j(\frac{\partial f}{\partial u^i}(p))u^i(\frac{d\psi}{dt}(0)) \\ &= \frac{\partial g}{\partial u'^j}(f(p))u'^j(\frac{\partial f}{\partial u^i}(p))u^i(\lambda\frac{d\gamma}{dt}(0) + \mu\frac{d\eta}{dt}(0)) \\ &= \lambda\frac{\partial g}{\partial u'^j}(f(p))u'^j(\frac{\partial f}{\partial u^i}(p))u^i(\frac{d\gamma}{dt}(0)) \\ &\quad + \mu\frac{\partial g}{\partial u'^j}(f(p))u'^j(\frac{\partial f}{\partial u^i}(p))u^i(\frac{d\eta}{dt}(0)) \\ &= \lambda\frac{dg \circ f \circ \gamma}{dt}(0) + \mu\frac{dg \circ f \circ \eta}{dt}(0) \\ &= \lambda Df(p)(X_{[\gamma]})(g) + \mu Df(p)(X_{[\eta]})(g) \end{aligned}$$

Thus the tangent map is linear. Furthermore it holds:

$$u'^j(\frac{\partial f}{\partial u^i}(p)) = \frac{\partial x'^j \circ h' \circ f \circ h^{-1}}{\partial x^i}(h(p)) = J(h' \circ f \circ h^{-1})(h(p))$$

So from the previous equations it also follows that $J(h' \circ f \circ h^{-1})(h(p))$ is the matrix representation of $Df(p)$. □

Since the tangent map is linear, it has a dual, called the pullback:

Definition 2.2.8. Let $f\colon M \to N$ be a differentiable map between manifolds. The pull back of f at $p \in M$ is the map $f^*(p)\colon T^*_{f(p)}N \to T^*_pM$ dual to $Df(p)$. It is thus given by $f^*(p)(\omega)(X) = \omega(Df(p)(X))$ for $\omega \in T^*_{f(p)}N$ and $X \in T_pM$.

A tangent space is a space assigned to a point of the manifold. This can be done for all points on the manifold. If one tangent vector is picked from every tangent space of the manifold, one gets a vector field:

Definition 2.2.9. Let M be a manifold. A vector field X on M is a map $X\colon M \to \bigsqcup_{p\in M} T_pM$, where $\bigsqcup$ denotes the disjoint union, such that $X_p := X(p) \in T_pM$.

Let $f\colon M \to \mathbb{R}$ be a differentiable function and X a vector field on M. Then Xf denotes the function on the manifold given by $(Xf)(p) = X_pf$.

Definition 2.2.10. Let M be a manifold and X a vector field on M, then X is called differentiable if Xf is differentiable for every differentiable $f\colon M \to \mathbb{R}$. The set of all differentiable vector fields on M is denoted by $\mathfrak{X}(M)$ and forms a vector space under pointwise operations.

With this in mind the following map can be defined:

Definition 2.2.11. Let M be a manifold. The map $[\cdot,\cdot]\colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ given by $[X,Y] = XY - YX$ is called the Lie-bracket.

Since Yf is a differentiable function on M, $(XY)f = X(Yf)$ is well defined. XY is thus again a vector field. Since $\mathfrak{X}(M)$ is a vector space, it follows that the Lie-bracket is well defined. With the Lie-bracket as multiplication, $\mathfrak{X}(M)$ becomes a non-commutative algebra. Furthermore, if $(u^i)_i$ is a local coordinate system and $X_p = X^i \frac{\partial}{\partial u^i}|_p$ and $Y_p = Y^j \frac{\partial}{\partial u^j}|_p$ for all p, then the coordinate expression of the Lie-bracket is, for all p:

$$\begin{aligned}
[X,Y]|_p &= X_pY_p - Y_pX_p = X^i \frac{\partial}{\partial u^i}|_p Y^j \frac{\partial}{\partial u^j}|_p - Y^j \frac{\partial}{\partial u^j}|_p X^i \frac{\partial}{\partial u^i}|_p \\
&= X^i \frac{\partial Y^j}{\partial u^i}|_p \frac{\partial}{\partial u^j}|_p - Y^j \frac{\partial X^i}{\partial u^j}|_p \frac{\partial}{\partial u^i}|_p = (X^j \frac{\partial Y^i}{\partial u^j}|_p - Y^j \frac{\partial X^i}{\partial u^j}|_p) \frac{\partial}{\partial u^i}|_p
\end{aligned}$$

Summarizing the tangent space essentially encodes in which direction the manifolds extends onward. Globally points on a manifolds cannot be compared, but the tangent space makes it possible to do this on an infinitesimally close range. Later on in this chapter this comparison will be made clearer.

2.3 Tensors

The Laws of Physics shouldn't depend on the particular coordinate system chosen to describe the physical system. This means that a coordinate change puts a restriction on the transformation behaviour the objects considered in the theory. Objects that obey these transformation restrictions are called tensor fields. Before tensor fields can be discussed however, tensors need to be introduced:

Definition 2.3.1. Let V be a finite dimensional vector space. A type-(k,l) tensor T on V is a multilinear function:

$$T\colon (V^*)^k \times V^l \to \mathbb{R}$$

with the V^* and V not necesarilly in this order. The set of all type-(k,l) tensors is called $\boldsymbol{T}^k_l(V)$ and forms a vector space under pointwise operations. A general tensor is an element of the direct sum $\boldsymbol{T}(V) = \bigoplus_{k,l} \boldsymbol{T}^k_l(V)$ of the vector spaces of tensors of all types.

Some examples:

Example 2.3.2. By definition a type-$(0,0)$ tensor T is just a number $T \in \mathbb{R}$. This is also called a scalar.

A type-$(1,0)$ tensor T is a linear function $V^* \to \mathbb{R}$. Since the (vector) space of linear functions $V^* \to \mathbb{R}$ is isomorphic to V, if V is finite dimensional, T is just a vector in V.

On the other hand a type-$(0,1)$ tensor T is a linear function $V \to \mathbb{R}$, which is by definition a covector in V^*.

A type-$(1,k)$ tensor T is a multilinear function $V^* \times V^k \to \mathbb{R}$. The (vector) space of linear functions $V^* \times V^k \to \mathbb{R}$ is isomorphic to the (vector) space of linear functions $V^k \to V$. So T is a multilinear map from V^k to V. In particular, a type-$(1,1)$ tensor is just a linear map on V, which is often represented by a matrix.

In analogy to the matrix representation of tensors of type $(1,1)$, tensors of other types can as well be represented by a set of numbers. To do this, let $(e_1, \dots, e_n)$ be a basis for V and $(\epsilon^1, \dots, \epsilon^n)$ be a basis for V^*. Then, with respect to this basis, a type-(k,l) tensor T is uniquely determined by the set of numbers $T^{i_1 \dots i_k}{}_{j_1 \dots j_l} = T(\epsilon^{i_1}, \dots, \epsilon^{i_k}, e_{j_1}, \dots, e_{j_l})$, with $1 \leq i_a, j_b \leq n$ for all $1 \leq a \leq k$ and $1 \leq b \leq l$, due to its multilinearity. These numbers are called its components. Often tensors will be given only in terms of their components. To avoid ambiguity, it is customary to write the basis vectors of V with lower indices and the basis vectors of V^* with upper indices. In this way the upper indices on a tensor component always refer to covector entries and the lower indices refer to vector entries. The upper and lower indices are also called contravariant and covariant indices respectively. A tensor that only has upper or lower indices is also called a contravariant respectively covariant tensor. A vector for example is a contravariant tensor.

There exists a natural product on the set of tensors:

Definition 2.3.3. Let $S \in \boldsymbol{T}^k_l(V)$ and $T \in \boldsymbol{T}^r_s(V)$. Then the tensor product $S \otimes T$ of S and T is a tensor of type $(k+r, l+s)$ defined by:

$$(S \otimes T)(\alpha_1, \ldots, \alpha_{k+r}, a_1, \ldots, a_{l+s}) = S(\alpha_1, \ldots, \alpha_k, a_1, \ldots, a_l) \cdot T(\alpha_{k+1}, \ldots, \alpha_{k+r}, a_{l+1}, \ldots, a_{l+s})$$

where $\alpha_i \in V^*$ and $a_j \in V$ for all i and j. With this multiplication $\boldsymbol{T}(V)$ becomes a non-commutative algebra.

From this definition it is clear that every type-(k, l) tensor can be decomposed as the tensor product of k type-$(1, 0)$ tensors and l type-$(0, 1)$ tensors.

Using the Einstein summation convention, the tensor product also gives a new way to write tensors. Due to its multilinearity the evaluation of a tensor T of type (k, l) can be written as, for $\alpha_1, \ldots, \alpha_k \in V^*$ and $a_1, \ldots, a_l \in V$:

$$\begin{aligned}
&T(\alpha_1, \ldots, \alpha_k, a_1, \ldots, a_l) \\
&= T(\epsilon^{i_1}, \ldots, \epsilon^{i_k}, e_{j_1}, \ldots, e_{j_l})\alpha_{1,i_1} \cdot \ldots \cdot \alpha_{k,i_k} \cdot a_1^{j_1} \cdot \ldots \cdot a_l^{j_l} \\
&= T^{i_1 \ldots i_k}{}_{j_1 \ldots j_l} \cdot \alpha_{1,i_1} \cdot \ldots \cdot \alpha_{k,i_k} \cdot a_1^{j_1} \cdot \ldots \cdot a_l^{j_l} \\
&= T^{i_1 \ldots i_k}{}_{j_1 \ldots j_l} e_{i_1}(\alpha_1) \cdot \ldots \cdot e_{i_k}(\alpha_{k,i_k}) \cdot \epsilon^{j_1}(a_1) \cdot \ldots \cdot \epsilon^{j_l}(a_l) \\
&= T^{i_1 \ldots i_k}{}_{j_1 \ldots j_l}(e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes \epsilon^{j_1} \otimes \cdots \otimes \epsilon^{j_l})(\alpha_1, \ldots, \alpha_k, a_1, \ldots, a_l)
\end{aligned}$$

The tensor T can thus be written as $T = T^{i_1 \ldots i_k}{}_{j_1 \ldots j_l} e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes \epsilon^{j_1} \otimes \cdots \otimes \epsilon^{j_l}$. This shows that the tensors $e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes \epsilon^{j_1} \otimes \cdots \otimes \epsilon^{j_l}$ form a basis for $\boldsymbol{T}^r_s(V)$. If V is n-dimensional, this shows that $\boldsymbol{T}^k_l(V)$ is n^{k+l}-dimensional. Moreover the previous calculation shows that an evaluation of a tensor can be completely expressed in terms of the components of the tensors involved. It is easy to see that the same holds true for the scalar multiplication, addition and tensor product multiplication of tensors. This means that any tensor equation can be solved completely in terms of components.

The final basic operation on tensors often encountered is the contraction:

Definition 2.3.4. Let $T \in \boldsymbol{T}^k_l(V)$. The contraction C^a_b of the a-th upper index with the b-th lower index of T is the type-$(k-1,l-1)$ tensor:

$$\begin{aligned}C^a_b(T) =&\epsilon^{j_b}(e_{i_a})T^{i_1\dots i_a\dots i_k}{}_{j_1\dots j_b\dots j_l}e_{i_1}\otimes\dots\otimes e_{i_{a-1}}\otimes e_{i_{a+1}}\\ &\otimes\dots\otimes e_{i_k}\otimes\epsilon^{j_1}\otimes\dots\otimes\epsilon^{j_{b-1}}\otimes\epsilon^{j_{b+1}}\otimes\dots\otimes\epsilon^{j_l}\\ =&\delta^{j_b}_{i_a}T^{i_1\dots i_a\dots i_k}{}_{j_1\dots j_b\dots j_l}e_{i_1}\otimes\dots\otimes e_{i_{a-1}}\otimes e_{i_{a+1}}\\ &\otimes\dots\otimes e_{i_k}\otimes\epsilon^{j_1}\otimes\dots\otimes\epsilon^{j_{b-1}}\otimes\epsilon^{j_{b+1}}\otimes\dots\otimes\epsilon^{j_l}\\ =&T^{i_1\dots i_{a-1}xi_{a+1}\dots i_k}{}_{j_1\dots j_{b-1}xj_{b+1}\dots j_l}e_{i_1}\otimes\dots\otimes e_{i_{a-1}}\otimes e_{i_{a+1}}\\ &\otimes\dots\otimes e_{i_k}\otimes\epsilon^{j_1}\otimes\dots\otimes\epsilon^{j_{b-1}}\otimes\epsilon^{j_{b+1}}\otimes\dots\otimes\epsilon^{j_l}\end{aligned}$$

where δ^i_j is the Kronecker delta.

Since the contraction also has a component expression, this means that any tensor equation can be solved completely in terms of components.

If now the tangent space T_pM is chosen as vector space, then a tensor can be assigned to every point on the manifold:

Definition 2.3.5. Let M be a manifold. A tensor field T of type-(k,l) on M is a map $T\colon M\to\bigsqcup_{p\in M}\boldsymbol{T}^k_l(T_pM)$. By pointwise operations all tensor fields of type-(k,l) form a vector space. A general tensor field T is a map $T\colon M\to\bigsqcup_{p\in M}(\bigoplus_{k,l}\boldsymbol{T}^k_l(T_pM))$. With pointwise multiplication $\otimes$, the set of all tensor fields forms an algebra.

If one chooses a chart (U,h,V) of M, then, if $(u^i)_i$ is the associated local coordinate system, a tensor field T of type (k,l) on U can be written as $T = T^{i_1\dots i_k}{}_{j_1\dots j_l}(p)\frac{\partial}{\partial u_{i_1}}|_p\otimes\dots\otimes\frac{\partial}{\partial u_{i_k}}|_p\otimes(du^{j_1})_p\otimes\dots\otimes(du^{j_l})_p$, where $T^{i_1\dots i_k}{}_{j_1\dots j_l}$ is now a function on U.

In physics often a condition to check whether something is a tensor field is given in terms of a transformation rule:

Proposition 2.3.6. *Let M be a manifold, $p\in M$ and (U,h,V) and (U',h',V') two chart of M around p. Let $(u^i)_i$ be the local coordinate system defined by (U,h,V) and $(u'^i)_i$ the one defined by (U,h,V). Let T be a tensor field of type-(k,l) on $U\cap U'$. Then the components of T transform as:*

$$T'^{i_1\dots i_k}{}_{j_1\dots j_l}(p)=\frac{\partial u'^{i_1}}{\partial u^{a_1}}\cdot\dots\cdot\frac{\partial u'^{i_k}}{\partial u^{a_k}}\cdot\frac{\partial u^{b_1}}{\partial u'^{j_1}}\cdot\dots\cdot\frac{\partial u^{b_l}}{\partial u'^{j_l}}\cdot T^{a_1\dots a_k}{}_{b_1\dots b_l}(p)$$

Proof. The basis vectors of the tangent space associated to $(u'^i)_i$ are $\frac{\partial}{\partial u'^i}$. They transform as $\frac{\partial}{\partial u'^i}=(\frac{\partial u^j}{\partial u'^i})^j_i\frac{\partial}{\partial u^j}$. The associated dual basis then transforms as $du'^i=((\frac{\partial u^j}{\partial u'^i})^{-1})^i_j du^j=(\frac{\partial u'^i}{\partial u^j})^i_j du^j$. The desired property then follows from the multilinearity. □

Since tensor fields are maps, a differentiability condition can be formulated:

Definition 2.3.7. Let M be a manifold and $(u^i)_i$ a local coordinate system with coordinate neighbourhood U. A tensor field is called differentiable on U if its components with respect to $(u^i)_i$ are differentiable on U.

By the transformation property of tensor fields and the differentiability of the gluing maps, this notion is independent from the chosen coordinate system.

Finally, a tensor is called symmetric respectively antisymmetric in the indices a and b if switching the indices a and b leaves the tensor component the same respectively flips the sign of the tensor component.

A tensor can be made symmetric as follows:

Definition 2.3.8. Let T be a tensor (field) of type $(k, 1)$ with components $T^{i_1 \dots i_k}{}_{j_1}$, then the tensor (field) defined in components as follows is symmetric in the indices between the brackets:

$$T^{(i_1 \dots i_n) i_{n+1} \dots i_k}{}_{j_1} = \frac{1}{n!} \sum_{\sigma \in S_n} T^{\sigma(i_1 \dots i_n) i_{n+1} \dots i_k}{}_{j_1}$$

where S_n is the permutation group of n elements. The same definition is made for lower indices.

A tensor can also be made antisymmetric:

Definition 2.3.9. Let T be a tensor (field) of type $(k, 1)$ with components $T^{i_1 \dots i_k}{}_{j_1}$, then the tensor (field) defined in components as follows is antisymmetric in the indices between the square brackets:

$$T^{[i_1 \dots i_n] i_{n+1} \dots i_k}{}_{j_1} = \frac{1}{n!} \sum_{\sigma \in S_n} sign(\sigma) T^{\sigma(i_1 \dots i_n) i_{n+1} \dots i_k}{}_{j_1}$$

where S_n is the permutation group of n elements. The same definition is made for lower indices.

2.4 Vector bundles

In the previous two sections constructions were given for defining a vector space related to a point on a manifold, the tangent space and the space of tensors. When the union of all these vector spaces is considered, an object is created that turns out to have a manifold structure. This object is called a vector bundle.

Definition 2.4.1. Let M be an n-dimensional manifold. A differentiable vector bundle of rank r over M is a triple $(E, \pi, \mathcal{A}^E)$ such that:

1. E is a $(n+r)$-dimensional differentiable manifold
2. $\pi\colon E \to M$ is a differentiable map
3. $\mathcal{A}^E = \{(U_i^E, h_i^E, U_i \times \mathbb{R}^r)\colon i \in I\}$, where I is an index set, is a bundle atlas for E, meaning:
   (a) $\{U_i\colon i \in I\}$ is an open cover for M, the U_i^Ecover E and $U_i^E = \pi^{-1}(U_i)$ for all i
   (b) $h_i^E\colon U_i^E \to U_i \times \mathbb{R}^r$ is a diffeomorphism for all i
   (c) If $p_i\colon U_i \times \mathbb{R}^r \to U_i$ is the projection on the first factor, then $\pi|_{U_i^E} = p_i \circ h_i^E$ for all i
   (d) The maps $h_i^E|_{\pi^{-1}(p)} \circ (h_j^E|_{\pi^{-1}(p)})^{-1}\colon \{p\} \times \mathbb{R}^r \to \{p\} \times \mathbb{R}^r$ are linear on the second factor for all $p \in U_i \cap U_j$ and all i and j

To see that the map $h_i^E|_{\pi^{-1}(p)} \circ (h_j^E|_{\pi^{-1}(p)})^{-1}$ is well defined, let $p \in U_i \subset M$ for some i. Then $\pi^{-1}(p) \subset U_i^E$, so :

$$p = \pi|_{U_i^E}(\pi^{-1}(p)) = p_i(h_i^E(\pi^{-1}(p)))$$

Since p_i is the projection on the first factor, it must hold that $h_i^E(\pi^{-1}(p)) \subset \{p\} \times \mathbb{R}^r$. So the maps $h_i^E|_{\pi^{-1}(p)} \circ (h_j^E|_{\pi^{-1}(p)})^{-1}$ are well defined. Moreover, since the h_i^E are diffeomorphisms, it follows that the maps $h_i^E|_{\pi^{-1}(p)}$ from $\pi^{-1}(p)$ to $\{p\} \times \mathbb{R}^r$ are bijections. Hence the maps $h_i^E|_{\pi^{-1}(p)} \circ (h_j^E|_{\pi^{-1}(p)})^{-1}$ are isomorphisms.

Often the bundle atlas is implied and a differentiable vector bundle $(E, \pi, \mathcal{A}^E)$ over a manifold M is denoted as $\pi\colon E \to M$. If the projection is clear, then also E itself is called a differentiable vector bundle. The set $E_p := \pi^{-1}(p)$ is also called the fibre of E over p. The bijection $h_i^E|_{E_p}$ defines an r-dimensional real vector space structure on E_p. This justifies the name vector bundle. The vector bundle E is basically an addition of a vector space to each point on the base manifold M.

It is very useful to define maps $M \to E$ as follows:

Definition 2.4.2. Let $\pi\colon E \to M$ differentiable vector bundle. A section s is a map $s\colon M \to E$ such that $s(p) \in E_p$ for all $p \in M$. If s is differentiable, then s is called a differentiable section. The space of all differentiable sections on E is denoted by $\Gamma(E)$. It is a vector space under pointwise operations.

Since a vector bundle E is in particular a manifold, it carries all the structure of manifold. In particular it has a tangent space. This leads to the following definition:

Definition 2.4.3. Let $\pi\colon E \to M$ differentiable vector bundle of rank r and let $v \in E$. $E_{\pi(v)}$ has a natural r-dimensional vector space structure and using the subspace topology it is a second countable Hausdorff topological space. This means that $E_{\pi(v)}$ is a manifold by Example 2.1.6. The vertical space V_vE of E at v is now defined as $V_vE = T_vE_{\pi(v)}$.

The vertical space is part of the tangent space of E:

Proposition 2.4.4. *Let $\pi\colon E \to M$ be a differentiable vector bundle of rank r and let $v \in E$. The vertical space V_vE is a linear subspace of tangent space T_vE.*

Proof. Let $\mathcal{A}^E = \{(U_i^E, h_i^E, U_i \times \mathbb{R}^r)\colon i \in I\}$ be a bundle atlas for E and $\mathcal{A} = \{(U_i, h_i, V_i)\colon i \in I\}$ be an atlas for M. Then it is clear that $\{(U_i^E, (h_i \times id_{\mathbb{R}^n}) \circ h_i^E, V_i \times \mathbb{R}^r)\colon i \in I\}$ is a differentiable atlas for E, (U_i^E is open in E since π is continuous.) Take E to be a manifold using differentiable structure associated to this atlas. Let $v \in E$ and let $(U^E, (h \times id_{\mathbb{R}^n}) \circ h^E, V \times \mathbb{R}^r)$ be a chart around v. Let $(u^i)_i$ be the local coordinate system defined by this chart. It holds $h^E(E_{\pi(v)} \cap U^E) \subset \{\pi(v)\} \times \mathbb{R}^r$, since $\pi(v) = \pi|_{U^E}(E_{\pi(v)} \cap U^E) = p \circ h^E(E_{\pi(v)} \cap U^E)$. This gives:

$$\begin{aligned} u^i(E_{\pi(v)} \cap U^E) &= (x^i \circ (h \times id_{\mathbb{R}^n}) \circ h^E)(E_{\pi(v)} \cap U^E) \\ &\subset (x^i \circ (h \times id_{\mathbb{R}^n}))(\{\pi(v)\} \times \mathbb{R}^r) = x^i(\{h(\pi(v))\} \times \mathbb{R}^r) \end{aligned}$$

which shows that on $E_{\pi(v)}$ only the last r coordinate functions u^i vary. Moreover $E_{\pi(v)}$ is a r-dimensional vector space, so by Example 2.2.4 $T_vE_{\pi(v)}$ is also a r-dimensional. This means that $V_vE = T_vE_{\pi(v)}$ has as a basis the r vectors $\frac{\partial}{\partial u^{n+1}}|_v$ to $\frac{\partial}{\partial u^{n+r}}|_v$. Since T_vE has $(\frac{\partial}{\partial u^i}|_v)_{i=1}^{n+r}$ as a basis, this shows that V_vE is a r-dimensional linear subspace of tangent space T_vE. $\square$

Furthermore one can define a complement of vertical space:

Definition 2.4.5. Let $\pi\colon E \to M$ be a differentiable vector bundle and let $v \in E$. A horizontal space H_vE is a linear subspace of T_vE such that $T_vE = V_vE \oplus H_vE$.

Let M be an n-dimensional manifold. Here are some important examples for vector bundles over M:

Example 2.4.6. *The trivial bundle*

Consider $E = M \times \mathbb{R}^r$. If E is given the product topology, it is second countable and Hausdorff, since M and $\mathbb{R}^r$ are. There is a natural projection $\pi\colon E \to M$ onto the first factor. Now take a differentiable atlas $\{(U_i, h_i, V_i)\colon i \in I\}$ for M. Then consider the set $\{(U_i \times \mathbb{R}^r, h_i \times id_{\mathbb{R}^r}, V_i \times \mathbb{R}^r)\colon i \in I\}$. It is clear that this is an $(n+r)$-dimensional differentiable atlas for E. With the associated differentiable structure it is also clear that π becomes differentiable. Finally take $\mathcal{A}^E = \{(E, id_{M\times\mathbb{R}^r}, M \times \mathbb{R}^r)\}$. It is clear that this is a bundle atlas, hence $(E, \pi, \mathcal{A}^E)$ is a vector bundle. E is called the trivial bundle.

Example 2.4.7. *The tangent bundle*

(This construction is an extension of the one shown in Lübke [2] and Tu [3].) Consider the disjoint union $TM = \bigsqcup_{p\in M} T_pM$. There exists a natural projection $\pi\colon TM \to M$, namely for all $v \in TM$ there exists a unique $p \in M$ such that $v \in T_pM$. π is then given by $\pi(v) = p$. Now take a differentiable atlas $\mathcal{A} = \{(U_i, h_i, V_i)\colon i \in I\}$ for M. Then this defines a local coordinate system $(u_i^1, \dots u_i^n)$ on U_i for each i. Define for all i the sets $U_i^{TM} := \pi^{-1}(U_i) = \bigsqcup_{p\in U_i} T_pM$, the map $g\colon M \times \mathbb{R}^n \to TM$ given by $g(p, x) = x^k \frac{\partial}{\partial u_i^k}|_p$, where x^k is the k-th component of x in $\mathbb{R}^n$, and its restrictions $g_i := g|_{U_i\times\mathbb{R}^n}\colon U_i \times \mathbb{R}^n \to g(U_i \times \mathbb{R}^n) = U_i^{TM}$. Because $\mathcal{A}$ is an atlas for M, $\{U_i\colon i \in I\}$ is an open cover for M and by construction it holds $TM = \bigcup_i U_i^{TM}$. Given the fact that T_pM is an n-dimensional vector space with basis $(\frac{\partial}{\partial u_i^k}|_p)_k$ (see Theorem 2.2.3), it follows that g, and thus also the g_i, are bijective. By construction of U_i^{TM}, it holds $\pi \circ g_i = p_i$, where $p_i\colon U_i \times \mathbb{R}^r \to U_i$ is the projection on the first factor. This gives $\pi|_{U_i^{TM}} = p_i \circ g_i^{-1}$.

Topology of the tangent bundle:

Now a topology on TM is going to be created. Take for all i the product topology on $U_i \times \mathbb{R}^n$. Since the g_i are bijective, this induces a topology on U_i^{TM} by setting $X \subset U_i^{TM}$ open if and only if $g_i^{-1}(X)$ is open in $U_i \times \mathbb{R}^n$. Now suppose X_i is open in U_i^{TM} and X_j is open in U_j^{TM}. Then $X_i \cap U_j^{TM}$ and $X_j \cap U_i^{TM}$ are open in $U_i^{TM} \cap U_j^{TM}$. It holds $X_i \cap X_j \subset U_i^{TM} \cap U_j^{TM}$, so $X_i \cap X_j = (X_i \cap X_j) \cap (U_i^{TM} \cap U_j^{TM}) = (X_i \cap U_j^{TM}) \cap (X_j \cap U_i^{TM})$. This means that $X_i \cap X_j$ is open in $U_i^{TM} \cap U_j^{TM}$. From this it follows that the set $B_{TM} = \{X \subset TM | \exists i\colon X \subset U_i^{TM} \text{ open}\}$ is a topological basis. Take on TM the topology generated by B_{TM}. The maps g_i are now by construction homeomorphisms. Now suppose another compatible atlas

$\mathcal{A}' = \{(U_i', h_i', V_i') : i \in I'\}$ was chosen, which gave rise to another basis B'_{TM}. Then for $p \in U_i \cap U_j'$ it holds:

$$\begin{aligned} g_i(p,y) &= y^k \frac{\partial}{\partial u_i^k}|_p = y^k \frac{\partial u'^l_j}{\partial u_i^k}(h_i(p)) \cdot \frac{\partial}{\partial u'^l_j}|_p \\ &= y^k \frac{\partial x^l \circ h_j' \circ h_i^{-1}}{\partial x^k}(h_i(p)) \cdot \frac{\partial}{\partial u'^l_j}|_p = g_j'(p, J(h_j' \circ h_i^{-1}) \cdot y) \end{aligned}$$

where the x^i denote the coordinate functions on $\mathbb{R}^n$ and $J(f)$ denotes the Jacobian of f. This thus gives $g'^{-1}_j \circ g_i(p,y) = (p, J(h_j' \circ h_i^{-1}) \cdot y)$. Since $h_j' \circ h_i^{-1}$ is differentiable, because of the compatibility of the atlases, $J(h_j' \circ h_i^{-1})$ is an isomorphism and thus a homeomorphism. This implies that $g'^{-1}_j \circ g_i = id_M \times J(h_j' \circ h_i^{-1})$ is a homeomorphism, since $U_i \times \mathbb{R}^n$ and $U_j' \times \mathbb{R}^n$ have the product topology. Moreover, by the same token it is then found that $g_j^{-1} \circ g_i = id_M \times J(h_j \circ h_i^{-1})$ (without prime) is linear on the second factor for all $p \in U_i \cap U_j$ and all i and j. Going back to the primed case, the map $g_j' \circ (g'^{-1}_j \circ g_i) \circ g_i^{-1}$, from U_i^{TM} with topology induced by B_{TM} to U'^{TM}_j with topology induced by B'_{TM}, is then a homeomorphism. This implies that the topologies induced by B_{TM} and B'_{TM} on TM are the same. The topology is thus independent from the chosen atlas. This also means that g is a homeomorphism.

Topological properties of the tangent bundle:

First note that if $Y \subset M$ is open, then the triples $(Y \cap U_i, h_i|_{Y \cap U_i}, h_i(Y \cap U_i))$ are charts for M for all i and they are compatible with the atlas $\mathcal{A}$ for M. Hence the differentiable structure contains all such charts for M. Now take $p, q \in TM$ with $p \neq q$. Because M is Hausdorff, if $\pi(p) \neq \pi(q)$, there exist open $P, Q \subset M$ such that $\pi(p) \in P$, $\pi(q) \in Q$ and $P \cap Q = \emptyset$. Moreover there are$i$ and j such that $\pi(p) \in U_i$ and $\pi(q) \in U_j$. Therefore the charts given by the restriction to the sets $P \cap U_i =: U_p$ and $Q \cap U_j =: U_q$ are part of the differentiable structure of M. Since the g_i are homeomorphisms, $U_p^{TM} = g_p(U_p \times \mathbb{R}^n)$ and $U_q^{TM} = g_q(U_q \times \mathbb{R}^n)$ are open in TM, it holds $p \in U_p^{TM}$ and $q \in U_q^{TM}$ and, since g is bijective, $U_p^{TM} \cap U_q^{TM} = \emptyset$. Because the topology is independent from the chosen atlas, this means that TM is Hausdorff.

Since M is second countable, it has a countable basis B_M. This means that for all i and $p \in U_i$ a $B_{p,i} \in B_M$ can be picked such that $p \in B_{p,i} \subset U_i$. The set $\tilde{B}_M$ of these $B_{p,i}$ forms a subset of B_M that covers M and hence is a countable basis for M. Moreover, since each $B_{p,i}$ is open, it forms a chart as a restriction of the chart of U_i. The set of all these charts then forms a compatible atlas. M thus has a countable basis $\{U_b\}_b$ consisting of coordinate charts. Since M is second countable, the U_b are second countable and, because also $\mathbb{R}^n$ is second countable, the $U_b \times \mathbb{R}^n$ are second countable. Since the g_b are homeomorphisms, also the U_b^{TM} are second countable. Therefore the basis of TM given by the union of the bases of the U_b^{TM} is countable. Since the topology is independent from the chosen atlas,TM is second countable.

Manifold structure of the tangent bundle:

Now TM has a topology, a manifold structure can be defined. For this purpose, define $h_i^{TM} := (h_i \times id_{\mathbb{R}^n}) \circ g_i^{-1} \colon U_i^{TM} \to V_i \times \mathbb{R}^n$ for all i. Since h_i and g_i are homeomorphisms, the h_i^{TM} are also all homeomorphisms. For all i and j and all $(x, y) \in h_i^{TM}(U_i^{TM} \cap U_j^{TM})$ it holds:

$$\begin{aligned} h_j^{TM} \circ (h_i^{TM})^{-1}(x,y) &= ((h_j \times id_{\mathbb{R}^n}) \circ g_j^{-1} \circ g_i \circ (h_i^{-1} \times id_{\mathbb{R}^n}))(x,y) \\ &= (h_j \times id_{\mathbb{R}^n})(h_i^{-1}(x), J(h_j \circ h_i^{-1}) \cdot y) \\ &= ((h_j \circ h_i^{-1})(x), J(h_j \circ h_i^{-1}) \cdot y) \end{aligned}$$

Since the gluing maps $h_j \circ h_i^{-1}$ for M are differentiable, it follows that the gluing maps $h_j^{TM} \circ (h_i^{TM})^{-1}$ for TM are also differentiable. This means that $\{(U_i^{TM}, h_i^{TM}, V_i \times \mathbb{R}^n) \colon i \in I\}$ is a differentiable atlas for TM and TM is thus a $2n$-dimensional differentiable manifold. If another compatible atlas $\mathcal{A}' = \{(U_i', h_i', V_i') \colon i \in I'\}$ for M was chosen, then in the same way as the $h_j^{TM} \circ (h_i^{TM})^{-1}$ are differentiable, the functions $h'^{TM}_j \circ (h_i^{TM})^{-1}$ are differentiable. This means that the atlas $\{(U'^{TM}_i, h'^{TM}_i, V_i' \times \mathbb{R}^n) \colon i \in I\}$ found for TM is compatible and the same differentiable structure would have been defined on TM. So the differentiable structure on TM is independent of the chosen atlas.

Using the trivial bundle structure on $U_i \times \mathbb{R}^n$, it is clear that with the defined manifold structure on TM the maps g_i become diffeomorphisms. The set $\{(U_i^{TM}, g_i^{-1}, U_i \times \mathbb{R}^n) \colon i \in I\}$ is thus a bundle atlas for TM. Finally, for all i and all $(x, y) \in h_i^{TM}(U_i^{TM} \cap U_j^{TM})$ it holds:

$$\begin{aligned} h_i \circ \pi \circ (h_i^{TM})^{-1} &= h_i \circ \pi \circ g_i \circ (h_i^{-1} \times id_{\mathbb{R}^n}) \\ &= h_i \circ p_i \circ (h_i^{-1} \times id_{\mathbb{R}^n}) = p_i \circ id_{V_i \times \mathbb{R}^n} \end{aligned}$$

Since the projection p_i is differentiable, it follows that π is differentiable. Thus it can be concluded that TM is a vector bundle of rank n over M. TM is called the tangent bundle.

Example 2.4.8. *The tensor bundle*

Consider the disjoint union $\boldsymbol{T}^k_l M = \bigsqcup_{p\in M} \boldsymbol{T}^k_l(T_pM)$. There exists a natural projection $\pi\colon \boldsymbol{T}^k_l M \to M$, namely for all $v \in \boldsymbol{T}^k_l M$ there exists a unique $p \in M$ such that $v \in \boldsymbol{T}^k_l(T_pM)$. π is then given by $\pi(v) = p$. Now take a differentiable atlas $\mathcal{A} = \{(U_i, h_i, V_i)\colon i \in I\}$ for M. Then this defines a local coordinate system $(u^1_i, \dots u^n_i)$ on U_i for each i. Define for all i the sets $U_i^{\boldsymbol{T}^k_l M} = \pi^{-1}(U_i) = \bigsqcup_{p\in U_i} \boldsymbol{T}^k_l(T_pM)$ and the maps $g_i\colon U_i \times \mathbb{R}^{n^{k+l}} \to U_i^{\boldsymbol{T}^k_l M}$ given by $g_i(p, x) = x_{a_1\dots a_k b_1\dots b_l} \frac{\partial}{\partial u_i^{a_1}}|_p \otimes \cdots \otimes \frac{\partial}{\partial u_i^{a_k}}|_p \otimes du_i^{b_1}|_p \otimes \cdots \otimes du_i^{b_l}|_p$, where, if x is represented as a $(k+l)$-dimensional array of numbers, $x_{a_1\dots a_k b_1\dots b_l}$ is the entry of x at place a_1 in the first dimension, a_2 in the second dimension, b_1 in the $(k+1)$-th dimension and so on. Given the fact that $\boldsymbol{T}^k_l(T_pM)$ is an n^{k+l}-dimensional vector space with basis given by the vectors $\frac{\partial}{\partial u_i^{a_1}}|_p \otimes \cdots \otimes \frac{\partial}{\partial u_i^{a_k}}|_p \otimes du_i^{b_1}|_p \otimes \cdots \otimes du_i^{b_l}|_p$ (see right after Definition 2.3.3), it follows that the g_i are bijective. Then, completely analogous to the case of TM, it is found that $\boldsymbol{T}^k_l M$ is a vector bundle of rank $n(k+l)$ over M. $\boldsymbol{T}^k_l M$ is called the tensor bundle of type (k, l).

Example 2.4.9. *The vertical and horizontal bundle*

Let $\pi\colon E \to M$ be a differentiable vector bundle of rank r with E having the differentiable structure as prescribed in the proof of Proposition 2.4.4. Now consider the disjoint union $VE = \bigsqcup_{p\in E} V_pE$. As the previous two examples show, if a vector space is added to every point on a manifold in such a way that a basis can be given in terms of the coordinate functions of a local coordinate system, then the structure obtains a differentiable vector bundle structure. Precisely this has been done in the proof of Proposition 2.4.4. This means that VE becomes a rank r vector bundle over E. If for all $p \in E$ a horizontal space H_pE is chosen, then in the same way $HE = \bigsqcup_{p\in E} H_pE$ becomes a vector bundle of rank n called a horizontal bundle, by definition of the horizontal space and the fact that T_pE has a basis in terms of coordinate functions on E.

This shows that (differentiable) vector fields and tensor fields can be reinterpreted as (differentiable) sections of the tangent bundle and tensor bundle respectively.

2.5 Differentiation & integration

There is no obvious way to take a derivative of a tensor field on a manifold. A tensor field namely gives a tensor on each point on a manifold, but these tensors are incomparable since they live in different vector spaces. To make the comparison, a way is needed to relate the vector spaces at neighbouring points on the manifold. To solve this problem the covariant derivative is called into being. It however needs additional structure to be defined in the form of a vector. There however turns out to be a way to define a notion of differentiability that not suffers from this problem called the exterior derivative.

To define the exterior derivative, the k-th exterior power $\Lambda^k V$ of the vector space V has to be defined:

Definition 2.5.1. Let V be a vector space. The sets $\Lambda^k V$ are defined inductively. Set $\Lambda^0 V = \mathbb{R}$ and $\Lambda^1 V = V$. Then define inductively the map $\wedge\colon \Lambda^k V \times \Lambda^l V \to \Lambda^{k+l} V$ for all $k, l \in \mathbb{N}$ as the alternating bilinear map given by $\wedge(v, w) = v \wedge w$, where for $k = 0$ it is defined as $\wedge(a, w) = a \wedge w = aw$ and for $l = 0$ as $\wedge(v, b) = v \wedge b = bv$. Then $\Lambda^k V$ is a vector space. The map $\wedge$ is called the wedge product. If V is n-dimensional and has a basis $(e_i)_{i=1}^n$, then $\Lambda^k V$ is $\binom{n}{k}$-dimensional and has a basis $(e_{i_1} \wedge \cdots \wedge e_{i_k})_{1 \leq i_1 \leq \cdots \leq i_k = n}$.

Since $\wedge$ is alternating, it holds $\Lambda^k V = \{0\}$ for all $k > n$. Furthermore $\Lambda^n V$ is 1-dimensional with basis vector $e_1 \wedge \cdots \wedge e_n$.

Now let M be an n-dimensional manifold. Take $V = T_p^* M$ and consider the set $\Lambda^k T^* M = \bigsqcup_{p \in M} \Lambda^k T_p^* M$. Since $T_p^* M$ has a basis in terms of the coordinate functions of a local coordinate system on M and $\Lambda^k T_p^* M$ has a basis in terms of the basis vectors of $T_p^* M$, this defines a rank $\binom{n}{k}$ vector bundle structure on $\Lambda^k T^* M$, analogous to Example 2.4.7. Then:

Definition 2.5.2. Let M be a manifold. The vector space of differentiable sections of $\Lambda^k T^* M$ is denoted as $\Omega^k M$. An element of $\Omega^k M$ is called a differential form of degree k, or just k-form in short.

Let $k > n$, since $\Lambda^k T^* M = \{0\}$, $\Omega^k M$ contains only one section, namely the zero section, thus $\Omega^k M = \{0\}$. Similarly, since $\Lambda^0 T^* M = \mathbb{R}$, $\Omega^0 M$ is just the space of differentiable $\mathbb{R}$-valued functions on M. Moreover, for $f\colon M \to \mathbb{R}$ a differentiable function and ω a k-form, the k-form $f\omega$ can be defined as $(f\omega)(p) = f(p)\omega(p) \in \Lambda^k T_p^* M$. The wedge product $\wedge$ can be extended to differential forms by pointwise application, thus if

$\omega \in \Omega^k M$ and $\eta \in \Omega^l M$ then $\omega \wedge \eta \in \Omega^{k+l} M$. With this as multiplication, $\Omega M = \bigoplus_{i=0}^n \Omega^i M$ becomes an algebra, called the exterior algebra of M.

The pull back can now be extended to ΩM:

Definition 2.5.3. Let $f : M \to N$ be a differentiable map between manifolds and $(u^i)_i$ a local coordinate system of N. Let $\omega \in \Omega^k N$ for some k, then ω can locally be written as $\omega = a \cdot du^{i_1} \wedge \cdots \wedge du^{i_k}$ for some function a on N. The pull back of ω by f is now defined as:

$$\begin{aligned} f^*(\omega)(p) &= f^*(a \cdot du^{i_1} \wedge \cdots \wedge du^{i_k})(p) \\ &= (a \circ f)(p) \cdot f^*(du^{i_1}|_{f(p)}) \wedge \cdots \wedge f^*(du^{i_k}|_{f(p)}) \in \Lambda^k T_p^* M \end{aligned}$$

The pull back $f^* \colon \Omega N \to \Omega M$ is then defined as the linear extension of this map.

Now the exterior derivative can be defined. Let M be a manifold and $f \colon M \to \mathbb{R}$ be a differentiable function, then the 1-form $df \in \Omega^1 M$ can be defined as $df(p)(X) = Xf$ for all$p \in M$ and $X \in T_p M$. Since f is differentiable, df is also differentiable and thus well defined. The 1-form df is called the total differential of f. Then:

Definition 2.5.4. Let M be a manifold. The exterior derivative d is a linear endomorphism on ΩM such that:

1. for $f \in \Omega^0 M$ a differentiable function df is the total derivative of f
2. for $\omega \in \Omega^k M$ and $\eta \in \Omega^l M$ it holds $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$
3. $d^2 = 0$

Let $(u^i)_i$ be some local coordinate system on M and write $\omega = f du^{i_1} \wedge \cdots \wedge du^{i_k}$for some differentiable function f. Then $\omega \in \Omega^k M$. Now it holds $d\omega = df \wedge du^{i_1} \wedge \cdots \wedge du^{i_k} + f d(du^{i_1} \wedge \cdots \wedge du^{i_k}) = df \wedge du^{i_1} \wedge \cdots \wedge du^{i_k}$ by properties 2 and 3, so $d\omega \in \Omega^{k+1} M$. Since d is linear and $\Omega^k M$ is a vector space, this implies $d(\Omega^k M) \subset \Omega^{k+1} M$ for all k.

It may feel a bit strange to call d a derivative, so here is an clarifying example:

Example 2.5.5. Let $f \colon M \to \mathbb{R}$ be a differentiable function, then $df(p)(\frac{\partial}{\partial u^i}|_p) = \frac{\partial f}{\partial u^i}(p) = \frac{\partial f \circ h^{-1}}{\partial x^i}(h(p))$ is the derivative of f in the direction of u^i, i.e. the derivative of $f \circ h^{-1} \colon V \to \mathbb{R}$ in the direction of x^i. This means that df can be thought of as the gradient of f. In particular if $f = u^j$ is taken, one gets $du^j(p)(\frac{\partial}{\partial u^i}|_p) = \frac{\partial u^j}{\partial u^i}(p) = \frac{\partial x^j}{\partial x^i}(h(p)) = \delta_i^j$. This clarifies the notation of $(du^i|_p)_i$ being the dual basis of $(\frac{\partial}{\partial u^i}|_p)_i$ of $T_p M$.

Furthermore, d commutes with the pullback. For a proof see Lübke [2].

The exterior derivative does not need any extra structure to be defined. The other type of derivative, the covariant derivative, does need that:

Definition 2.5.6. Let M be a manifold and $X \in T_pM$ a tangent vector for some $p \in M$. The covariant derivative ∇_X with respect to X is a map $\nabla_X \colon \Gamma(\boldsymbol{T}M) \to \Gamma(\boldsymbol{T}M)$ such that:

1. for f a differentiable function it holds $(\nabla_X f)(p) = Xf$
2. for Y a vector field $(\nabla_X Y)(p)$ is a vector linear in X and additive in Y
3. for T and S tensor fields it holds $(\nabla_X(T \otimes S))(p) = (\nabla_X(T) \otimes S + T \otimes \nabla_X(S))(p)$
4. ∇_X commutes with contractions

From this it follows that for a covector field ω, a vector field Y and a vector X, it holds:

$$\begin{aligned} X\omega(Y) &= \nabla_X(\omega(Y)) = \nabla_X(C_1^1(\omega_a Y^b)) = C_1^1(\nabla_X(\omega_a Y^b)) \\ &= C_1^1(\nabla_X(\omega_a)Y^b + \omega_a \nabla_X(Y^b)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y) \end{aligned}$$

Hence $\nabla_X \omega$ is a covector field given by $(\nabla_X \omega)(Y) = X\omega(Y) - \omega(\nabla_X Y)$. For $\lambda \in \mathbb{R}$ it also holds $\nabla_X(\lambda Y) = X(\lambda)Y + \lambda \nabla_X Y = \lambda \nabla_X Y$, since λ can be seen as a constant function on the manifold. It now follows that ∇_X is $\mathbb{R}$-linear. Moreover, since Xf is a function for f a function, $\nabla_X Y$ a vector field and $\nabla_X \omega$ a covector field, it follows that ∇_X preserves the tensor type. Furthermore, the symbol ∇ can be thought of as a covector acting on the vector X to give a function that acts as a derivative, since ∇_X is tensor type preserving. This is why ∇_X is called the covariant derivative along X.

Now let $(u^i)_i$ be some local coordinate system around p and write $X = X^i \frac{\partial}{\partial u^i}|_p$, $Y = Y^j \frac{\partial}{\partial u^j}|_p$and $\omega = \omega_j du^j(p)$ for the vector X, the vector field Y and the covector field ω (the Y^j and ω_j are thus functions). By applying the properties one then finds:

$$\begin{aligned} \nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial u^i}|_p}(Y^j \frac{\partial}{\partial u^j}|_p) = X^i(\nabla_{\frac{\partial}{\partial u^i}|_p} Y^j)\frac{\partial}{\partial u^j}|_p + X^i Y^j \nabla_{\frac{\partial}{\partial u^i}|_p}\frac{\partial}{\partial u^j}|_p \quad (2.5.1) \\ &= X^i \frac{\partial Y^j}{\partial u^i}(p)\frac{\partial}{\partial u^j}|_p + X^i Y^j \Gamma^k_{ij}\frac{\partial}{\partial u^k}|_p = (X^i \frac{\partial Y^k}{\partial u^i}(p) + X^i Y^j \Gamma^k_{ij})\frac{\partial}{\partial u^k}|_p \quad (2.5.2) \end{aligned}$$

and thus also:

$$(\nabla_X\omega)(Y) = X\omega(Y) - \omega\nabla_X Y = X^i\frac{\partial \omega_j Y^j}{\partial u^i}(p) - \omega_k(X^i\frac{\partial Y^k}{\partial u^i}(p) + X^iY^j\Gamma^k_{ij})$$
$$= X^iY^j\frac{\partial \omega_j}{\partial u^i}(p) - \omega_k X^iY^j\Gamma^k_{ij}$$

where Γ^k_{ij} is the k-th component of the vector $\nabla_{\frac{\partial}{\partial u^i}|_p}\frac{\partial}{\partial u^j}|_p$. The Γ^k_{ij} are called connection coefficients. They determine the covariant derivative. There is thus not one unique covariant derivative. The components of ∇_X are thus $(\nabla_X)^k_j = X^i\frac{\partial}{\partial u^i}|_p\delta^k_j + X^i\Gamma^k_{ij} = X^i\nabla_i$, where $\nabla_i = \frac{\partial}{\partial u^i}|_p\delta^k_j + \Gamma^k_{ij}$, in the case of a vector entry and $(\nabla_X)^k_j = X^i\frac{\partial}{\partial u^i}|_p\delta^k_j - X^i\Gamma^k_{ij} = X^i\nabla_i$, where $\nabla_i = \frac{\partial}{\partial u^i}|_p\delta^k_j - \Gamma^k_{ij}$, in the case of a covector entry. Using the decomposition of a tensor field in tensor fields of type $(1,0)$ and $(0,1)$ and the product rule for the covariant derivative gives the components of ∇_X for a general tensor field entry. For a type-$(1,1)$ tensor field T, with decomposition $T = A\otimes B$ with $A \in \Gamma(\boldsymbol{T}^1_0 M)$ and $B \in \Gamma(\boldsymbol{T}^0_1 M)$, e.g. it holds in components:

$$(\nabla_X T)^a{}_b = (\nabla_X(A\otimes B))^a{}_b = ((\nabla_X A)\otimes B + A\otimes\nabla_X B)^a{}_b$$
$$= B_b(\nabla_X)^a_j A^j + A^a(\nabla_X)^j_b B_j$$
$$= B_b(X^i\frac{\partial}{\partial u^i}|_p\delta^a_j A^j + X^i\Gamma^a_{ij}A^j) + A^a(X^i\frac{\partial}{\partial u^i}|_p\delta^j_b B_j - X^i\Gamma^j_{ib}B_j)$$
$$= X^i(\frac{\partial A^a}{\partial u^i}(p)B_b + A^a\frac{\partial B_b}{\partial u^i}(p)) + X^i(\Gamma^a_{ij}A^jB_b - \Gamma^j_{ib}A^aB_j)$$
$$= X^i\frac{\partial T^a{}_b}{\partial u^i}(p) + X^i(\Gamma^a_{ij}T^j{}_b - \Gamma^j_{ib}T^a{}_j)$$

The connection coefficients don't form a tensor, even though they have indices! That's why they are called coefficients. For a proof of this see Carroll [4].

The connection coefficients can be thought of as being those coefficients that compensate for the non tensorial nature of the partial derivative is such a way that the resulting object is tensorial. A geometric interpretation of the covariant derivative ∇_X in the direction of X is then that it describes the change of the coefficients of the tensor in the direction of X whereby it compensates for the change of the basis by use of Γ.

Given a vector field, all three defined derivatives act the same on functions on the manifold, giving the partial derivative of the function in the direction of the vector field. The exterior derivative stands out because it is only defined on k-forms. The Lie-derivative and the covariant derivative on the other hand are defined on all types of tensor fields. The difference between those two is that the covariant derivative along a single vector is well defined, as opposed to the Lie-derivative. The Lie-derivative can only be taken with respect to a full vector field around the point of differentiation, because it also involves the change of the vector field along which the Lie-derivative is taken. Furthermore, the covariant derivative is the only one that is not unique.

Functions on manifolds can be integrated. The next part of this section is devoted to defining this integral and stating the Theorem of Stokes for it. The integral will however not be completely atlas independent. To make this precise the following definition is given:

Definition 2.5.7. Let $\mathcal{A} = \{(U_i, h_i, V_i) : i \in I\}$ be an atlas for a manifold M. $\mathcal{A}$ is called oriented if for all $i, j \in I$ and $p \in U_i \cap U_j$ the gluing maps satisfy $\det(J(h_j \circ h_i^{-1})(h_i(p))) > 0$. If such an atlas exists, M is called orientable.

Let M be a manifold with differentiable structure $\mathcal{S}$. Define the set of all compatible oriented atlases $\mathbb{A} = \{\mathcal{A} \subset \mathcal{S} : \mathcal{A} \text{ is oriented}\}$. Then define the relation $\sim$ on $\mathbb{A}$ by $\mathcal{A} \sim \mathcal{A}'$ if and only if $\mathcal{A} \cap \mathcal{A}' \in \mathbb{A}$. This relation is clearly reflexive and symmetric. By the chain rule and the multiplicative nature of the determinant it is also transitive, and thus an equivalence relation on $\mathbb{A}$. There are only two equivalence classes in $\mathbb{A}/\sim$. An orientation of M is a choice of equivalence class in $\mathbb{A}/\sim$. M is called oriented if it has an atlas that is an element of the chosen orientation.

The integral makes use of the charts of a manifold. The integral must therefore be cut into pieces along the charts. Since charts in general overlap, something is needed to not overcount these regions:

Definition 2.5.8. Let X be a topological space with open cover $\mathcal{U} = \{U_i : i \in I\}$. A partition of unity of X subordinate to $\mathcal{U}$ is a set of continuous functions $\{\tau_i : X \to [0,1] | i \in I\}$ such that:

1. $supp(\tau_i) \subset U_i$ for all i
2. for every $p \in X$ there exists an open neighbourhood U of p in X for which $supp(\tau_i) \cap U \neq \emptyset$ holds for at most finitely many i
3. $\sum_{i \in I} \tau_i(p) = 1$ for all $p \in X$

If X is a manifold, a partition of unity is called differentiable if the τ_i are differentiable. Such a partition of unity always exists for a manifold:

Lemma 2.5.9. *If M is a paracompact manifold with open cover $\mathcal{U} = \{U_i : i \in I\}$, then there exists a partition of unity subordinate to this cover. Moreover, if $K \subset M$ is compact, then there exists a partition of unity subordinate to a finite open cover of a subset $U \subset M$, such that $K \subset U$, given by charts of M.*

Proof. For the first statement see Kobayashi and Nomizu [1] and for the second statement see Lübke [2]. □

Now let $\mathcal{A} = \{(U_i, h_i, V_i) : i \in I\}$ be an atlas for the n-dimensional manifold M and let ω be an n-form. Then locally ω can be written as $\omega|_{U_i} = \omega_i du_i^1 \wedge \cdots \wedge du_i^n$, where $(u_i^j)_j$ is the local coordinate system defined by (U_i, h_i, V_i). Locally ω can then be pulled back to an n-form on V_i as:

$$\begin{aligned}(h_i^{-1})^*(\omega|_{U_i}) &= (\omega_i \circ h_i^{-1})(h_i^{-1})^*(du_i^1) \wedge \cdots \wedge (h_i^{-1})^*(du_i^n)\\ &= (\omega_i \circ h_i^{-1}) dx_i^1 \wedge \cdots \wedge dx_i^n\end{aligned}$$

since

$$(h_i^{-1})^*(du_i^j)(p)(X) = du_i^j(p)(Dh_i^{-1}(h_i(p))X) = du_i^j(p)(IX) = dx_i^j(p)(X)$$

The integral of ω is then defined as:

Definition 2.5.10. Let M be an oriented n-dimensional manifold with oriented atlas $\mathcal{A} = \{(U_i, h_i, V_i) : i \in I\}$, $A \subset M$ open and $\omega \in \Omega^n M$ an n-form. Let $\{U_j : j \in J \subset I\}$ be an open cover of $supp(\omega)$ and $\{\tau_j : j \in J\}$ a partition of unity subordinate to this cover. The integral of ω over M is then defined as:

$$\begin{aligned}\int_A \omega &= \sum_{j \in J} \int_{h_j(U_j \cap A)} (h_j^{-1})^*(\tau_j \cdot \omega|_{U_j})\\ &= \sum_{j \in J} \int_{h_j(U_j \cap A)} ((\tau_j \cdot \omega_j) \circ h_j^{-1}) dx_j^1 \wedge \cdots \wedge dx_j^n\\ &= \sum_{j \in J} \int_{h_j(U_j \cap A)} ((\tau_j \cdot \omega_j) \circ h_j^{-1}) d^n x^j\end{aligned}$$

where the last integral is just the standard Lebesgue integral.

This definition is independent of the chosen atlas and partition of unity, hence this integral is well defined. For a proof see Lübke [2]. Moreover, if ω has compact support in M, the integral of ω is finite. This follows from Lemma 2.5.9. Furthermore, this integral is also well defined for non-differentiable n-forms. The only real requirement is that the functions ω_i are Lebesgue integrable.

Often one also wants to define the integral of a function on a manifold. Since the product of a function and an n-form is again an n-form, this can be done by choosing a particular n-form.

Definition 2.5.11. Let M be an oriented n-dimensional manifold. A volume form on M is a particular choice of $Vol \in \Omega^n M$ such that it differs from the n-forms $du^1 \wedge \cdots \wedge du^n$, defined by the charts, only by a positive function. If f is a function on M its integral over an open $A \subset M$ is then defined as:

$$\int_A f = \int_A f \cdot Vol$$

Note that in general there is no natural choice for a volume form. Since $\Lambda^n V$, for V an n-dimensional vector space, is one dimensional, different choices of volume forms differ only by a function. In particular this defines the volume $Vol(M)$ of the manifold M, if it is compact, as:

$$Vol(M) = \int_M Vol$$

To state Stokes Theorem, the boundary of a subset of a manifold has to be defined:

Definition 2.5.12. Let M be a manifold, A an open subset of M and ∂A the topological boundary of A in M, then A is said to have a smooth boundary if for every $p \in \partial A$ there exists a chart (U, h, V) for M around p such that $h(U \cap A) = \{(x^1, \ldots, x^n) \in V \colon x^1 < 0\}$.

Lemma 2.5.13. *Let M be an n-dimensional manifold and A an open subset of M with smooth boundary. Then ∂A is a $(n-1)$-dimensional submanifold of M. Moreover, an orientation of M induces an orientation of ∂A.*

Proof. Take charts as given in Definition 2.5.12. Suppose now that there exists a $p \in \partial A \cap U$ such that $h(p) \in \{(x^1, \ldots, x^n) \in V \colon x^1 > 0\}$. Since h is an homeomorphism, U is open and V is Hausdorff, there exists an open neighbourhood $P \subset U$ around p such that $h(P) \subset \{(x^1, \ldots, x^n) \in V \colon x^1 > 0\}$. On the other hand $P \cap A \neq \emptyset$, since $\partial A = \overline{A} \backslash A$, by definition of the closure. This contradicts the assumption that the chart is of the form given in Definition 2.5.12. It thus holds $h(\partial A \cap U) = \{(x^1, \dots, x^n) \in V : x^1 = 0\}$. By definition ∂A is thus an $(n-1)$-dimensional submanifold of M (the coordinates can be rearranged such that $x^n = 0$ in stead of $x^1 = 0$).

For the proof of the fact that M induces an orientation on ∂A see Lübke [2]. □

Now the Theorem of Stokes can be stated:

Theorem 2.5.14. *Let M be an oriented n-dimensional manifold and A an open subset of M with smooth boundary ∂A. Fix on ∂A the orientation induced by M. Let $\iota\colon \partial A \to M$ denote the inclusion map. Then for every $\omega \in \Omega^{n-1}M$ with compact support it holds:*

$$\int_A d\omega = \int_{\partial A} \iota^*(\omega)$$

Proof. See Lübke [2]. □

In terms of covariant derivatives Stokes's Theorem can be restated as:

Theorem 2.5.15. *Let M be an oriented n-dimensional manifold and A an open subset of M with smooth boundary ∂A. Fix on ∂A the orientation induced by M, giving rise to the normal vector n of ∂A. Let g be a metric on M, let γ be the induced metric on ∂A and let ∇ be the covariant derivative associated to the Levi-Civita connection. Then for every vector field X with compact support it holds:*

$$\int_A \nabla_a X^a \sqrt{|g|} du^1 \wedge \cdots \wedge du^n = \int_{\partial A} n_a X^a \sqrt{|\gamma|} du^1 \wedge \cdots \wedge du^{n-1}$$

where $du^1 \wedge \cdots \wedge du^n$ is the n-form locally given by the chart under consideration when evaluating the integral.

Proof. See Wald [5]. For the definition of the metric and the Levi-Civita connection see next section. □

2.6 Curvature

One of the most important features of manifolds is that manifolds allow for a general definition of curvature. That is also the reason manifolds are often used to model physical theories. They do not restrict the shape of the space to have some particular type of curvature. The notion of curvature on a manifold will be made clear in this section.

There are two independent ways of defining a curvature on a manifold. The first is by way of a metric:

Definition 2.6.1. Let M be a manifold. A metric $g \in \Gamma(\boldsymbol{T}^0_2 M)$ on M is a non-degenerate symmetric type-$(0,2)$ tensor field. In a local coordinate system $(u^i)_i$ it is given by $g = g_{ij} du^i \otimes du^j$. This is often also written as $ds^2 = g_{ij} du^i du^j$. The inverse of g is the tensor field $g^{-1} \in \Gamma(\boldsymbol{T}^2_0 M)$ such that $g^{-1}(\omega, g(X, \cdot)) = \omega(X)$ for ω a covector field and X a vector field, thus in components, $(g^{-1})^{ab} g_{bc} = \delta^a_c$. Usually the inverse sign $^{-1}$ is left out in component notation, since there is no confusion because the indices are in an other spot.

Since the metric gives a symmetric bilinear map on every tangent space that takes two vectors and outputs a number, it gives rise to an inner product on those spaces. In this way it defines for example lengths and angles on the manifold.

Example 2.6.2. *Lengths and angles*

The length L of a curve γ on a manifold M with metric g is given by:

$$L = \int \sqrt{g(X_{[\gamma]}(t), X_{[\gamma]}(t))} dt$$

where $X_{[\gamma]}(t) = X_{[\gamma(\cdot - t)]}$.

The angle θ between two curves γ and η on a manifold M with metric g intersecting at $p \in M$ is given by:

$$\begin{aligned} &g(X_{[\gamma]}(\gamma^{-1}(p)), X_{[\eta]}(\eta^{-1}(p))) \\ &= \cos(\theta) g(X_{[\gamma]}(\gamma^{-1}(p)), X_{[\gamma]}(\gamma^{-1}(p))) g(X_{[\eta]}(\eta^{-1}(p)), X_{[\eta]}(\eta^{-1}(p))) \end{aligned}$$

An important way to characterise the metric is with the following concept:

Definition 2.6.3. Let g be a metric on an n-dimensional manifold M. The signature (p, q) of g is the amount p of positive eigenvalues and q negative eigenvalues of g when viewed as a matrix when it is expressed in some coordinate system. The signature is clearly basis independent and the same on whole M, since g is differentiable and non-degenerate, so it is well defined. A metric is called Euclidean if its signature is $(n, 0)$ and Lorentzian if it is $(n-1, 1)$.

When the metric is Lorentzian, vectors can be categorised into three different types:

Definition 2.6.4. Let g be a Lorentzian metric on a manifold M and $p \in M$. A vector $v \in T_pM$ is called spacelike/timelike/null if $g(p)(v,v)$ is respectively positive/negative/zero. Similarly a vector field is called spacelike/timelike/null if it consists of spacelike/timelike/null vectors and a hypersurface is called spacelike/timelike/null if every possible vector field tangent to it is spacelike/timelike/null.

Furthermore note that a metric defines a natural volume form on an oriented manifold as follows. The inverse metric can be extended to $\Omega^k M$ by setting $g^{-1}(a_1 \wedge \cdots \wedge a_k, b_1 \wedge \cdots \wedge b_k) = \prod_i g^{-1}(a_i, b_i)$ for all $a_i, b_i \in \Gamma(\boldsymbol{T}^0_1 M)$. The natural volume form defined by the metric is then given by $g^{-1}(Vol, Vol) = 1$. This condition fixes the volume form up to sign. The sign however is determined by the orientation, so this means that the volume form is uniquely determined. Given a local coordinate system $(u^i)_i$ Vol can be written as $Vol = \tilde{Vol} du^1 \wedge \cdots \wedge du^n$ where $\tilde{Vol}$ is a function. It holds:

$$\begin{aligned} 1 &= g^{-1}(Vol, Vol) = \tilde{Vol}^2 |g^{-1}(du^1 \wedge \cdots \wedge du^n, du^1 \wedge \cdots \wedge du^n)| \\ &= \tilde{Vol}^2 |\det((g^{ij})_{ij})| = \tilde{Vol}^2 |\det(((g_{ij})_{ij})^{-1})| = \tilde{Vol}^2 |g^{-1}| = \tilde{Vol}^2 |g|^{-1} \end{aligned}$$

where $(A_{ij})_{ij}$ is the matrix with at position (i,j) the element A_{ij} and g is the determinant of the matrix representation of the metric. The fact that $(g^{ij})_{ij} = ((g_{ij})_{ij})^{-1}$ follows from $g^{ab} g_{bc} = \delta^a_c$. This thus gives:

$$Vol = \tilde{Vol} du^1 \wedge \cdots \wedge du^n = \sqrt{|g|} du^1 \wedge \cdots \wedge du^n$$

The pullback can be extended to general covariant tensors:

Definition 2.6.5. Let $f : M \to N$ be a differentiable map between manifolds and $(u^i)_i$ a local coordinate system of N. Let $T \in \Gamma(\boldsymbol{T}^0_k N)$ for some k, then T can locally be written as $T = a \cdot du^{i_1} \otimes \cdots \otimes du^{i_k}$ for some function a on N. The pullback of T by f is now defined as:

$$\begin{aligned} f^*(T)(p) &= f^*(a \cdot du^{i_1} \otimes \cdots \otimes du^{i_k})(p) \\ &= (a \circ f)(p) \cdot f^*(du^{i_1}|_{f(p)}) \otimes \cdots \otimes f^*(du^{i_k}|_{f(p)}) \in \boldsymbol{T}^0_k(T^*_p M) \end{aligned}$$

The pull back $f^* \colon \Gamma(\bigoplus \boldsymbol{T}^0_k N) \to \Gamma(\bigoplus_k \boldsymbol{T}^0_k M)$, where $\bigoplus$ denotes a fibrewise direct sum (it is clear from Example 2.4.7 that this is a vector bundle), is then defined as the linear extension of this map.

Given a manifold with a metric, there is a natural metric on any submanifold:

Definition 2.6.6. Let M be a manifold with metric g and X be a submanifold of M. Let $\iota\colon X \to M$ be the inclusion map. Then g induces a metric on X defined by $\gamma = \iota^*(g)$. This is called the induced metric.

The metric is often used to change the tensor type:

Definition 2.6.7. Let M be a manifold and $T \in \Gamma(\boldsymbol{T}^k_l M)$. Then a $(k-1, l+1)$-type tensor field can be defined by:

$$T(\omega_1, \ldots, g(X_i, \cdot), \ldots, \omega_k, X_{k+1}, \ldots, X_{k+l})$$

for $\omega_1, \ldots, \omega_k \in \boldsymbol{T}^0_1 M$ and $X_i, X_{k+1}, \ldots, X_{k+l} \in \boldsymbol{T}^1_0 M$. In components this tensor field is given by $g_{ca_i} T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$. This is called lowering an index. Conversely a $(k+1, l-1)$-type tensor can be defined by:

$$T(\omega_1, \ldots, \omega_k, X_{k+1}, \ldots, g^{-1}(\omega_i, \cdot), \ldots, X_{k+l})$$

for $\omega_1, \ldots, \omega_k, \omega_i \in \boldsymbol{T}^0_1 M$ and $X_{k+1}, \ldots, X_{k+l} \in \boldsymbol{T}^1_0 M$. In components this tensor is given by $g^{cb_i} T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$. This is called raising an index.

Because g is non-degenerate it is clear that these operations are isomorphisms on the fibres of the tensor bundles.

Another way to characterise curvature on a manifold is by way of a connection:

Definition 2.6.8. Let $\pi\colon E \to M$ be a vector bundle over M. A connection ∇ on E is an bilinear map $\nabla\colon \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E)$ satisfying $\nabla(X, fs) = \nabla_X fs = f\nabla_X s + (Xf)s$.

The covariant derivative can thus be seen as a particular connection on the tensor bundle. In the same way as for the covariant derivative, a connection can be expressed in terms of connection coefficients Γ^k_{ij}. If $(u^i)_i$ is some local coordinate system of M and $(e_i(p))_i$ is some basis of the fibre E_p, then $X = X^i \frac{\partial}{\partial u^i}$ and $s(p) = s^i e_i(p)$. In the same way as equation 2.5.1 one finds:

$$\nabla_X s = (X^i \frac{\partial s^k}{\partial u^i}(p) + X^i s^j \Gamma^k_{ij}) e_k(p)$$

where the connections coefficients are defined as $\Gamma^k_{ij} = \nabla_{\frac{\partial}{\partial u^i}} e_j(p)$.

Curves on manifolds extend to curves on vector bundles:

Definition 2.6.9. Let $\pi\colon E \to M$ be a vector bundle and $\gamma\colon (-\epsilon, \epsilon) \to M$ a curve. A curve $\eta\colon (-\epsilon, \epsilon) \to E$ on E is called a curve above γ if $\eta(t) \in E_{\gamma(t)}$ for all t.

The connection then defines a notion of parallel transport of an element of E along a curve in M:

Definition 2.6.10. Let $\pi\colon E \to M$ be a vector bundle, ∇ a connection on E and $\gamma\colon (-\epsilon, \epsilon) \to M$ a curve. A curve $\eta\colon (-\epsilon, \epsilon) \to E$ is called parallel along γ if $\nabla_{\dot{\gamma}(t)}\eta(t) = 0$ for all $t \in (-\epsilon, \epsilon)$, where $\dot{\gamma}(t) = X_{[\gamma(\cdot - t)]} \in T_{\gamma(t)}M$. A section $s \in \Gamma(E)$ is called parallel along γ if $\nabla_{\dot{\gamma}(t)}(s \circ \gamma)(t) = 0$ for all $t \in (-\epsilon, \epsilon)$.

Proposition 2.6.11. *Let $\pi\colon E \to M$ be a vector bundle, ∇ a connection on E, $\gamma\colon (-\epsilon, \epsilon) \to M$ a curve and $e_0 \in E_{\gamma(0)}$. Then there exists a unique curve η above γ that is parallel along γ and has $\eta(0) = e_0$.*

Proof. See Kobayashi & Nomizu [1]. □

In this way all fibres of the tangent bundle can be connected to one another:

Definition 2.6.12. Let $\pi\colon E \to M$ be a vector bundle, ∇ a connection on E and $\gamma\colon (-\epsilon, \epsilon) \to M$ a curve. Parallel transport along γ from time t_0 to t_1 is the map $T_\gamma^{t_0,t_1}\colon E_{\gamma(t_0)} \to E_{\gamma(t_1)}$ given by $T_\gamma^{t_0,t_1}(e_0) = \eta(t_1)$ where η is the unique curve above γ that is parallel along γ and has $\eta(t_0) = e_0$.

Proposition 2.6.13. *The parallel transport map $T_\gamma^{t_0,t_1}\colon E_{\gamma(t_0)} \to E_{\gamma(t_1)}$ is a linear isomorphism.*

Proof. See Kobayashi & Nomizu [1]. □

A connection thus gives a way of relating one fibre to the next in a vector bundle. It can be thought of as defining the direction of the normal to a fibre and thus it gives a way to make a directional derivative for sections of the fibre bundle. This directional derivative for sections is what is called the covariant derivative associated to the connection.

A connection on the tangent bundle has two properties, a torsion and a curvature:

Definition 2.6.14. Let $\pi\colon TM \to M$ be the tangent bundle and ∇ a connection on TM. The torsion $\hat{T}$ of ∇ is a map $\hat{T}\colon \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ given by:

$$\hat{T}(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$$

The curvature $\hat{R}$ of ∇ is the map $\hat{R}\colon (\Gamma(TM))^3 \to \Gamma(TM)$ given by:

$$\hat{R}(X,Y)Z = \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]}Z$$

Both the torsion and the curvature are $\mathcal{C}^\infty(M)$-linear:

Lemma 2.6.15. *The torsion and the curvature are $\mathcal{C}^\infty(M)$-linear in all their entries.*

Proof. Let M be a manifold, let $X, Y, Z \in \Gamma(TM)$ and let $f \in \mathcal{C}^\infty(M)$. Then it holds:

$$\begin{aligned}
T(fX,Y) &= \nabla_{fX}Y - \nabla_Y(fX) - [fX,Y] \\
&= f\nabla_X Y - (\nabla_Y f)X - f\nabla_Y X - fXY + Y(fX) \\
&= f\nabla_X Y - f\nabla_Y X - Y(f)X - fXY + Y(f)X + fYX \\
&= f(\nabla_X Y - \nabla_Y X - [X,Y]) = fT(X,Y) \\
\hat{R}(fX,Y)Z &= \nabla_{fX}(\nabla_Y Z) - \nabla_Y(\nabla_{fX}Z) - \nabla_{[fX,Y]}Z \\
&= f\nabla_X(\nabla_Y Z) - \nabla_Y(f\nabla_X Z) - \nabla_{fXY-Y(fX)}Z \\
&= f\nabla_X(\nabla_Y Z) - (\nabla_Y f)\nabla_X Z - f\nabla_Y(\nabla_X Z) \\
&\quad - \nabla_{fXY-Y(f)X-fXY}Z \\
&= f\nabla_X(\nabla_Y Z) - (Yf)\nabla_X Z - f\nabla_Y(\nabla_X Z) \\
&\quad - \nabla_{f[X,Y]}Z + Y(f)\nabla_X Z \\
&= f(\nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]}Z) = fR(X,Y)Z \\
\hat{R}(X,Y)(fZ) &= \nabla_X\nabla_Y(fZ) - \nabla_Y\nabla_X(fZ) - \nabla_{[X,Y]}(fZ) \\
&= \nabla_X((\nabla_Y f)Z) + \nabla_X(f\nabla_Y Z) - \nabla_Y((\nabla_X f)Z) \\
&\quad - \nabla_Y(f\nabla_X fZ) - (\nabla_{[X,Y]}f)Z - f\nabla_{[X,Y]}Z \\
&= \nabla_X((Yf)Z) + \nabla_X(f\nabla_Y Z) - \nabla_Y((Xf)Z) \\
&\quad - \nabla_Y(f\nabla_X Z) - ([X,Y]f)Z - f\nabla_{[X,Y]}Z \\
&= (\nabla_X(Yf))Z + (Yf)\nabla_X Z + (\nabla_X f)\nabla_Y Z + f\nabla_X(\nabla_Y Z) \\
&\quad - (\nabla_Y(Xf))Z - (Xf)\nabla_Y Z - (\nabla_Y f)\nabla_X Z \\
&\quad - f\nabla_Y(\nabla_X Z) - ([X,Y]f)Z - f\nabla_{[X,Y]}Z \\
&= (X(Yf))Zf\nabla_X(\nabla_Y Z) - (Y(Xf))Z - f\nabla_Y(\nabla_X Z) \\
&\quad - ([X,Y]f)Z - f\nabla_{[X,Y]}Z \\
&= f(\nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]}Z) = fR(X,Y)Z
\end{aligned}$$

Furthermore it holds $\hat{T}(Y,X) = -\hat{T}(X,Y)$ and $\hat{R}(Y,X)Z = -\hat{R}(X,Y)Z$. This means that T and $\hat{R}$ are $\mathcal{C}^\infty(M)$-linear in all their entries. □

From this it follows that $\hat{T}$ and $\hat{R}$ can be seen as giving multilinear maps $T_p\colon T_pM \times T_pM \to T_pM$ and $R_p\colon (T_pM)^3 \to T_pM$ smoothly varying in p. By Example 2.3.2, $\hat{T}$ and $\hat{R}$ give rise to tensor fields $T \in \Gamma(\boldsymbol{T}^1_2 M)$ and $R \in \Gamma(\boldsymbol{T}^1_3 M)$.

Definition 2.6.16. The tensor field $R \in \Gamma(\boldsymbol{T}^1_3 M)$ is called the Riemann curvature tensor. It is given by $R(\omega, X, Y, Z) = \omega(R(Y, Z)X)$.

The components of these tensor fields are in some local coordinate system $(u^i)_i$:

$$\begin{aligned}
T^a{}_{bc} &= du^a(T(\frac{\partial}{\partial u^b}, \frac{\partial}{\partial u^c})) = du^a(\nabla_{\frac{\partial}{\partial u^b}} \frac{\partial}{\partial u^c} - \nabla_{\frac{\partial}{\partial u^c}} \frac{\partial}{\partial u^b} - [\frac{\partial}{\partial u^b}, \frac{\partial}{\partial u^c}]) \\
&= du^a(\Gamma^k_{bc} - \Gamma^l_{cb}) = \Gamma^a_{bc} - \Gamma^a_{cb} \\
R^a{}_{bcd} &= du^a(R(\frac{\partial}{\partial u^c}, \frac{\partial}{\partial u^d})\frac{\partial}{\partial u^b}) \\
&= du^a(\nabla_{\frac{\partial}{\partial u^c}}(\nabla_{\frac{\partial}{\partial u^d}} \frac{\partial}{\partial u^b}) - \nabla_{\frac{\partial}{\partial u^d}}(\nabla_{\frac{\partial}{\partial u^c}} \frac{\partial}{\partial u^b}) - \nabla_{[\frac{\partial}{\partial u^c}, \frac{\partial}{\partial u^d}]} \frac{\partial}{\partial u^b}) \\
&= du^a(\nabla_{\frac{\partial}{\partial u^c}} \Gamma^k_{db} - \nabla_{\frac{\partial}{\partial u^d}} \Gamma^l_{cb}) \\
&= du^a((\frac{\partial}{\partial u^c} + \Gamma^m_{ck})\Gamma^k_{db} - ((\frac{\partial}{\partial u^d} + \Gamma^n_{dl}))\Gamma^l_{cb}) \\
&= du^a(\frac{\partial}{\partial u^c}\Gamma^k_{db} + \Gamma^m_{ck}\Gamma^k_{db} - \frac{\partial}{\partial u^d}\Gamma^l_{cb} - \Gamma^n_{dl}\Gamma^l_{cb}) \\
&= \frac{\partial}{\partial u^c}\Gamma^a_{db} - \frac{\partial}{\partial u^d}\Gamma^a_{cb} + \Gamma^a_{ck}\Gamma^k_{db} - \Gamma^a_{dl}\Gamma^l_{cb}
\end{aligned}$$

where the $[\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}]$ vanish because partial derivatives commute on $\mathcal{C}^\infty$-functions. The contractions of the Riemann tensor are also often encountered:

Definition 2.6.17. Let M be a manifold with metric g. The Ricci tensor is a tensor field $Ric \in \Gamma(\boldsymbol{T}^0_2 M)$ given in components by $Ric_{ab} = R^i{}_{aib}$. Since the amount of indices of the Riemann tensor and the Ricci tensor is different, the components of Ric are often just written as R_{ab}. The Ricci tensor is also known as the intrinsic curvature. The Ricci scalar $R \in \Gamma(\boldsymbol{T}^0_0 M)$ is a differentiable function whose value is given by $R = g^{ab} Ric_{ab}$.

Curvature can also be defined on a submanifold. Note that if M is a manifold with submanifold N, then it is clear that TN is a subspace of TM in a natural way. Using this fact two notions of curvature can be defined for a submanifold.

Definition 2.6.18. Let M be a manifold with connection ∇. Let N be a submanifold of M and let $\iota\colon N \to M$ be the inclusion map. The induced connection $\hat{\nabla}$ on N is a connection on N defined by, for $v, w \in \Gamma(TN)$:

$$\hat{\nabla}_v w = \nabla_{D\iota(v)}(D\iota(w))$$

Definition 2.6.19. Let M be a manifold with metric g and connection ∇. Let N be a submanifold of M with induced metric γ and induced connection $\hat{\nabla}$. The intrinsic curvature of N is the Ricci tensor $\hat{Ric}$ on N with respect to the induced connection on N.

Note that if the submanifold N is one dimension smaller than M, then there clearly exists a differentiable vector field n normal to N, meaning that $T_pM = T_pN \oplus Span(n(p))$ holds for all $p \in N$, and that its restriction $n|_N$ is unique up to a scalar. This allows for the definition of the second kind of curvature.

Definition 2.6.20. Let M be an r-dimensional manifold with metric g and connection ∇. Let N be a $(r-1)$-dimensional submanifold of M with inclusion map $\iota\colon N \to M$. Finally let n be a unit vector field normal to N, so its restriction $n|_N$ is unique up to sign. Then the extrinsic curvature $K \in \Gamma(\boldsymbol{T}^0_2 N)$ is the type-$(0,2)$ tensor field given by, for $v, w \in \Gamma(TN)$:

$$K(v,w) = g(D\iota(v), \nabla_{D\iota(w)} n|_N)$$

In physics often a special connection is chosen:

Definition 2.6.21. Let M be a manifold with metric g. The Levi-Civita connection ∇ is a connection on the tangent bundle of M such that its torsion vanishes for all vector fields and it is compatible with the metric, meaning $\nabla_X g(Y,Z) = 0$, where $g(Y,Z)$ is a function, for all $X, Y, Z \in \Gamma(TM)$, or in components $\nabla_a g_{bc} = 0$.

The connection coefficients of the Levi-Civita connection are called Christoffel symbols.

Proposition 2.6.22. *Let M be a manifold with metric g, then there exists a unique Levi-Civita connection on TM.*

Proof. See Kobayashi & Nomizu [1]. □

Christoffel symbols can be expressed in terms of the metric.

Proposition 2.6.23. *For Christoffel symbols it holds:*

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} \left(\frac{\partial g_{jl}}{\partial u^k} + \frac{\partial g_{kl}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^l} \right)$$

Proof. See Carroll [4]. □

For the Levi-Civita connection the following relations hold:

Proposition 2.6.24. *Consider the Levi-Civita connection on a manifold with a metric. For this connection it holds:*

$$\begin{aligned}
\mathcal{L}_X Y &= \nabla_X Y - \nabla_Y X \\
\Gamma^a_{bc} &= \Gamma^a_{cb} \\
R_{abcd} &= -R_{bacd} \\
R_{abcd} &= -R_{abdc} \\
R_{abcd} &= R_{cdab} \\
R_{abcd} &= -R_{bacd} \\
R_{abcd} + R_{adbc} + R_{acdb} &= 0 \\
R_{a[bcd]} &= 0 \\
R_{[abcd]} &= 0 \\
\nabla_{[k} R_{ab]cd} &= 0 \\
R_{ab} &= R_{ba}
\end{aligned}$$

Proof. See Carroll [4]. □

Chapter 3

General Relativity

The current understanding of Nature at her biggest scales is given by the theory of General Relativity. General Relativity is namely the theory that describes the origin of gravity. Since at galactic scales, ranging from the solar system to the entirety of he universe, the other three fundamental forces play a negligible role, General Relativity alone can predict the evolution of Nature at these scales. It does this by describing the universe as a four dimensional manifold and relating its curvature to the energy-momentum content of the universe. Gravity is then just the physical effect of this curvature on objects inhabiting the universe. Though the predictive power of General Relativity is great, especially in smaller scale regions about the size of the solar system, it has its known limitations. The search for a sound theory with greater predictive power, of which this thesis is a part, is thus ongoing. However, since General Relativity offers such a good description in smaller scale regions, it should be a limiting case for any new theory. It is therefore fruitful to investigate the theory of General Relativity.

3.1 Geometry of the Universe

General Relativity thus describes the universe as a four dimensional manifold. This section will explain why this description was chosen and what it looks like.

In Special Relativity, space is modelled by the vector space $\mathbb{R}^4$. There are three spatial dimensions defined on it and one temporal one and an inner product is defined on $\mathbb{R}^4$ by the matrix $diag(-1, 1, 1, 1)$, given in a basis such that the first basis vector spans the temporal dimension. Moreover, gravity is thought of as just a force F_G given by some potential Φ. Each object then has a "gravitational charge" m_G associated to it, i.e. a measure of how much it is affected by the force of gravity, analogous to electric charge for electromagnetism. Newton's Second Law then states that:

$$m_I a = F_G = m_G \nabla \Phi$$

where a is the acceleration of the object and ∇ is just the gradient for differentiable functions on $\mathbb{R}^4$. Moreover, m_I, called the inertial mass of the object, is a property of the object defined as the constant in Newton's Second Law. It thus measures the objects accelerations reaction to a force. The quotient $\frac{m_G}{m_I}$ is therefore a measure of the effect of the force on the motion of the object. There is a priori no relation between m_I and m_G other than the one given. From experiments however it is concluded that $m_I = m_G$, thus that $a = \nabla \Phi$, for all objects. This means that locally there can be made no distinction between gravitational acceleration and acceleration stemming from another source. This leads naturally to the following assumption:

Assumption 3.1.1. *Locally physics reduces to special relativity. In particular it is impossible to detect a gravitational field by means of local experiments. This assumption is called the Einstein Equivalence Principle.*

This however raises a problem. Suppose you want to measure some force in a local experiment. To do this you have to measure the acceleration of the object in question. Any gravitational field will also contribute to the acceleration of the object. However, since locally the gravitational field is undetectable, it is completely unclear what part of the acceleration of the object stems from gravity and what part stems from the force you wanted to measure. Hence under Einstein's Equivalence Principle it becomes impossible to measure forces locally. The problem encountered here is that it is not possible to find a frame of reference in which the motion is unaccelerated, since it is unclear what the acceleration of the object due to gravity is. This is because the acceleration due to gravity can in turn not be defined well since all objects are always subject to gravity in the same way, i.e. have the same quotient $\frac{m_G}{m_I} = 1$, so there is no frame of reference to measure it against. The concept of gravity as a force which gives rise to acceleration thus is very problematic in this setting. In stead of keeping to try to find a way to find unaccelerated motion, Einstein gave the following solution:

Solution 3.1.2. Unaccelerated motion is defined as that motion that an object exhibits in a gravitational field (free from other forces).

As a consequence this solution imposes a curved manifold structure on the space in which the object moves, i.e. the universe. Firstly the Einstein Equivalence Principle states that physics locally reduces to Special Relativity. This means three things, that the universe locally looks like $\mathbb{R}^4$, that velocity comparisons can only be done locally and that there exists a Lorentzian metric on the manifold. The first point leads to the idea of a manifold and the second to that of the tangent space. The Lorentzian metric arises from the fact that there is a Lorentzian inner product in Special Relativity which is used to measure lengths and angles. So the universe can be described by a manifold with a Lorentzian metric. The definition of unaccelerated motion as motion in a gravitational field then gives a unique set of curves on the manifold given by paths of unaccelerated test particles. These curves then naturally give rise to a covariant derivative along them and thus a connection on the manifold, defined by parallel transport along the curves. By choosing the Levi-Civita connection, the parallel transport becomes dependent on the metric and the metric can therefore be used to model gravity, i.e. the change of the curves as the gravitational field changes. Gravity is thus represented as the curvature of the manifold, via the metric and thus the connection and thus also the curvature tensor and the Riemann tensor. This is the setting in which the theory of General Relativity in formulated. The equations of motion given by General Relativity then dictate the evolution of the metric and thus gravity.

3.2 Principle of Least Action

In physics it is often difficult to obtain the equations of motion, that describe the behaviour of the system, directly. Luckily the equations of motion can be derived from the Lagrangian of the system, which, most of the time, is a lot simpler to do. The method of obtaining the equations of motion of a system from its Lagrangian is called the Principle of Least Action.

The Principle of Least Action is an algorithm that, if you input the Lagrangian of the system, outputs the equations of motion of that system. For classical mechanics it works as follows. In classical mechanics the object of study is the path that a particle or object takes. If M is the manifold of possible positions of the object, then its path in a finite time interval is described by a curve $\gamma\colon (-\epsilon, \epsilon) \to M$ for some $\epsilon > 0$. Let L be the Lagrangian of the system, not dependent on time. Then L is determined by the position and the velocity of the object, thus L can be thought of as a function from TM to $\mathbb{R}$. The action S of the system between two events A

and B with coordinates (x_A, t_A) respectively (x_B, t_B) is then defined by:

$$S = \int_{t_A}^{t_B} L dt$$

The only requirement made here is that the object moves from event A to event B, thus from position x_A at time t_A to position x_B at time t_B. The way it does this is arbitrary, hence the value of S is dependent of the path γ of the object. In general the action can thus be seen as a map $S\colon \mathcal{C}^\infty((-\epsilon, \epsilon), M) \to \mathbb{R}$ given by:

$$S[\gamma] = \int_{-\epsilon}^{\epsilon} L((\gamma(t), \dot{\gamma}(t)))dt$$

where $\dot{\gamma}(t) \in T_{\gamma(t)}M$ is the velocity of the object along its path. The Principle of Least Action now states that the path taken by the object is the path that is a critical point for the action. To find this path consider a parametrised family of curves $\Phi = \{\gamma_s \colon s \in (-\alpha, \alpha)\}$ for some $\alpha > 0$ depending differentiably on s and a differentiable map $\phi\colon (-\epsilon, \epsilon) \times (-\alpha, \alpha) \to M$ given by $\phi(t, s) = \gamma_s(t)$. Then for this family the action can be viewed as a function $S^\Phi\colon (-\alpha, \alpha) \to \mathbb{R}$ given by $S^\Phi(s) = S[\gamma_s]$. The path taken by an object in this system can then be found by setting the derivative of this function to zero:

$$\frac{dS[\gamma_s]}{ds} = \frac{d}{ds} \int_{-\epsilon}^{\epsilon} L((\phi(t, s), \frac{\partial}{\partial t}\phi(t, s)))dt = 0$$

Manipulation of this expression will result in obtaining a differential equation that describes all paths that an object in this system can take. If there are multiple independent variables, this procedure is repeated for each variable. The resulting set of differential equations are the equations of motion of the system.

This procedure is generalisable to General Relativity. There are however a couple of important differences. To make these precise, the notion of infinite dimensional manifolds is necessary. This however is beyond the scope of this thesis, so only a sketch of the ideas will be given here. For a more detailed description of infinite dimensional manifolds see Kriegl & Michor [6]. In General Relativity the variable of study is the metric, which describes gravity. The metric is a section of the tensor bundle $\boldsymbol{T}^0_2 M$. The space of sections $\Gamma(\boldsymbol{T}^0_2 M)$ is a vector space by Definition 2.4.2. It is clear that in general it is infinite dimensional. Choosing an appropriate topology (see Kriegl & Michor [6] for the specifics), $\Gamma(\boldsymbol{T}^0_2 M)$ becomes an infinite dimensional manifold, analogous to the case of Example 2.1.6.

Where as the position and velocity of a particle can be specified by giving a particular time and path, the metric depends on its position of evaluation on the whole space-time manifold. To specify a metric, its value on the whole manifold must be given as well as its path. Therefore the integral in the action is taken over the complete space-time manifold in stead. Since the whole manifold is taken as integration domain, General Relativity considers the Lagrangian density of the system in stead of the Lagrangian. The Lagrangian density is then a map $L\colon \boldsymbol{T}^0_2 M \to \mathbb{R}$. Furthermore, to guarantee that the integral exists the action is given by not a single map but a family of maps $\{S_K\colon \Gamma(\boldsymbol{T}^0_2 M) \to \mathbb{R} | K \subset M \text{ compact}\}$ given by:

$$S_K(g) = \int_K L((x, g(x)))Vol := \int_{K^\circ} L((x, g(x)))Vol$$

where$K^\circ = K \backslash \partial K$ is the interior of K and Vol a volume form on M.

Consider now a differentiable map $\phi\colon M \times (-\alpha, \alpha) \to \boldsymbol{T}^0_2 M$ for some $\alpha > 0$ such that $\phi(\cdot, s) \in \Gamma(\boldsymbol{T}^0_2 M)$ for all s. This map gives thus a family of metrics $\Phi = \{\phi(\cdot, s)\colon s \in (-\alpha, \alpha)\}$ depending differentiably on s. For this family the actions can be viewed as a functions $S^\Phi_K\colon (-\alpha, \alpha) \to \mathbb{R}$ given by $S^\Phi_K(s) = S_K[\phi(\cdot, s)]$. The Principle of Least Action then again states that the metric of the system can then be found by setting the derivative of this function to zero:

$$\frac{dS_K[\phi(\cdot, s)]}{ds} = \frac{d}{ds} \int_K L((x, \phi(x, s)))Vol = 0$$

for all K. Manipulation of this expression will give the equations of motion of the system for the metric. It is clear that the same procedure can be applied to any kind of tensor field on the manifold.

In order to simplify the expressions when manipulating the integral, the following operator is introduced. Let $\pi\colon E \to M$ be a vector bundle of rank r. Suppose $g \in \Gamma(E)$ and $\phi\colon M \times (-\alpha, \alpha) \to E$ a map for some $\alpha > 0$ such that $\phi(\cdot, s) \in \Gamma(E)$ for all s and $\phi(\cdot, 0) = g$. The operator δ acting on g is then given by $(\delta g)(x) = \frac{\partial \phi(x,s)}{\partial s}|_{s=0}$ for all $x \in M$. Note that this operation is well defined because for fixed $x \in M$ the map $\phi(x, \cdot)$ is just a map from $(-\alpha, \alpha)$ to E_x, which is a vector space, so the derivative in the standard vector calculus sense can indeed be applied. Since $\phi(\cdot, s) \in \Gamma(E)$ for all s, this defines a curve on $\Gamma(E)$. Hence $\delta g = \frac{\partial \phi(\cdot,s)}{\partial s}|_{s=0}$ can be thought of as a tangent vector of $\Gamma(E)$ at g in the direction of ϕ. This function ϕ is therefore called a variation of g. Moreover it follows that $\phi(p, s) \in E_p$ for all s. This means that for all x:

$$(\delta g)(x) = \frac{\partial \phi(x, s)}{\partial s}|_{s=0} \in T_{g(x)} E_x = V_{g(x)} E$$

Hence more specifically δg is a differentiable section of the pullback bundle $g^* VE$. Using this notation the requirement of the Principle of Least Action can be stated as:

$$\delta S_K^\Phi = 0$$

for all $K \subset M$ compact.

Since δ is basically given by a 'normal' derivative and $\Gamma(E)$, where E is a vector bundle, is a vector space, it follows that δ is linear. Moreover, from the definition of the tensor product and the Leibniz property of the 'normal' derivative, it follows that δ also satisfies the Leibniz property, i.e. $\delta(S \otimes T) = \delta S \otimes T + S \otimes \delta T$. This makes computations a lot easier.

As a final remark, when performing computations, one typically wants to apply the Theorem of Stokes and then neglect the boundary terms. When integrating over the whole manifold this can be done by choosing the variations to be become zero as they approach the boundary of the manifold. The integration however is not over the complete manifold, but just over (the interior of) a compact subset K. Therefore the variations shall in general not vanish on the boundary of K and the boundary terms cannot be neglected. However it turns out that the all variations possible are always compactly supported, if $\Gamma(E)$ where to become a differentiable manifold. See chapter 42.2 of Kriegl & Michor [6] for more information. A K can thus always be chosen such that the support of the variation is contained in K and the boundary term can be neglected. The only real boundary terms therefore come from the boundary of the manifold itself, which also can be chosen to be zero as stated earlier.

3.3 The Einstein-Hilbert action

Using the Principle of Least Action, the equations of motion of the theory of General Relativity can now be found from the Lagrangian. This section is dedicated to obtaining this result. It follows the derivation of Carroll [4].

In General Relativity the universe is described by an oriented 4-dimensional differentiable manifold M with metric g and corresponding Levi-Civita connection ∇. The Lagrangian of the system is given by $L = \frac{c^4}{16\pi G}R + L_M$, where c is the speed of light, G the gravitational constant, R the Ricci scalar and L_M is the Lagrangian of the matter. The corresponding action is then:

$$\begin{aligned} S_K = \int_K L Vol &= \int_K (\frac{c^4}{16\pi G}R + L_M)\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 \\ &= \frac{c^4}{16\pi G}\int_K R\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 + \int_K L_M\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 \end{aligned}$$

where the natural volume form defined by the metric is chosen. The left integral after the final equality is called the Einstein-Hilbert action. Given now a family of metrics $\Phi = \{\phi(\cdot, s) \colon s \in (-\alpha, \alpha)\}$ depending differentiably on s such that $\phi(\cdot, s) = g$. The Principle of Least Action now states that:

$$0 = \delta S_K^{\Phi} = \frac{c^4}{16\pi G}\delta\int_K R\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 + \delta\int_K L_M\sqrt{|g|}du^1 \wedge \cdots \wedge du^4$$

To obtain the equations of motion, first the first integral is manipulated. It holds:

$$\begin{aligned} \delta\int_K R\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 &= \delta\int_K g^{\mu\nu}R_{\mu\nu}\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 \\ &= \int_K ((\delta g^{\mu\nu})R_{\mu\nu}\sqrt{|g|} + g^{\mu\nu}(\delta R_{\mu\nu})\sqrt{|g|} \\ &\quad + R\delta\sqrt{|g|})du^1 \wedge \cdots \wedge du^4 \end{aligned}$$

So the integral splits in three parts. The goal is to express everything that is under the integral as a function times a variation of the inverse metric. The first term in the integral is thus already of the correct form.

The Ricci tensor is given by a contraction of the Riemann tensor. The variation of the Ricci tensor can therefore be obtained from the variation of the Riemann tensor. The Riemann tensor is given by:

$$R^{\rho}{}_{\mu\lambda\nu} = \frac{\partial}{\partial u^{\lambda}}\Gamma^{\rho}_{\nu\mu} - \frac{\partial}{\partial u^{\nu}}\Gamma^{\rho}_{\lambda\mu} + \Gamma^{\rho}_{\lambda\tau}\Gamma^{\tau}_{\nu\mu} - \Gamma^{\rho}_{\nu\tau}\Gamma^{\tau}_{\lambda\mu}$$

as shown under Definition 2.6.16. Since the Christoffel symbols $\Gamma^{\rho}_{\nu\mu}$ are just functions on the manifold, they can be considered as sections on an appropriate vector bundle (specifically $M \times \mathbb{R}^{4^3}$). Therefore the variation of the Riemann tensor is:

$$\begin{aligned}\delta R^{\rho}{}_{\mu\lambda\nu} = \frac{\partial}{\partial u^{\lambda}}\delta\Gamma^{\rho}_{\nu\mu} - \frac{\partial}{\partial u^{\nu}}\delta\Gamma^{\rho}_{\lambda\mu} + (\delta\Gamma^{\rho}_{\lambda\tau})\Gamma^{\tau}_{\nu\mu} \\ - (\delta\Gamma^{\rho}_{\nu\tau})\Gamma^{\tau}_{\lambda\mu} + \Gamma^{\rho}_{\lambda\tau}\delta\Gamma^{\tau}_{\nu\mu} - \Gamma^{\rho}_{\nu\tau}\delta\Gamma^{\tau}_{\lambda\mu}\end{aligned}$$

Now since $\Gamma(M \times \mathbb{R}^{4^3})$ is a vector space and a variation $\delta\Gamma^{\rho}_{\nu\mu}$ can be thought of as an element of the tangent space of $\Gamma(M \times \mathbb{R}^{4^3})$, by the infinite dimensional analogue of Example 2.2.4 it then follows that $\delta\Gamma^{\rho}_{\nu\mu}$ can be thought of as an element of $\Gamma(M \times \mathbb{R}^{4^3})$. Since $\Gamma(M \times \mathbb{R}^{4^3})$ is a vector space it follows that $\delta\Gamma^{\rho}_{\nu\mu}$ is the difference of two $\Gamma^{\rho}_{\nu\mu}$. Thus it holds:

$$\delta\Gamma^{\rho}_{\nu\mu} = \Gamma'^{\rho}_{\nu\mu} - \Gamma^{\rho}_{\nu\mu} = \frac{\partial}{\partial u^{\nu}} + \Gamma'^{\rho}_{\nu\mu} - \frac{\partial}{\partial u^{\nu}} - \Gamma^{\rho}_{\nu\mu} = \nabla'_{\nu} - \nabla_{\nu}$$

where the ∇ are the connections associated to the connection coefficients. Since the connections are tensorial objects, it follows that $\delta\Gamma^{\rho}_{\nu\mu}$ is also a tensor. Its covariant derivative is therefore well defined and is given by:

$$\nabla_{\lambda}\delta\Gamma^{\rho}_{\nu\mu} = \frac{\partial\delta\Gamma^{\rho}_{\nu\mu}}{\partial u^{\lambda}} + \Gamma^{\rho}_{\lambda\tau}\delta\Gamma^{\tau}_{\nu\mu} - \Gamma^{\tau}_{\lambda\nu}\delta\Gamma^{\rho}_{\tau\mu} - \Gamma^{\tau}_{\lambda\mu}\delta\Gamma^{\rho}_{\nu\tau}$$

which follows from the component expression of the covariant derivative given after Definition 2.5.6. Note that from this it follows that:

$$\begin{aligned}\delta R^{\rho}{}_{\mu\lambda\nu} &= \frac{\partial}{\partial u^{\lambda}}\delta\Gamma^{\rho}_{\nu\mu} - \frac{\partial}{\partial u^{\nu}}\delta\Gamma^{\rho}_{\lambda\mu} + \Gamma^{\tau}_{\nu\mu}\delta\Gamma^{\rho}_{\lambda\tau} - \Gamma^{\tau}_{\lambda\mu}\delta\Gamma^{\rho}_{\nu\tau} + \Gamma^{\rho}_{\lambda\tau}\delta\Gamma^{\tau}_{\nu\mu} - \Gamma^{\rho}_{\nu\tau}\delta\Gamma^{\tau}_{\lambda\mu} \\ &= \frac{\partial}{\partial u^{\lambda}}\delta\Gamma^{\rho}_{\nu\mu} + \Gamma^{\rho}_{\lambda\tau}\delta\Gamma^{\tau}_{\nu\mu} - \Gamma^{\tau}_{\lambda\mu}\delta\Gamma^{\rho}_{\nu\tau} - (\frac{\partial}{\partial u^{\nu}}\delta\Gamma^{\rho}_{\lambda\mu} + \Gamma^{\rho}_{\nu\tau}\delta\Gamma^{\tau}_{\lambda\mu} - \Gamma^{\tau}_{\nu\mu}\delta\Gamma^{\rho}_{\lambda\tau}) \\ &= \frac{\partial}{\partial u^{\lambda}}\delta\Gamma^{\rho}_{\nu\mu} + \Gamma^{\rho}_{\lambda\tau}\delta\Gamma^{\tau}_{\nu\mu} - \Gamma^{\tau}_{\lambda\nu}\delta\Gamma^{\rho}_{\tau\mu} - \Gamma^{\tau}_{\lambda\mu}\delta\Gamma^{\rho}_{\nu\tau} \\ &\quad - (\frac{\partial}{\partial u^{\nu}}\delta\Gamma^{\rho}_{\lambda\mu} + \Gamma^{\rho}_{\nu\tau}\delta\Gamma^{\tau}_{\lambda\mu} - \Gamma^{\tau}_{\nu\lambda}\delta\Gamma^{\rho}_{\tau\mu} - \Gamma^{\tau}_{\nu\mu}\delta\Gamma^{\rho}_{\lambda\tau}) \\ &= \nabla_{\lambda}\delta\Gamma^{\rho}_{\nu\mu} - \nabla_{\nu}\delta\Gamma^{\rho}_{\lambda\mu}\end{aligned}$$

where the fact that the Christoffel symbols, i.e. the connection coefficients of the Levi-Civita connection, are symmetric in the lower indices is used. The variation of the Ricci tensor is thus given by:

$$\delta R_{\mu\nu} = \delta R^{\lambda}{}_{\mu\lambda\nu} = \nabla_{\lambda}\delta\Gamma^{\lambda}_{\nu\mu} - \nabla_{\nu}\delta\Gamma^{\lambda}_{\lambda\mu}$$

Since the connection is compatible with the metric it follows that:

$$\begin{aligned}
&\int_K g^{\mu\nu}(\delta R_{\mu\nu})\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 \\
&= \int_K g^{\mu\nu}(\nabla_\lambda \delta\Gamma^\lambda_{\nu\mu} - \nabla_\nu \delta\Gamma^\lambda_{\lambda\mu})\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 \\
&= \int_K \nabla_\lambda (g^{\mu\nu}\delta\Gamma^\lambda_{\nu\mu} - g^{\mu\lambda}\delta\Gamma^\tau_{\tau\mu})\sqrt{|g|}du^1 \wedge \cdots \wedge du^4
\end{aligned}$$

By the Theorem of Stokes this equals a boundary term which can be taken to be zero by considering appropriate K as discussed at the end of Section 3.2. Note that since $\Gamma^\lambda_{\nu\mu}$ contains derivatives of the metric, also the variations of the derivative of the metric need to be chosen zero on the boundary of K.

For general matrices M it holds that $\ln(\det(M)) = Tr(\ln(M))$. Varying this gives $\det(M)^{-1}\delta\det(M) = Tr(M^{-1}\delta M)$. Now taking as M the matrix representation $(g_{\mu\nu})_{\mu\nu}$ of the metric, this gives $g^{-1}\delta g = Tr(g^{\mu\nu}\delta g_{\lambda\rho}) = g^{\mu\nu}\delta g_{\mu\nu}$, where $g = \det((g_{\mu\nu})_{\mu\nu})$. Now varying the identity $g^{\mu\lambda}g_{\lambda\nu} = \delta^\mu_\nu$ gives:

$$0 = \delta\delta^\mu_\nu = \delta(g^{\mu\lambda}g_{\lambda\nu}) = (\delta g^{\mu\lambda})g_{\lambda\nu} + g^{\mu\lambda}\delta g_{\lambda\nu}$$

since the Kronecker delta δ^μ_ν is constant. Plugging this into the previous equation gives:

$$\begin{aligned}
\delta\sqrt{|g|} &= sgn(g)\frac{1}{2\sqrt{|g|}}\delta g = sgn(g)\frac{g}{2\sqrt{|g|}}g^{\mu\nu}\delta g_{\mu\nu} \\
&= -sgn(g)\frac{g}{2\sqrt{|g|}}g_{\mu\nu}\delta g^{\mu\nu} = -\frac{1}{2}\sqrt{|g|}g_{\mu\nu}\delta g^{\mu\nu}
\end{aligned}$$

The result of these manipulations of the three terms in the integral is thus:

$$\begin{aligned}
&\delta\int_K R\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 \\
&= \int_K ((\delta g^{\mu\nu})R_{\mu\nu}\sqrt{|g|} - \frac{1}{2}\sqrt{|g|}g_{\mu\nu}R\delta g^{\mu\nu})du^1 \wedge \cdots \wedge du^4 \\
&= \int_K (R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R)\delta g^{\mu\nu}\sqrt{|g|}du^1 \wedge \cdots \wedge du^4
\end{aligned}$$

Now as a last step the energy-momentum tensor $T_{\mu\nu}$ is defined by the relation:

$$-2T_{\mu\nu}\delta g^{\mu\nu} = L_M\sqrt{|g|}$$

Its precise form thus depends on the dependence of the matter Lagrangian on the metric. For different types of matter this dependency is in general also different. The energy-momentum tensor is thus dependent of the type of matter considered.

Given this definition the Principle of Least Action states that:

$$0 = \delta S_K^\Phi = \frac{c^4}{16\pi G}\delta \int\limits_K R\sqrt{|g|}du^1 \wedge \cdots \wedge du^4 + \delta \int\limits_K L_M\sqrt{|g|}du^1 \wedge \cdots \wedge du^4$$
$$= \int\limits_K (\frac{c^4}{16\pi G}(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R) - 2T_{\mu\nu})\delta g^{\mu\nu}\sqrt{|g|}du^1 \wedge \cdots \wedge du^4$$

Since this equation must hold for all variations $\delta g^{\mu\nu}$ it follows that:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}$$

This equation is called the Einstein Field Equation and is the equation of motion of General Relativity. Since the Ricci tensor consists of derivatives of the metric, the Einstein Field Equation consists of a set of 16 coupled differential equations. Given a specific matter Lagrangian, the solution metrics to this differential equations describe the force of gravity in the universe according to the theory of General Relativity.

Chapter 4

The Covariant Galileon Model

Observations show that General Relativity does not describe our Universe correctly. On smaller scales, of the order of the solar system for example, General Relativity gives a good description, but on very large scales the theory deviates from the observations. A new theory is therefore needed. The Covariant Galileon model gives such a new theory. The conception of this model starts with the observation that on very large scales the matter densities are nearly completely homogeneous and isotropic. The spatial part metric of the Universe giving rise to this matter distribution must therefore be also homogeneous and isotropic. The metric on these large scales is therefore of the form:

$$ds^2 = -dt^2 + (a(t))^2(dx^2 + dy^2 + dz^2)$$

where t denotes the temporal coordinate, x, y and z the spatial coordinates and the function $a\colon \mathbb{R} \to \mathbb{R}$ is called the scale factor. Moreover the scale factor is positive.

From the form of this metric it is now clear that the time coordinate t plays a special role. It is set apart from the other three coordinates. This means that the general symmetry in the four coordinates of General Relativity is broken for the temporal direction. There is no temporal symmetry any more, just spatial symmetry. Goldstone's Theorem now states that there is a scalar field ϕ associated to this symmetry breaking. The Covariant Galileon Model takes this as a starting point. It introduces an additional scalar field to the theory of General Relativity in order to amend for the discrepancies found between General Relativity and made observations. This additional field can be thought of as being some kind of strange not yet observed form of energy, a 'dark energy' if you will. This makes the Covariant Galileon Model part of the wider range of a dark energy models.

In this Chapter the Covariant Galileon Model shall be set up, its equations of motion on the background will be given and solved and values for different parameters of the theory shall be discussed.

4.1 The model

The Covariant Galileon Model is a theory of gravity. As stated before it takes General Relativity as a basis and adds an additional scalar field ϕ, called the Galileon field. The mathematical basis of the theory is thus the same as that of General Relativity. The Universe is still described by a manifold with curvature such that the metric gives rise to gravity, only now the Galileon field gives rise to additional dynamics. The Galileon field can however not be any scalar field. The Covariant Galileon model imposes three conditions on the scalar field:

1. The equations of motion of the theory remain invariant under transformations of the form $\partial_\mu\phi \to \partial_\mu\phi + a_\mu$. This comes down to the idea that under everyday circumstances, i.e. objects having speeds much lower than the speed of light, the theory is invariant under Galilean transformations. Hence also the name of the theory.

2. The equations of motion arising in the theory are at most second order. This is to exclude the possibility of getting negative energies and other instabilities.

3. The equations of motion on small scales reduce to the case of General Relativity. This means that the effect of ϕ on the dynamics of the system is negligible on small scales.

These conditions heavily constrain the possibilities for the Lagrangian of the theory. To meet these constraints, the action that defines the Covariant Galileon Model is given by:

$$S = \int (\frac{c^4}{16\pi G}R - \frac{1}{2}\sum_{i=1}^{5} c_i L_i + L_M)\sqrt{g}d^4x$$

where c is the speed of light, G the gravitational constant, R the Ricci scalar, the c_i constants, L_M the matter Lagrangian and the Lagrangians

L_i are given by:

$$
\begin{aligned}
L_1 &= M^3\phi \\
L_2 &= X \\
L_3 &= \frac{2X}{M^3}\Box\phi \\
L_4 &= \frac{X}{M^6}(2(\Box\phi)^2 - 2(\nabla^\mu\nabla^\nu\phi)(\nabla_\mu\nabla_\nu\phi) - \frac{1}{2}RX) \\
L_5 &= \frac{X}{M^9}((\Box\phi)^3 - 3(\Box\phi)(\nabla^\mu\nabla^\nu\phi)(\nabla_\mu\nabla_\nu\phi) \\
&\quad + 2(\nabla_\mu\nabla^\nu\phi)(\nabla_\nu\nabla^\rho\phi)(\nabla_\rho\nabla^\mu\phi) - 6(\nabla_\mu\phi)(\nabla^\rho\phi)(\nabla^\mu\nabla^\nu\phi)G_{\nu\rho})
\end{aligned}
$$

where ϕ is the Galileon field, ∇ is the Levi-Civita connection, $X = \nabla_\mu\phi\nabla^\mu\phi$, $\Box\phi = \nabla_\mu\nabla^\mu\phi$, R is again the Ricci scalar and $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is the Einstein tensor. M is just a constant given by $M^3 = M_{pl}H_0^2$ where H_0 is the present day value of Hubble's constant and M_{pl} is the Planck mass given by $M_{pl}^2 = \frac{c^4}{8\pi G}$.

The Covariant Galileon Model is thus a theory of gravity that describes the Universe as a 4-dimensional oriented manifold M with metric g and associated Levi-Civita connection ∇ and with a function $\phi\colon M \to \mathbb{R}$ associated to the Universe. Moreover the Galileon field is such that its derivative $\nabla\phi$ is nowhere vanishing. The action that then further defines the Covariant Galileon Model is the one given above. Now note that if the L_i are considered part of the matter Lagrangian, then the action takes the same form as the action in the Theory of General Relativity. The dynamics of the system are thus given by the Einstein Field Equation only with a energy-momentum tensor modified such that it includes the Galileon field addition. Therefore the Galileon field can be thought of as a form of yet unknown energy, the 'dark energy' of the Covariant Galileon Model.

To solve the dynamics of this system in general is really hard. To simplify calculations only the background of the system will be considered. The background of the system is the spatial average of the system. Here the dynamical variables g and ϕ are thus replaced with their spatial averages. As stated in the introduction of this chapter it follows from observations that the metric on this scale is of the form:

$$ds^2 = -dt^2 + (a(t))^2(dx^2 + dy^2 + dz^2)$$

So the whole dynamics arising from the (16 components of the) metric on the background reduces to the dynamics arising from just the scale factor a. Since the scale factor is not directly observable, Hubble's constant

H, which is directly measurable, shall be used as a variable instead of it. Hubble's constant is defined by:

$$H = \frac{\dot{a}}{a}$$

This switch of variable is not problematic since the scale factor can be recovered from Hubble's constant via:

$$a = \exp(\int H dt)$$

Moreover the Galileon field ϕ can on the background effectively be seen as just a function of time, since it is constant in space by definition of considering only the spatial average. The Galileon field on the background is denoted by $\phi_0 \colon \mathbb{R} \to \mathbb{R}$ and on the background it thus holds $\phi((t, x, y, z)) = \phi_0(t)$ given a point $p \in M$ with coordinates $p = (t, x, y, z)$. Note that on the background the covariant derivative of ϕ reduces to $\nabla_\mu \phi = \frac{\partial \phi}{\partial t} = \dot{\phi}_0$.

Using the Principle of Least Action the equations of motion for the Covariant Galileon Model can now be found. On the background they are given by:

$$\begin{aligned}
0 &= R + 3H^2 - \kappa\rho \\
0 &= 3\dot{H} + 3H^2 + \frac{\kappa}{2}(\rho + 3p) \\
0 &= c_2(\ddot{\phi}_0 + 3H\dot{\phi}_0) + \frac{c_3}{M^3}(12H\dot{\phi}_0\ddot{\phi}_0 + 18H^2\dot{\phi}_0^2 + 6\dot{H}\dot{\phi}_0^2) \\
&\quad + \frac{c_4}{M^6}(54H^2\dot{\phi}_0^2\ddot{\phi}_0 + 36H\dot{H}\dot{\phi}_0^3 + 54H^3\dot{\phi}_0^3) \\
&\quad + \frac{c_5}{M^9}(45H^4\dot{\phi}_0^4 + 60H^3\dot{\phi}_0^3\ddot{\phi}_0 + 45H^2\dot{H}\dot{\phi}_0^4)
\end{aligned}$$

where R is the Ricci scalar, ρ and p are respectively the energy density and the pressure of the matter in the Universe, the c's are the ones from the action and κ is a constant given by $\kappa = M_{pl}^{-2}$.

Considering the Galileon field as dark energy, four types of energy can be distinguished in the Universe as described by the Covariant Galileon Model. Those are the energies associated to normal matter, radiation, neutrino's and the Galileon field. Their energy densities and pressures are on the background given by:

$$\begin{aligned}
\rho_m &= \frac{\rho_{m,0}}{a^3} & p_m &= 0 \\
\rho_r &= \frac{\rho_{r,0}}{a^4} & p_r &= \frac{1}{3}\rho_r = \frac{\rho_{r,0}}{3a^4} \\
\rho_\nu &= 0 & p_\nu &= 0
\end{aligned}$$

and

$$\begin{aligned}\rho_\phi =& \frac{1}{2}c_2\dot{\phi}_0^2 + 6\frac{c_3}{M^3}H\dot{\phi}_0^3 + \frac{45}{2}\cdot\frac{c_4}{M^6}H^2\dot{\phi}_0^4 + 21\frac{c_5}{M^9}H^3\dot{\phi}_0^5 \\ p_\phi =& \frac{1}{2}c_2\dot{\phi}_0^2 - 2\frac{c_3}{M^3}\dot{\phi}_0^2\ddot{\phi}_0 - \frac{3}{2}\cdot\frac{c_4}{M^6}(8H\dot{\phi}_0^3\ddot{\phi}_0 + 2\dot{H}\dot{\phi}_0^4 + 3H^2\dot{\phi}_0^4) \\ & - 3\frac{c_5}{M^9}(5H^2\dot{\phi}_0^4\ddot{\phi}_0 + 2H\dot{H}\dot{\phi}_0^5 + 2H^3\dot{\phi}_0^5)\end{aligned}$$

where a subscript m denotes normal matter, r denotes radiation, ν denotes neutrino's and ϕ denotes the Galileon field and the additional subscript 0 in the energy density and pressure denotes the present value.

The derivation of the equations of motion and the energy density and pressure of the Galileon field can be found in Barreira et al. [7] and the derivation of the energy density and pressure of the other energy types can be found in Carroll [4].

4.2 The ADM formalism

It turns out to be numerically very difficult to solve the general (i.e. not on the background) equations of motion, without choosing a particular set of coordinates. This particular choice of coordinates and the setting which it creates, in which the numerical calculations are relatively easy to perform, is called the Arnowit-Deser-Misner formalism or ADM formalism. In this formalism the manifold is sliced up in hyperspaces and the coordinates are chosen to align with the hyperspaces. The idea is then that the hyperspaces are chosen in such a way that the equations on the hyperspaces become easier to work with. This approach will be used in this thesis in order to be able to incorporate the results later on in a more general scheme used by many other theories as well which facilitates making comparisons.

So the Covariant Galileon Model describes the Universe as a 4-dimensional manifold M with metric g and scalar field ϕ. Since the derivative $\nabla\phi$ is nowhere vanishing, the Galileon field can be used to slice up the manifold. This happens in the following way. First notice that given a local coordinate chart the Galileon field can be decomposed locally as $\phi(t,x,y,z) = \phi_0(t) + \delta\phi(t,x,y,z)$ where ϕ_0 is the average in the spatial coordinates of ϕ, i.e. the background value of ϕ, on the coordinate patch and $\delta\phi$ its deviation around that average. The coordinate patch can then be split up in level surfaces of ϕ_0 which are by construction spacelike. Such decompositions of ϕ also exist globally.

As an example consider the following. The metric, viewed as a matrix, is diagonalisable at every point, since it is diagonal in normal coordinates (see Kobayashi & Nomizu [1]). therefore there is a basis of eigenvectors of the metric for the tangent space of the manifold at every point. Since the metric is Lorentzian, 3 of those eigenvectors are spacelike and one is timelike. Moreover this basis of eigenvectors can be chosen to change differentiably, since the metric is differentiable. This thus constitutes 3 spacelike vector fields and one timelike vector field all linearly independent from each other. The integral curves of the 3 spacelike vector fields form a spacelike hypersurface Σ_t. By letting the start point of the integral curves change with a parameter t along the integral curves of the timelike vector field, a splitting of the manifold in spacelike hypersurfaces Σ_t is created. One can then take $\phi_0(t)$ to be the average of ϕ on Σ_t. This shows that there indeed exists a such decomposition of ϕ. Note that this decomposition is not unique. Moreover, since $\nabla\phi \neq 0$, the Σ_t can be reparametrised such that $\phi_0(p_0) = t$ for all $p = (p_0, p_1, p_{2,} p_3) \in \Sigma_t$. Furthermore, if the manifold is globally hyperbolic, then it follows that $t \in \mathbb{R}$ - and not some smaller interval contained in $\mathbb{R}$ - for such a decomposition (see Gourghoulhon [8]).

The level sets of this decomposition are submanifolds:

Lemma 4.2.1. *Σ_t is a 3-dimensional submanifold of M.*

Proof. Let $p \in M$ and (U, h, V) be a chart of M around p with associated coordinates (t, x, y, z). Then the rank of ϕ at p is given by

$$\begin{aligned} rk_p\phi &= rk(J(\phi \circ h^{-1})(h(p))) \\ &= rk((\frac{\partial\phi}{\partial t}(h(p)), \frac{\partial\phi}{\partial x}(h(p)), \frac{\partial\phi}{\partial y}(h(p)), \frac{\partial\phi}{\partial z}(h(p)))) \\ &= rk(\nabla\phi(h(p))) = 1 \end{aligned}$$

since $\nabla\phi \neq 0$ everywhere. This holds for all $p \in M$, so by the Rank Theorem (see Lübke [2]) there exists for all $p \in M$ a chart (U', h', V') of M around p such that:

$$\phi \circ h'^{-1}(x_1, x_2, x_3, x_4) = x_1 \in \mathbb{R}$$

Choosing this chart around $p \in \Sigma_t$ for some $t \in \mathbb{R}$ gives:

$$h'(U' \cap \Sigma_t) = h'(U' \cap \phi_0^{-1}(t)) = \{(x_1, x_2, x_3, x_4) \in V' \colon x_1 = t\}$$

Since the function given by $f((x_1, x_2, x_3, x_4)) = (x_2, x_3, x_4, x_1 - t)$ is a diffeomorphism, there exists a chart $(\hat{U}, \hat{h}, \hat{V})$ of M around p such that $\hat{U} = U'$ and:

$$\hat{h}(\hat{U} \cap \Sigma_t) = f \circ h'(U' \cap \phi^{-1}(t)) = \{(x_1, x_2, x_3, x_4) \in \hat{V} \colon x_4 = 0\}$$

This means that Σ_t is an 3-dimensional submanifold of M. □

Since the Σ_t are submanifolds they have their own intrinsic and extrinsic curvature tensor fields denoted respectively by${}^3R$ and K (the subscript 3 to distinguish from the intrinsic curvature of the whole manifold.) They also have an induced metric h and a corresponding Levi-Civita connection D.

Furthermore this means that the whole manifold can be given as a union of submanifolds:

$$M = \bigcup_{t \in \mathbb{R}} \Sigma_t$$

such that $\Sigma_t \cap \Sigma_r = \emptyset$ if $t \neq r$. This is called a foliation of M and each hypersurface Σ_t is called a leaf.

From now on we will consider the only the background case, so $\delta\phi = 0$, and we will identify ϕ with ϕ_0.

Now a normal vector to Σ_t can be defined as follows. Since ϕ is a scalar function, $\nabla\phi = \nabla_\mu\phi$ is a covector field, so $\vec{\nabla}\phi = g^{-1}(\cdot, \nabla\phi) = \nabla^\mu\phi$ is the associated vector field given by index raising. Let V be a vector field tangent to the Σ_t. Then it holds:

$$\begin{aligned} g(V, \vec{\nabla}\phi) &= g_{\mu\nu} V^\mu \nabla^\nu \phi = g_{\mu\nu} V^\mu g^{\nu\sigma} \nabla_\sigma \phi \\ &= \delta^\sigma_\mu V^\mu \nabla_\sigma \phi = V^\mu \nabla_\mu \phi = \nabla_V \phi = 0 \end{aligned}$$

since $\phi = \phi_0$ is constant on Σ_t by definition. So $\vec{\nabla}\phi$ is normal to Σ_t.

The lapse function $N \colon M \to \mathbb{R}$ can now be defined as:

$$N = (-g(\vec{\nabla}\phi, \vec{\nabla}\phi))^{-\frac{1}{2}}$$

Since $\vec{\nabla}\phi$ is timelike it follows that the lapse function is real. This defines a unit normal n as $n = N\vec{\nabla}\phi$. It is clear that n is normal to Σ_t. It is a unit vector because:

$$g(n, n) = g(N\vec{\nabla}\phi, N\vec{\nabla}\phi) = N^2 g(\vec{\nabla}\phi, \vec{\nabla}\phi) = \frac{g(\vec{\nabla}\phi, \vec{\nabla}\phi)}{-g(\vec{\nabla}\phi, \vec{\nabla}\phi)} = -1$$

This offers a physical interpretation of the foliation of M. Since n is a timelike unit vector it can be thought of as a velocity vector of some observer. Such observers are called Eulerian observers. The world lines of these observers are thus orthogonal to the leaves Σ_t of the foliation. This means that the Σ_t are locally the set of events that occur at the same time as seen from the perspective of the Eulerian observer. The acceleration a of the Eulerian observer is given by $a = \nabla_n n$. It holds:

$$g(a, n) = g(\nabla_n n, n) = \frac{1}{2}\nabla_n g(n, n) = \frac{1}{2}\nabla_n(-1) = 0$$

So a is tangent to the Σ_t.

Now also what is called the normal evolution vector m can be defined. It is given by $m = -Nn$. It holds:

$$\begin{aligned}\nabla_m \phi &= (\nabla\phi)(m) = d\phi(m) = -Nd\phi(n) \\ &= -N^2 d\phi(\vec{\nabla}\phi) = -N^2 g(\vec{\nabla}\phi, \vec{\nabla}\phi) = 1\end{aligned}$$

This means that the flow generated by m lets the Σ_t flow into each other.

Now a specific coordinate system is going to be introduced on M. Let $(u^i_t)_i = (u^1_t, u^2_t, u^3_t)$ be some local coordinate system for Σ_t with coordinate neighbourhood U_t such that it smoothly varies with t, meaning that the map $f\colon \bigcup_{t\in I} U_t \to \mathbb{R}^3$ with $I \subseteq \mathbb{R}$ given by $f(p) = (u^1_t(p_t), u^2_t(p_t), u^3_t(p_t))$, where p_t is defined by $p = (t, p_t)$, is differentiable w.r.t. the whole space M. Such coordinate systems exist at least locally, i.e. for $I = (a, b)$ for some $a, b \in \mathbb{R}$ with $a < b$. Then $(u^\mu)_\mu = (t, u^1_t, u^2_t, u^3_t)$ constitutes a local coordinate system for M. The subscript t is from now on dropped from the u's.

This particular coordinate system thus defines a basis:

$$\left(\frac{\partial}{\partial t}|_p, \frac{\partial}{\partial u^1}|_p, \frac{\partial}{\partial u^2}|_p, \frac{\partial}{\partial u^3}|_p\right)$$

for T_pM for all $p \in \bigcup_{t\in(a,b)} U_t$. The vector $\frac{\partial}{\partial t}|_p$ is called the time vector. Note that it is not necessarily timelike. Since $\phi(p) = t$ and t can be seen as a scalar field and ϕ and t can be identified with each other. In particular it thus holds:

$$\nabla_{\frac{\partial}{\partial t}}\phi = (\nabla\phi)\left(\frac{\partial}{\partial t}\right) = d\phi\left(\frac{\partial}{\partial t}\right) = 1$$

so the flow of generated by $\frac{\partial}{\partial t}$ lets the Σ_t flow into each other in the same way as m does. In general these vectors are however not equal. Their difference is called the shift vector denoted by $\beta = \frac{\partial}{\partial t} - m$. It holds:

$$g(n, \beta) = Ng(\vec{\nabla}\phi, \beta) = Nd\phi(\beta) = N(d\phi(\frac{\partial}{\partial t}) - d\phi(m)) = N(1-1) = 0$$

so β is tangent to the leaves Σ_t. Since $\frac{\partial}{\partial t}$ points along curves of fixed u^i, the shift vector β can thus be interpreted as the shift of the coordinate system as it is pushed along m.

Note that:

$$g(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}) = g(\beta + Nn, \beta + Nn) = g(\beta, \beta) + N^2 g(n, n) = g(\beta, \beta) - N^2$$

thus $\frac{\partial}{\partial t}$ can be spacelike, timelike or null.

This completes the description of the ADM formalism. The goal is now to describe tensors on M in terms of ADM quantities, i.e. in terms of the unit normal n, the lapse function N, the normal evolution vector m, the shift vector β and tensors on the leaves Σ_t. This will be done for the metric g, the intrinsic curvature R and the extrinsic curvature K.

To do this it is helpful to make the following observation. The tangent bundle TM can be decomposed as $TM = (\bigsqcup_{t\in\mathbb{R}} T\Sigma_t) \oplus Span(n)$, where the direct sum is taken fibrewise and $\bigsqcup$ denotes the disjoint union. This decomposition gives rise to a projection $\gamma\colon TM \to \bigsqcup_{t\in\mathbb{R}} T\Sigma_t$ given by $\gamma(v) = v + g(v, n)n$. This is the orthogonal projection onto the Σ_t. Its dual $\gamma^*\colon \bigsqcup_{t\in\mathbb{R}} T^*\Sigma_t \to T^*M$ is given by $\gamma^*(\omega)(v) = \omega(\gamma(v))$ for all $v \in TM$. For all k this can be extended to tensor bundles of type $(0, k)$ by entrywise application. γ^* can be used to extend tensors defined on the leaves Σ_t to the whole manifold M.

Consider now the induced metric h on the leaves $\bigsqcup_{t\in\mathbb{R}} \Sigma_t$. Let $\iota\colon \bigcup_{t\in\mathbb{R}} \Sigma_t \to M$ be the inclusion map, then $h = \iota^* g$ by Definition 2.6.6. Let $u, v \in \bigsqcup_{t\in\mathbb{R}} T\Sigma_t \subset TM$, then:

$$\begin{aligned}(\gamma^* h)(u, v) &= h(\gamma(u), \gamma(v)) = h(u + g(u, n)n, v + g(v, n)n) \\ &= h(u, v) = (\iota^* g)(u, v) = g(D\iota u, D\iota v) = g(u, v)\end{aligned}$$

Now let $v \in TM$ and $\lambda \in \mathbb{R}$, then:

$$\begin{aligned}(\gamma^* h)(\lambda n, v) &= h(\gamma(\lambda n), \gamma(v)) = h(\lambda n + g(\lambda n, n)n, v + g(v, n)n) \\ &= h(\lambda n - \lambda n, v + g(v, n)n) = 0\end{aligned}$$

From this it follows that:

$$\gamma^* h = g + g(\cdot, n) \otimes g(\cdot, n)$$

The induced metric gets now redefined as this extension, i.e. $h := \gamma^* h$. In components this gives:

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu$$

On the other hand the metric g can be written in terms of the lapse function and the shift vector as follows. First note that, since β is tangent to the leaves Σ_t, it holds $\beta = \beta^i \frac{\partial}{\partial u^i}$. This gives:

$$g_{00} = g(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}) = g(\beta, \beta) - N^2 = \beta_i \beta^i - N^2$$

Moreover:

$$\begin{aligned} g_{0i} &= g(\frac{\partial}{\partial t}, \frac{\partial}{\partial u^i}) = g(\beta, \frac{\partial}{\partial u^i}) - Ng(n, \frac{\partial}{\partial u^i}) \\ &= g(\beta, \frac{\partial}{\partial u^i}) = (\beta_j du^j)(\frac{\partial}{\partial u^i}) = \beta_j \delta_i^j = \beta_i \end{aligned}$$

and

$$g_{ij} = g(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}) = h(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}) - g(\frac{\partial}{\partial u^i}, n) \otimes g(\frac{\partial}{\partial u^j}, n) = h_{ij}$$

Thus:

$$g_{\mu\nu} = \begin{pmatrix} \beta_i \beta^i - N^2 & \beta_j \\ \beta_i & h_{ij} \end{pmatrix} \tag{4.2.1}$$

The metric inverse is then given by:

$$g^{\mu\nu} = \begin{pmatrix} -\frac{1}{N^2} & \frac{\beta^j}{N^2} \\ \frac{\beta^i}{N^2} & h^{ij} - \frac{\beta^i \beta^j}{N^2} \end{pmatrix}$$

Now for the extrinsic curvature K. For $u, v \in TM$ it holds:

$$\begin{aligned} \gamma^* K(u, v) &= K(\gamma(u), \gamma(v)) = g(D\iota(\gamma(u)), \nabla_{D\iota(\gamma(v))} n) \\ &= g(u + g(u, n)n, \nabla_{v + g(v,n)n} n) = g(u + g(u, n)n, \nabla_v n + g(v, n)a) \\ &= g(u, \nabla_v n) + g(u, n) g(n, \nabla_v n) \\ &\quad + g(v, n) g(u, a) + g(u, n) g(v, n) g(n, a) \\ &= g(u, \nabla_v n) + g(v, n) g(u, a) = g(u, \nabla n)(v) + g(v, n) g(u, a) \end{aligned}$$

From this it follows that, if K gets redefined as its extension, i.e. $K := \gamma^* K$:

$$K = g(\cdot, \nabla n)(\cdot) + g(\cdot, n) \otimes g(\cdot, a)$$

In components this gives:

$$K_{\mu\nu} = \nabla_\mu n_\nu + n_\mu a_\nu = \nabla_\mu n_\nu + n_\mu n^\sigma \nabla_\sigma n_\nu$$

For the intrinsic curvature it holds in components:

$$R + 2R_{\mu\nu}n^\mu n^\nu = {}^3R + K^2 - K_{\mu\nu}K^{\mu\nu}$$

This relation is called the Gauss equation and its derivation can be found in Gourghoulhon [8]. Moreover it holds:

$$\begin{aligned} R_{\mu\nu}n^\mu n^\nu &= R^\alpha{}_{\mu\alpha\nu}n^\mu n^\nu = (\nabla_\alpha \nabla_\nu n^\alpha - \nabla_\nu \nabla_\alpha n^\alpha)n^\nu \\ &= \nabla_\alpha(n^\nu \nabla_\nu n^\alpha) - \nabla_\nu n^\alpha \nabla_\alpha n^\nu - \nabla_\nu(n^\nu \nabla_\alpha n^\alpha) + \nabla_\alpha n^\alpha \nabla_\nu n^\nu \\ &= \nabla_\alpha a^\alpha - \nabla^\nu n^\alpha \nabla_\alpha n_\nu - \nabla_\nu(K n^\nu) + K^2 \\ &= \nabla_\alpha a^\alpha - K_{\mu\nu}K^{\mu\nu} - \nabla_\nu(K n^\nu) + K^2 \end{aligned}$$

So the Gauss equation can also be written as:

$$R + 2\nabla_\alpha a^\alpha - 2\nabla_\nu(K n^\nu) = {}^3R - K^2 + K_{\mu\nu}K^{\mu\nu}$$

4.3 The EFT formalism

The EFT formalism is a general way to treat a physical theory. This allows for easy comparison of different theories describing the same phenomena. The idea behind the EFT formalism is to Taylor expand the Lagrangian of the theory around some commonly agreed upon base case, called the background, of the theories, e.g. a low energy limit of the theories. The different theories can then be compared by comparing the coefficients of the variations of the various orders. These coefficients are also called the EFT functions. Here the formalism will be applied to dark energy theories.

The dark energy theories considered are all extensions of General Relativity and thus all abide to the basic mathematical theory behind General Relativity. Therefore they can all be cast into the ADM formalism. Following Frusciante et al. [9] the general Lagrangian of such a theory is of the form:

$$L = L(N, {}^3R, S, K, Z, U, Z_1, Z_2, \alpha_1, \dots, \alpha_5; t)$$

where the variables are defined as:

$$\begin{aligned}
S &= K_{\mu\nu}K^{\mu\nu} \\
Z &= {}^3R_{\mu\nu}{}^3R^{\mu\nu} \\
U &= {}^3R_{\mu\nu}K^{\mu\nu} \\
Z_1 &= D_i{}^3RD^{i3}R \\
Z_2 &= D_i{}^3R_{jk}D^{i3}R^{jk} \\
\alpha_1 &= a_ia^i \\
\alpha_2 &= a_iD_jD^ja^i \\
\alpha_3 &= {}^3RD_ia^i \\
\alpha_4 &= a_i(D_jD^j)^2a^i \\
\alpha_5 &= D_iD^{i3}RD_ja^j
\end{aligned}$$

The base case around which this Lagrangian will be varied is the FLRW metric, i.e. the case when the metric is of the form:

$$ds^2 = -dt^2 + a(t)^2\delta_{ij}du^idu^j$$

where $a\colon \mathbb{R} \to \mathbb{R}$ is the scale factor. By comparing this to equation 4.2.1 it follows that on the background it holds $N = 1$, $\beta = 0$ and $h_{ij} = a(t)^2\delta_{ij}$. This means in particular that $n = (1,0,0,0)$ and that $a = 0$, since its a derivative of n. Furthermore note that on the background it holds ${}^3R_{\mu\nu} = 0$, since h is flat. For the extrinsic curvature the following holds:

$$K_{\mu\nu} = \nabla_\nu n_\mu + a_\mu n_\nu = \frac{\partial n_\mu}{\partial u^\nu} - \Gamma^\sigma_{\nu\mu}n_\sigma = -\Gamma^\sigma_{\nu\mu}n_\sigma = -\Gamma^0_{\nu\mu}$$

since n is constant on the background. This gives:

$$K^\mu{}_\nu = g^{\mu\alpha}K_{\alpha\nu} = -g^{\mu\alpha}\Gamma^\sigma_{\nu\alpha}n_\sigma = -g^{\mu\alpha}g^{\sigma\beta}\Gamma_{\beta\nu\alpha}n_\sigma = -\Gamma^\mu_{\beta\nu}n^\beta = -\Gamma^\mu_{0\nu}$$

which in turn gives for the scalar extrinsic curvature:

$$\begin{aligned}
K = K^\mu{}_\mu = -\Gamma^\mu_{0\mu} &= -\frac{1}{2}g^{\mu\alpha}\left(\frac{\partial g_{\alpha 0}}{\partial u^\mu} + \frac{\partial g_{\mu\alpha}}{\partial u^0} - \frac{\partial g_{0\mu}}{\partial u^\alpha}\right) \\
&= -\frac{1}{2}g^{\mu\mu}\left(\frac{\partial g_{\mu 0}}{\partial u^\mu} + \frac{\partial g_{\mu\mu}}{\partial u^0} - \frac{\partial g_{0\mu}}{\partial u^\mu}\right) = -\frac{1}{2}g^{\mu\mu}\frac{\partial g_{\mu\mu}}{\partial u^0} \\
&= -\frac{3}{2}a^{-2}\frac{da^2}{dt} = -3\frac{\dot{a}}{a} = -3H
\end{aligned}$$

where there is made use of the fact that the FLRW metric is diagonal. Furthermore it also holds on the background:

$$
\begin{aligned}
S = K_{\mu\nu}K^{\mu\nu} &= \Gamma^0_{\nu\mu}g^{\nu\alpha}\Gamma^\mu_{0\alpha} = g^{\nu\nu}\Gamma^0_{\nu\mu}\Gamma^\mu_{0\nu} \\
&= \frac{1}{4}g^{\nu\nu}g^{0\alpha}(\frac{\partial g_{\alpha\nu}}{\partial u^\mu} + \frac{\partial g_{\mu\alpha}}{\partial u^\nu} - \frac{\partial g_{\nu\mu}}{\partial u^\alpha})g^{\mu\beta}(\frac{\partial g_{\beta 0}}{\partial u^\nu} + \frac{\partial g_{\nu\beta}}{\partial u^0} - \frac{\partial g_{0\nu}}{\partial u^\beta}) \\
&= \frac{1}{4}g^{\nu\nu}g^{00}(\frac{\partial g_{0\nu}}{\partial u^\mu} + \frac{\partial g_{\mu 0}}{\partial u^\nu} - \frac{\partial g_{\nu\mu}}{\partial u^0})g^{\mu\mu}(\frac{\partial g_{\mu 0}}{\partial u^\nu} + \frac{\partial g_{\nu\mu}}{\partial u^0} - \frac{\partial g_{0\nu}}{\partial u^\mu}) \\
&= -\frac{1}{4}g^{\nu\nu}g^{00}g^{\mu\mu}\frac{\partial g_{\nu\mu}}{\partial u^0}\cdot\frac{\partial g_{\nu\mu}}{\partial u^0} = \frac{1}{4}(g^{\mu\mu})^2(\frac{\partial g_{\mu\mu}}{\partial u^0})^2 = \frac{3}{4}a^{-4}(\frac{da^2}{dt})^2 = 3H^2
\end{aligned}
$$

Now following Frusciante et al. [9] the action of the theories in the EFT formalism up to second order variations is then given by:

$$
\begin{aligned}
S = \int d^4x\sqrt{-g}\Big(&\frac{m_0^2}{2}(1+\Omega(t))R + \Lambda(t) - c(t)\delta g^{00} + \frac{M_2^4(t)}{2}(\delta g^{00})^2 \\
&- \frac{\bar{M}_1^3(t)}{2}\delta g^{00}\delta K - \frac{\bar{M}_2^2(t)}{2}(\delta K)^2 - \frac{\bar{M}_3^2(t)}{2}\delta K^\mu{}_\nu\delta K^\nu{}_\mu \\
&+ \frac{\hat{M}^2(t)}{2}\delta g^{00}\delta^3 R + m_2^2(t)h^{\mu\nu}\partial_\mu g^{00}\partial_\nu g^{00} + \frac{\bar{m}_5(t)}{2}\delta^3 R\delta K \\
&+ \lambda_1(t)(\delta^3 R)^2 + \lambda_2(t)\delta^3 R^\mu{}_\nu\delta^3 R^\nu{}_\mu + \lambda_3(t)\delta^3 R h^{\mu\nu}\nabla_\mu\partial_\nu g^{00} \\
&+ \lambda_4(t)h^{\mu\nu}\partial_\mu g^{00}\nabla_\rho\nabla^\rho\partial_\nu g^{00} + \lambda_5(t)h^{\mu\nu}\nabla_\mu{}^3R\nabla_\nu{}^3R \\
&+ \lambda_6(t)h^{\mu\nu}\nabla_\mu{}^3R_{ij}\nabla_\nu{}^3R^{ij} + \lambda_7(t)h^{\mu\nu}\partial_\mu g^{00}(\nabla_\rho\nabla^\rho)^2\partial_\nu g^{00} \\
&+ \lambda_8(t)h^{\mu\nu}\nabla_\rho\nabla^{\rho 3}R\nabla_\mu\partial_\nu g^{00}\Big)
\end{aligned}
$$

Now let a subscript denote a partial derivative, i.e.L_X denotes the derivative $\frac{\partial L}{\partial X}$. Then the following functions are defined:

$$
\begin{aligned}
\mathcal{A} &= L_{KK} + 4H^2 L_{SS} - 4HL_{SK} \\
\mathcal{B} &= L_{KN} - 2HL_{SN} \\
\mathcal{C} &= L_{K^3R} - 2HL_{S^3R} + \frac{1}{2}L_U - HL_{KU} + 2H^2 L_{SU} \\
\mathcal{D} &= L_{N^3R} + \frac{1}{2}\dot{L}_U - HL_{NU} \\
\mathcal{E} &= L_{^3R} - \frac{3}{2}HL_U - \frac{1}{2}\dot{L}_U \\
\mathcal{F} &= L_K - 2HL_S
\end{aligned}
$$

Then Frusciante et al. [9] shows that the following EFT functions are given by:

$$
\begin{aligned}
\bar{M}_2^2(t) &= -\mathcal{A} - 2\mathcal{E} \\
M_2^4(t) &= \frac{3}{4}L_N + \frac{1}{4}L_{NN} - \frac{1}{4}\dot{\mathcal{F}} - \frac{1}{2}H\dot{\mathcal{E}} + \frac{1}{2}\ddot{\mathcal{E}} + \mathcal{E}\dot{H} \\
\bar{M}_1^3(t) &= -\mathcal{B} - 2\dot{\mathcal{E}} \\
\bar{M}_3^2(t) &= -2L_S + 2\mathcal{E} \\
m_2^2(t) &= \frac{1}{4}L_{\alpha_1} \\
\bar{m}_5(t) &= 2\mathcal{C} \\
\hat{M}^2(t) &= \mathcal{D}
\end{aligned}
$$

These are the EFT functions of that will be discussed further on in the thesis. The expressions of the other EFT function can be found in Frusciante et al. [9].

4.4 The mapping

In this section the EFT functions for the Covariant Galileon Model on the background will be derived. In order to do this, the Lagrangian of the model will be written in terms of ADM quantities and then compared to the EFT action. By immediately making the comparison with the EFT action, it follows that only the functions L_1 to L_5 need to be expressed in terms of ADM quantities. The first term is namely already a part of the EFT action and the matter Lagrangian can be disregarded in this process. The functions are given by:

$$
\begin{aligned}
L_1 &= M^3\phi \\
L_2 &= X \\
L_3 &= \frac{2X}{M^3}\Box\phi \\
L_4 &= \frac{X}{M^6}(2(\Box\phi)^2 - 2(\nabla^\mu\nabla^\nu\phi)(\nabla_\mu\nabla_\nu\phi) - \frac{1}{2}RX) \\
L_5 &= \frac{X}{M^9}((\Box\phi)^3 - 3(\Box\phi)(\nabla^\mu\nabla^\nu\phi)(\nabla_\mu\nabla_\nu\phi) \\
&\quad + 2(\nabla_\mu\nabla^\nu\phi)(\nabla_\nu\nabla^\rho\phi)(\nabla_\rho\nabla^\mu\phi) - 6(\nabla_\mu\phi)(\nabla^\rho\phi)(\nabla^\mu\nabla^\nu\phi)G_{\nu\rho})
\end{aligned}
$$

Looking at L_1, it already is written in terms of ADM quantities, since the only variable is the Galileon field. Since ϕ is the only variable of L_1, it follows that L_1 does not contribute to the EFT functions under consideration.

The only variable of L_2 is X, which, when identifying ϕ with ϕ_0, can be written as:

$$X = \nabla^\mu \phi \nabla_\mu \phi = g^{\mu\nu} \nabla_\nu \phi \nabla_\mu \phi = g^{00} (\nabla_0 \phi)^2 = -\frac{\dot{\phi}^2}{N^2}$$

in terms of ADM quantities. From this it follows that L_2 also has no non-zero contributions to the EFT functions that are considered.

For the L_3 and L_4 cases the following relations are needed:

$$\begin{aligned} \nabla_\mu N^{-1} &= -N^{-2} \nabla_\mu N = -N^{-2} \nabla_\mu (-X)^{-\frac{1}{2}} \\ &= -\frac{1}{2N^2} (-X)^{-1\frac{1}{2}} \nabla_\mu X = -\frac{1}{2} N \nabla_\mu X \end{aligned}$$

which gives:

$$\begin{aligned} \nabla_\mu \nabla_\nu \phi &= \nabla_\mu \frac{n_\nu}{N} = N^{-1} \nabla_\mu n_\nu + n_\nu \nabla_\mu N^{-1} \\ &= N^{-1} (K_{\mu\nu} - n_\mu a_\nu) - \frac{1}{2} N n_\nu \nabla_\mu X \\ &= N^{-1} (K_{\mu\nu} - n_\mu a_\nu) - \frac{1}{2} N n_\nu \nabla_\mu (\nabla^\lambda \phi \nabla_\lambda \phi) \\ &= N^{-1} (K_{\mu\nu} - n_\mu a_\nu) - N n_\nu \nabla^\lambda \phi \nabla_\lambda \nabla_\mu \phi \\ &= N^{-1} (K_{\mu\nu} - n_\mu a_\nu) - N n_\nu \nabla^\lambda \phi \nabla_\lambda \frac{n_\mu}{N} \\ &= N^{-1} (K_{\mu\nu} - n_\mu a_\nu) - n_\nu \nabla^\lambda \phi \nabla_\lambda n_\mu - N n_\nu \nabla^\lambda \phi n_\mu \nabla_\lambda N^{-1} \\ &= N^{-1} (K_{\mu\nu} - n_\mu a_\nu) - N^{-1} n_\nu n^\lambda \nabla_\lambda n_\mu + \frac{1}{2} N^2 n_\nu (\nabla^\lambda \phi) n_\mu \nabla_\lambda X \\ &= N^{-1} (K_{\mu\nu} - n_\mu a_\nu - n_\nu a_\mu) + \frac{1}{2} N^2 (\nabla^\lambda \phi \nabla_\lambda X) n_\nu n_\mu \end{aligned}$$

and its contraction:

$$\Box \phi = \nabla^\mu \nabla_\mu \phi = N^{-1} K - \frac{1}{2} N^2 \nabla^\lambda \phi \nabla_\lambda X$$

where there is made use of the fact that n and a are orthogonal. Furthermore, using the fact that:

$$\begin{aligned} K_{\mu\nu} n^\mu &= (\nabla_\mu n_\nu + n_\mu a_\nu) n^\mu = a_\nu - a_\nu = 0 \\ K_{\mu\nu} n^\nu &= K_{\nu\mu} n^\nu = K_{\mu\nu} n^\mu = 0 \end{aligned}$$

this gives:

$$\begin{aligned}\nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi =&(N^{-1}(K_{\mu\nu}-n_\mu a_\nu-n_\nu a_\mu)+\frac{1}{2}N^2(\nabla^\lambda\phi\nabla_\lambda X)n_\nu n_\mu))\\ &\cdot((N^{-1}(K^{\mu\nu}-n^\mu a^\nu-n^\nu a^\mu)+\frac{1}{2}N^2(\nabla^\lambda\phi\nabla_\lambda X)n^\mu n^\nu\\ =&N^{-2}(S-2a^\mu a_\mu)+\frac{1}{4}N^4(\nabla^\lambda\phi\nabla_\lambda X)^2\end{aligned}$$

For L_3 the following relation then holds:

$$\begin{aligned}L_3 &= \frac{2X}{M^3}\Box\phi = \frac{2X}{M^3}(N^{-1}K-\frac{1}{2}N^2\nabla^\lambda\phi\nabla_\lambda X)\\ &= \frac{2}{M^3}X\sqrt{-X}K+\frac{1}{M^3}\nabla^\lambda\phi\nabla_\lambda X = \frac{-2}{M^3}(-X)^{\frac{3}{2}}K-\frac{X}{M^3}\Box\phi\\ &= -\frac{1}{2}L_3-\frac{2}{M^3}(-X)^{\frac{3}{2}}K\end{aligned}$$

where there is made use of partial integration to move the covariant derivative from X to ϕ (and the boundary terms are neglected.) Hence in ADM quantities:

$$L_3 = -\frac{4}{3M^3}(-X)^{\frac{3}{2}}K = -\frac{4}{3M^3}\cdot\frac{\dot\phi^3}{N^3}K$$

The nonzero contributions to the EFT functions given by L_3 are thus:

$$\begin{aligned}M_2^4(t) &= \frac{3}{4}L_{3N}+\frac{1}{4}L_{3NN}-\frac{1}{4}\dot L_{3K}\\ &= \frac{3}{M^3}\cdot\frac{\dot\phi^3}{N^4}K-\frac{4}{M^3}\cdot\frac{\dot\phi^3}{N^5}K+\frac{1}{M^3}\cdot\frac{\dot\phi^2}{N^3}\ddot\phi-\frac{1}{M^3}\cdot\frac{\dot\phi^3}{N^4}\dot N\\ \bar M_1^3(t) &= -L_{3KN} = -\frac{4}{M^3}\cdot\frac{\dot\phi^3}{N^4}\end{aligned}$$

On the background they are thus given by:

$$\begin{aligned}M_2^4(t) &= -\frac{9}{M^3}H\dot\phi^3+\frac{12}{M^3}H\dot\phi^3+\frac{1}{M^3}\dot\phi^2\ddot\phi = \frac{3}{M^3}H\dot\phi^3+\frac{1}{M^3}\dot\phi^2\ddot\phi\\ \bar M_1^3(t) &= -\frac{4}{M^3}\dot\phi^3\end{aligned}$$

Now for L_4 consider a function f of the form:

$$f(g,\phi) = G_4(\phi,X)R-2G_{4X}(\phi,X)((\Box\phi)^2-\nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi)$$

where G_4 is a differentiable function and G_{4X} is its derivative w.r.t. X. Noting that:

$$\begin{aligned} a^\mu \nabla_\mu X &= a^\mu \nabla_\mu (\nabla^\lambda \phi \nabla_\lambda \phi) = 2a^\mu \nabla^\lambda \phi \nabla_\mu \nabla_\lambda \phi = 2a^\mu \nabla^\lambda \phi \nabla_\lambda \nabla_\mu \phi \\ &= 2a^\mu N^{-1} n^\lambda \nabla_\lambda \frac{n_\mu}{N} = 2a^\mu N^{-1} n^\lambda (n_\mu \nabla_\lambda N^{-1} + N^{-1} \nabla_\lambda n_\mu) \\ &= 2a^\mu N^{-2} n^\lambda \nabla_\lambda n_\mu = 2N^{-2} a^\mu a_\mu \end{aligned}$$

f can be written as:

$$\begin{aligned} f(g,\phi) &= G_4 R - 2G_{4X}((N^{-1}K - \frac{1}{2}N^2 \nabla^\lambda \phi \nabla_\lambda X)^2 \\ &\quad - N^{-2}(S - 2a^\mu a_\mu) - \frac{1}{4}N^4 (\nabla^\lambda \phi \nabla_\lambda X)^2) \\ &= G_4 R + 2G_{4X}(NK \nabla^\lambda \phi \nabla_\lambda X + N^{-2}(S - K^2 - 2a^\mu a_\mu)) \\ &= G_4 R + 2G_{4X}((Kn^\lambda - a^\lambda)\nabla_\lambda X + N^{-2}(S - K^2)) \\ &= G_4 R + 2XG_{4X}(K^2 - S) + 2G_{4X}(Kn^\lambda - a^\lambda)\nabla_\lambda X \end{aligned}$$

Now, using $\nabla_\mu G_4 = G_{4\phi} \nabla_\mu \phi + G_{4X} \nabla_\mu X$, this gives:

$$\begin{aligned} f &= G_4 R + 2XG_{4X}(K^2 - S) + 2G_{4X}(Kn^\lambda - a^\lambda)\nabla_\lambda X \\ &= G_4 R + 2XG_{4X}(K^2 - S) + 2(\nabla_\lambda G_4 - G_{4\phi} \nabla_\lambda \phi)(Kn^\lambda - a^\lambda) \\ &= G_4 R + 2XG_{4X}(K^2 - S) + 2(\nabla_\lambda G_4 - G_{4\phi} N^{-1} n_\lambda)(Kn^\lambda - a^\lambda) \\ &= G_4 R + 2XG_{4X}(K^2 - S) + 2(\nabla_\lambda G_4)(Kn^\lambda - a^\lambda) + 2G_{4\phi} N^{-1} K \\ &= G_4 R + 2XG_{4X}(K^2 - S) - 2G_4 \nabla_\lambda (Kn^\lambda - a^\lambda) + 2G_{4\phi} N^{-1} K \end{aligned}$$

where in the last line partial integration was used. Using the Gauss equation gives:

$$\begin{aligned} f &= G_4 R + 2XG_{4X}(K^2 - S) - 2G_4 \nabla_\lambda (Kn^\lambda - a^\lambda) + 2G_{4\phi} N^{-1} K \\ &= G_4 (R - 2\nabla_\lambda (Kn^\lambda - a^\lambda)) + 2XG_{4X}(K^2 - S) + 2G_{4\phi} N^{-1} K \\ &= G_4 ({}^3R - K^2 + K_{\mu\nu} K^{\mu\nu}) + 2XG_{4X}(K^2 - S) + 2G_{4\phi} N^{-1} K \\ &= G_4 {}^3R + (2XG_{4X} - G_4)(K^2 - S) + 2G_{4\phi} \sqrt{-X} K \end{aligned}$$

If $G_4 = -\frac{1}{2}M^{-6}X^2$ then f equals L_4. This means that L_4 can be written in terms of ADM quantities as:

$$\begin{aligned}
L_4 &= -\frac{1}{2M^6}X^2 \cdot {}^3R + (-\frac{2}{M^6}X^2 + \frac{1}{2M^6}X^2)(K^2 - S) \\
&= -\frac{1}{2M^6}X^2({}^3R + 3(K^2 - S)) \\
&= -\frac{1}{2M^6} \cdot \frac{\dot{\phi}^4}{N^4}({}^3R + 3(K^2 - S))
\end{aligned}$$

The non-zero contributions from L_4 to the EFT functions are thus:

$$\begin{aligned}
\bar{M}_2^2 &= -L_{4KK} - 2L_{4^3R} = \frac{3}{M^6} \cdot \frac{\dot{\phi}^4}{N^4} + \frac{1}{M^6} \cdot \frac{\dot{\phi}^4}{N^4} = \frac{4}{M^6} \cdot \frac{\dot{\phi}^4}{N^4} \\
\bar{M}_1^3 &= -L_{4KN} + 2HL_{4SN} - 2\dot{L}_{4^3R} \\
&= -\frac{12}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}K - \frac{12}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}H + \frac{4}{M^6} \cdot \frac{\dot{\phi}^3}{N^4}\ddot{\phi} - \frac{4}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}\dot{N} \\
\bar{M}_3^2 &= -2L_S + 2L_{^3R} = -\frac{3}{M^6} \cdot \frac{\dot{\phi}^4}{N^4} - \frac{1}{M^6} \cdot \frac{\dot{\phi}^4}{N^4} = -\frac{4}{M^6} \cdot \frac{\dot{\phi}^4}{N^4} \\
\hat{M}^2 &= L_{4N^3R} = \frac{2}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}
\end{aligned}$$

and

$$\begin{aligned}
M_2^4 =& \frac{3}{4}L_N + \frac{1}{4}L_{NN} - \frac{1}{4}\dot{L}_{4K} + \frac{1}{2}(H\dot{L}_{4S} + \dot{H}L_{4S}) - \frac{1}{2}H\dot{L}_{4^3R} + \frac{1}{2}\ddot{L}_{4^3R} + L_{4^3R}\dot{H} \\
=& \frac{3}{2M^6} \cdot \frac{\dot{\phi}^4}{N^5}({}^3R + 3(K^2 - S)) - \frac{5}{2M^6} \cdot \frac{\dot{\phi}^4}{N^6}({}^3R + 3(K^2 - S)) + \frac{3}{M^6} \cdot \frac{\dot{\phi}^3}{N^4}K\ddot{\phi} \\
& - \frac{3}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}K\dot{N} + \frac{3}{4M^6} \cdot \frac{\dot{\phi}^4}{N^4}\dot{K} + \frac{1}{2}(\frac{6}{M^6} \cdot \frac{\dot{\phi}^3}{N^4}H\ddot{\phi} - \frac{6}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}H\dot{N} \\
& + \frac{3}{2M^6} \cdot \frac{\dot{\phi}^4}{N^4}\dot{H}) + \frac{1}{M^6} \cdot \frac{\dot{\phi}^3}{N^4}H\ddot{\phi} - \frac{1}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}H\dot{N} - \frac{3}{M^6} \cdot \frac{\dot{\phi}^2}{N^4}\ddot{\phi}^2 \\
& + \frac{4}{M^6} \cdot \frac{\dot{\phi}^3}{N^5}\dot{N}\ddot{\phi} - \frac{1}{M^6} \cdot \frac{\dot{\phi}^3}{N^4}\dddot{\phi} + \frac{4}{M^6} \cdot \frac{\dot{\phi}^3}{N^5}\dot{N}\ddot{\phi} - \frac{5}{M^6} \cdot \frac{\dot{\phi}^4}{N^6}\dot{N}^2 \\
& + \frac{1}{M^6} \cdot \frac{\dot{\phi}^4}{N^5}\ddot{N} - \frac{1}{2M^6} \cdot \frac{\dot{\phi}^4}{N^4}\dot{H}
\end{aligned}$$

So on the background these contributions are given by:

$$\bar{M}_2^2 = \frac{4}{M^6}\dot{\phi}^4$$
$$\bar{M}_1^3 = \frac{36}{M^6}H\dot{\phi}^4 - \frac{12}{M^6}H\dot{\phi}^4 + \frac{4}{M^6}\dot{\phi}^3\ddot{\phi} = \frac{24}{M^6}H\dot{\phi}^4 + \frac{4}{M^6}\dot{\phi}^3\ddot{\phi}$$
$$\bar{M}_3^2 = -\frac{4}{M^6}\dot{\phi}^4$$
$$\hat{M}^2 = \frac{2}{M^6}\dot{\phi}^4$$

and

$$\begin{aligned} M_2^4 =& \frac{9}{2M^6}(9H^2 - 3H^2)\dot{\phi}^4 - \frac{15}{2M^6}(9H^2 - 3H^2)\dot{\phi}^4 \\ &- \frac{9}{M^6}H\dot{\phi}^3\ddot{\phi} - \frac{9}{4M^6}\dot{\phi}^4\dot{H} + \frac{1}{2}(\frac{6}{M^6}H\dot{\phi}^3\ddot{\phi} + \frac{3}{2M^6}\dot{\phi}^4\dot{H}) \\ &+ \frac{1}{M^6}H\dot{\phi}^3\ddot{\phi} - \frac{3}{M^6}\dot{\phi}^2\ddot{\phi}^2 - \frac{1}{M^6}\dot{\phi}^3\dddot{\phi} - \frac{1}{2M^6}\dot{\phi}^4\dot{H} \\ =& - \frac{18}{M^6}H^2\dot{\phi}^4 - \frac{5}{M^6}H\dot{\phi}^3\ddot{\phi} - \frac{2}{M^6}\dot{\phi}^4\dot{H} - \frac{3}{M^6}\dot{\phi}^2\ddot{\phi}^2 - \frac{1}{M^6}\dot{\phi}^3\dddot{\phi} \end{aligned}$$

Since the total Lagrangian of the system is given by $L = \frac{c^4}{16\pi G}R - \frac{1}{2}\sum_{i=1}^{5} c_i L_i + L_M$, the individual contributions can be summed to give the EFT functions of the system:

$$\bar{M}_2^2 = -\frac{2c_4}{M^6}\dot{\phi}^4$$
$$\begin{aligned} M_2^4 =& -\frac{c_3}{M^3}(\frac{3}{2}H\dot{\phi}^3 + \frac{1}{2}\dot{\phi}^2\ddot{\phi}) \\ &+ \frac{c_4}{M^6}(9H^2\dot{\phi}^4 + \frac{5}{2}H\dot{\phi}^3\ddot{\phi} + \dot{\phi}^4\dot{H} + \frac{3}{2}\dot{\phi}^2\ddot{\phi}^2 + \frac{1}{2}\dot{\phi}^3\dddot{\phi}) \end{aligned}$$
$$\bar{M}_1^3 = \frac{2c_3}{M^3}\dot{\phi}^3 - \frac{c_4}{M^6}(12H\dot{\phi}^4 + 2\dot{\phi}^3\ddot{\phi})$$
$$\bar{M}_3^2 = \frac{2c_4}{M^6}\dot{\phi}^4$$
$$m_2^2(t) = 0$$
$$\bar{m}_5(t) = 0$$
$$\hat{M}^2 = -\frac{c_4}{M^6}\dot{\phi}^4$$

in the case that $c_5 = 0$.

Chapter 5

Simulations

In this chapter the the Covariant Galileon model, as explained in the previous chapter, will be solved numerically and the corresponding EFT functions will be given.

5.1 The equations

As shown in Barreira et al. [7] the equations of motion of the model at the background are:

$$\begin{aligned}
0 &= 3H^2 - \kappa\rho \\
0 &= 3\dot{H} + 3H^2 + \frac{\kappa}{2}(\rho + 3p) \\
0 &= c_2(\ddot{\phi}_0 + 3H\dot{\phi}_0) + \frac{c_3}{M^3}(12H\dot{\phi}_0\ddot{\phi}_0 + 18H^2\dot{\phi}_0^2 + 6\dot{H}\dot{\phi}_0^2) \\
&\quad + \frac{c_4}{M^6}(54H^2\dot{\phi}_0^2\ddot{\phi}_0 + 36H\dot{H}\dot{\phi}_0^3 + 54H^3\dot{\phi}_0^3) \\
&\quad + \frac{c_5}{M^9}(45H^4\dot{\phi}_0^4 + 60H^3\dot{\phi}_0^3\ddot{\phi}_0 + 45H^2\dot{H}\dot{\phi}_0^4)
\end{aligned}$$

Note that $\hat{R} = 0$ on the background since the metric is flat. This is a system of three equations in the variables ϕ and H as a function of time. As such the system is thus overdetermined. To fix this the first equation will be used as a means to constrain one of the parameters of the model in stead of as a equation to evolve the system. Furthermore the energy density and pressure are given by:

$$\rho_m = \frac{\rho_{m,0}}{a^3} \qquad p_m = 0$$
$$\rho_r = \frac{\rho_{r,0}}{a^4} \qquad p_r = \frac{\rho_{r,0}}{3a^4}$$
$$\rho_\nu = 0 \qquad p_\nu = 0$$

and

$$\begin{aligned}
\rho_\phi =& \frac{1}{2}c_2\dot{\phi}^2 + 6\frac{c_3}{M^3}H\dot{\phi}^3 + \frac{45}{2}\cdot\frac{c_4}{M^6}H^2\dot{\phi}^4 + 21\frac{c_5}{M^9}H^3\dot{\phi}^5 \\
p_\phi =& \frac{1}{2}c_2\dot{\phi}^2 - 2\frac{c_3}{M^3}\dot{\phi}^2\ddot{\phi} - \frac{3}{2}\frac{c_4}{M^6}(8H\dot{\phi}^3\ddot{\phi} + 2\dot{H}\dot{\phi}^4 + 3H^2\dot{\phi}^4) \\
& - 3\frac{c_5}{M^9}(5H^2\dot{\phi}^4\ddot{\phi} + 2H\dot{H}\dot{\phi}^5 + 2H^3\dot{\phi}^5)
\end{aligned}$$

These equations make up the system that is to be solved.

In order to solve these equations it is beneficial to use the logarithm of the scale factor, $\ln(a)$, as variable in stead of time. Hence time derivatives transform as:

$$\frac{d}{dt} = \frac{da}{dt}\cdot\frac{d\ln(a)}{da}\cdot\frac{d}{d\ln(a)} = \frac{\dot{a}}{a}\cdot\frac{d}{d\ln(a)} = H\frac{d}{d\ln(a)}$$

and the derivative of a quantity X w.r.t. $\ln(a)$ will be denoted as $\mathring{X} = \frac{dX}{d\ln(a)}$. Since the equations only involve derivatives of ϕ, but not ϕ itself, it is useful to introduce the quantity $\psi = \dot{\phi}$. Then it holds e.g. $\ddot{\phi} = \dot{\psi} = H\mathring{\psi}$ and $\dddot{\phi} = H\frac{dH\mathring{\psi}}{d\ln(a)} = H\mathring{\psi}\mathring{H} + H^2\mathring{\mathring{\psi}}$.

To make the equations involved dimensionless the following dimensionless variables are introduced:

$$\bar{\phi} = \frac{\phi}{M_{pl}} \qquad \bar{\rho} = \frac{\rho}{3M_{pl}^2H_0^2}$$
$$\bar{H} = \frac{H}{H_0} \qquad \bar{p} = \frac{p}{M_{pl}^2H_0^2}$$
$$\bar{\psi} = \frac{\psi}{M_{pl}H_0} \qquad \Omega = \frac{\bar{\rho}}{\bar{H}^2} = \frac{\rho}{3M_{pl}^2H^2}$$

In terms of these new variables the equations become:

$$\bar{\rho} = \bar{H}^2 \tag{5.1.1}$$

$$0 = \bar{H}\mathring{\bar{H}} + \bar{H}^2 + \frac{1}{2}(\bar{\rho} + \bar{p}) \tag{5.1.2}$$

$$\begin{aligned} 0 = {} & c_2(\bar{H}\mathring{\bar{\psi}} + 3\bar{H}\bar{\psi}) + c_3(12\bar{H}^2\bar{\psi}\mathring{\bar{\psi}} + 18\bar{H}^2\bar{\psi}^2 + 6\bar{H}\mathring{\bar{H}}\bar{\psi}^2) \\ & + c_4(54\bar{H}^3\bar{\psi}^2\mathring{\bar{\psi}} + 36\bar{H}^2\mathring{\bar{H}}\bar{\psi}^3 + 54\bar{H}^3\bar{\psi}^3) \\ & + c_5(45\bar{H}^4\bar{\psi}^4 + 60\bar{H}^4\bar{\psi}^3\mathring{\bar{\psi}} + 45\bar{H}^3\mathring{\bar{H}}\bar{\psi}^4) \end{aligned} \tag{5.1.3}$$

$$\bar{\rho}_\phi = \frac{1}{6}c_2\bar{\psi}^2 + 2c_3\bar{H}\bar{\psi}^3 + \frac{15}{2}c_4\bar{H}^2\bar{\psi}^4 + 7c_5\bar{H}^3\bar{\psi}^5 \tag{5.1.4}$$

$$\begin{aligned} \bar{p}_\phi = {} & \frac{1}{2}c_2\bar{\psi}^2 - 2c_3\bar{H}\bar{\psi}^2\mathring{\bar{\psi}} - \frac{3}{2}c_4(8\bar{H}^2\bar{\psi}^3\mathring{\bar{\psi}} + 2\bar{H}\mathring{\bar{H}}\bar{\psi}^4 + 3\bar{H}^2\bar{\psi}^4) \\ & - 3c_5(5\bar{H}^3\bar{\psi}^4\mathring{\bar{\psi}} + 2\bar{H}^2\mathring{\bar{H}}\bar{\psi}^5 + 2\bar{H}^3\bar{\psi}^5) \end{aligned} \tag{5.1.5}$$

and

$$\begin{aligned} \bar{\rho}_m &= \frac{\Omega_{m,0}}{e^{3\ln(a)}} & \bar{p}_m &= 0 \\ \bar{\rho}_r &= \frac{\Omega_{r,0}}{e^{4\ln(a)}} & \bar{p}_r &= \frac{\Omega_{r,0}}{e^{4\ln(a)}} \\ \bar{\rho}_n &= 0 & \bar{p}_n &= 0 \end{aligned}$$

If now the following quantities are defined:

$$\begin{aligned} \alpha &= 6c_3\bar{H}\bar{\psi}^2 + 36c_4\bar{H}^2\bar{\psi}^3 + 45c_5\bar{H}^3\bar{\psi}^4 \\ \beta &= c_2\bar{H} + 12c_3\bar{H}^2\bar{\psi} + 54c_4\bar{H}^3\bar{\psi}^2 + 60c_5\bar{H}^4\bar{\psi}^3 \\ \gamma &= 3c_2\bar{H}\bar{\psi} + 18c_3\bar{H}^2\bar{\psi}^2 + 54c_4\bar{H}^3\bar{\psi}^3 + 45c_5\bar{H}^4\bar{\psi}^4 \\ \delta &= -3c_4\bar{H}\bar{\psi}^4 - 6c_5\bar{H}^2\bar{\psi}^5 \\ \epsilon &= -2c_3\bar{H}\bar{\psi}^2 - 12c_4\bar{H}^2\bar{\psi}^3 - 15c_5\bar{H}^3\bar{\psi}^4 \\ \zeta &= \frac{1}{2}c_2\bar{\psi}^2 - \frac{9}{2}c_4\bar{H}^2\bar{\psi}^4 - 6c_5\bar{H}^3\bar{\psi}^5 \\ \Pi &= \bar{\rho} + \bar{p} - \bar{p}_\phi \end{aligned}$$

then the equations of motion used to evolve the system, i.e. equations 5.1.2 till 5.1.5, can be expressed as:

$$\begin{aligned} \mathring{\bar{H}} &= \frac{\beta(2\bar{H}^2 + \Pi + \zeta) - \gamma\epsilon}{\alpha\epsilon - \beta(2\bar{H} + \delta)} \\ \mathring{\bar{\psi}} &= \frac{\alpha(2\bar{H}^2 + \Pi + \zeta) - \gamma(2\bar{H} + \delta)}{\beta(2\bar{H} + \delta) - \alpha\epsilon} \end{aligned}$$

Written in this way the system can be easily solved numerically.

Lastly it is useful to give the EFT functions in a dimensionless form. To this purpose the dimensionless EFT functions are defined by and in the quartic case given by:

$$\gamma_1 = \frac{M_2^4}{M_{pl}^2 H_0^2} = -c_3(\frac{3}{2}\bar{H}\bar{\psi}^3 + \frac{1}{2}\bar{H}\bar{\psi}^2\mathring{\bar{\psi}}) + c_4(9\bar{H}^2\bar{\psi}^4 + \frac{5}{2}\bar{H}^2\bar{\psi}^3\mathring{\bar{\psi}}$$
$$+ \bar{H}\bar{\psi}^4\mathring{\bar{H}} + \frac{3}{2}\bar{H}^2\bar{\psi}^2\mathring{\bar{\psi}}^2 + \frac{1}{2}\bar{H}\bar{\psi}^3\mathring{\bar{H}}\mathring{\bar{\psi}} + \frac{1}{2}\bar{H}^2\bar{\psi}^3\mathring{\mathring{\bar{\psi}}})$$
$$\gamma_2 = \frac{M_1^3}{M_{pl}^2 H_0} = 2c_3\bar{\psi}^3 - c_4(12\bar{H}\bar{\psi}^4 + 2\bar{H}\bar{\psi}^3\mathring{\bar{\psi}})$$
$$\gamma_3 = \frac{\bar{M}_2^2}{M_{pl}^2} = -2c_4\bar{\psi}^4$$
$$\gamma_4 = \frac{\bar{M}_3^2}{M_{pl}^2} = 2c_4\bar{\psi}^4$$
$$\gamma_5 = \frac{\hat{M}^2}{M_{pl}^2} = -c_4\bar{\psi}^4$$
$$\gamma_6 = \frac{m_2^2}{M_{pl}^2} = 0$$

5.2 The parameters

Now that the equations that have to be simulated are established, the values of the parameters of the system need to be determined in order to be able to perform the simulations. The Covariant Galileon Model contains twelve parameters:

$$\{c_2,\ c_3,\ c_4,\ c_5,\ \Omega_{m,0},\ \Omega_{r,0},\ \bar{H}_i,\ \bar{\psi}_i,\ M_{pl},\ H_0,\ z_i,\ z_e\}$$

where $\bar{H}_i$ and $\bar{\psi}_i$ denote the values of the dimensionless Hubble constant and time derivative of the Galileon field at the start of the simulation and z_i and z_e denote the start and end redshift used in the simulation. Note that there are two other parameters that implicitly have been taken to be zero, namely the neutrino energy density $\Omega_{\nu,0}$ - for observational reasons - and the coefficient c_1 - for mathematical reasons. The choice of the twelve parameters will closely follow the choice of Barreira et al. [7].

Leaving out z_i and z_e, a distinction can be made between two types of parameters. On the one hand there are the parameters that can be directly measured to find their value. Those are $\Omega_{m,0}$, $\Omega_{r,0}$, $\bar{H}_i$, M_{pl} and H_0. They obtain their value from established models of well understood phenomena. The other five parameters, i.e. c_2 to c_5 and $\bar{\psi}_i$, can not be determined by measurement directly, rather their values are determined by the predictions of this model. Their values are such that the predictions of the model are in agreement with the observations. The last two parameters z_i and z_e just determine the length of the simulation. The parameters c_2 to c_5 and $\bar{\psi}_i$ can thus in principle only be determined by running the model and comparing the results, the other parameters can be fixed beforehand.

Luckily this five dimensional parameter space can reduced by noting the following. Firstly, as stated in Barreira et al. [10], this system has an attractor, characterised by the so called tracker solution. This is a particular solution of the system to which many other neighbouring solutions are attracted. The tracker solution is determined by the relation $\bar{\psi}\bar{H} = \xi$, where $\xi \in \mathbb{R}$ is just some fixed number. This relation will be used to relate the two initial conditions to each other, thus replacing $\bar{\psi}_i$ with ξ as a variable. Secondly there is a scaling degeneracy in the system. If $\lambda \in \mathbb{R}$ is some number and the variables $\phi' = \lambda^{-1}\phi$ and $c_i' = \lambda^i c_i$ (no summation intended here) for $i \in \{2, \ldots, 5\}$ are introduced, then when using these variables instead of their unprimed counterparts the Lagrangian of the system retains its form, meaning the dynamics of the system remain unchanged. This freedom can be used to fix one of the parameters. As in Barreira et al. [10], here c_2 will be set to -1. Next, the fact that the present day Universe is flat can be used to fix one parameter by using the constraint equation 5.1.1. This can be done by choosing parameter values, running the simulation and in the end check whether or not the constraint equation is satisfied. If not the parameter values are changed and the model is run again. In this way c_3 is fixed. Lastly, assuming flatness at the starting time of the simulation, by using the constraint equation and the expression of the Galileon energy density, a relation between the initial conditions and the c_i coefficients can be found. This relation will be used to fix $\bar{\psi}_i$. This leaves two free parameters, namely c_4 and c_5.

5.3 The simulations

In this section the results of the simulations are shown. The model is simulated by using a Python script created for this purpose. The simulations concern the quartic Covariant Galileon Model, meaning that c_5 is fixed to 0, and the model will be run from initial redshift $z_i = 10^6$ to end redshift $z = 0$. Two sets of parameter values have been simulated. The first one is being given by:

$$c_3 = 0,10104 \qquad \Omega_{m0} = 0,249$$
$$c_4 = -4,4523 \cdot 10^{-3} \qquad \Omega_{r0} = 7,18 \cdot 10^{-5}$$
$$\xi = 2,43$$

This set of parameter values can also be found in Barreira et al. [10]. The evolution of the dimensionless Hubble constant $\bar{H}$ and the dimensionless time derivative of the Galileon field $\bar{\psi}$ are found to be:

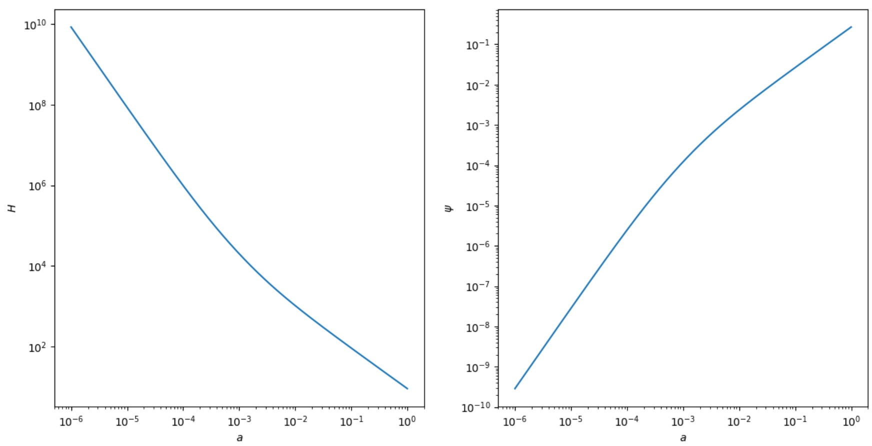

Figure 5.3.1: *The evolution of the dimensionless Hubble constant on the left and the dimensionless time derivative of the Galileon field on the right as a function of the scale factor in the quartic Covariant Galileon Model using parameter values as described directly above the figure.*

For this solution the simulated present day Hubble constant is 9,0, which thus makes it 9 times the observed present day Hubble constant. Furthermore for the constraint equation it holds $\bar{H}^2(1) - \bar{\rho}(1) = 82$.

The solution can be further characterised by the equation of state w of the Galileon field, which is defined as the pressure divided by the energy density. It is given by:

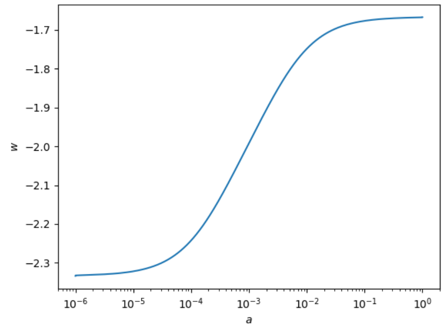

Figure 5.3.2: *The evolution of the equation of state as a function of the scale factor in the quartic Covariant Galileon Model using parameter values as described above.*

The found solutions can be compared to the tracker solution, the solution to which solutions for other parameter values are attracted. The tracker solution can be found in Barreira et al. [10]. Comparing this solution to the solution found by the full simulation of the equations gives:

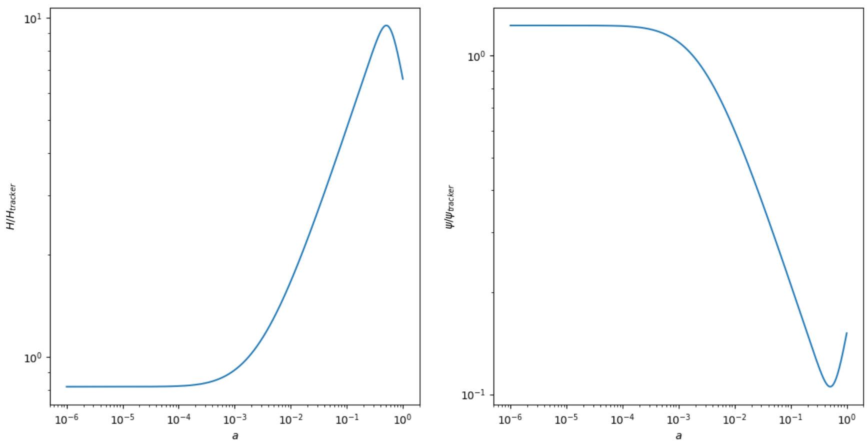

Figure 5.3.3: *The quotient of the Hubble constant as derived by the simulation and as given by the tracker solution on the left and the quotient of the time derivative of the Galileon field as derived by the simulation and as given by the tracker solution on the right as a function of the scale factor in the quartic Covariant Galileon Model using parameter values as described above.*

The EFT functions of this solution are given by:

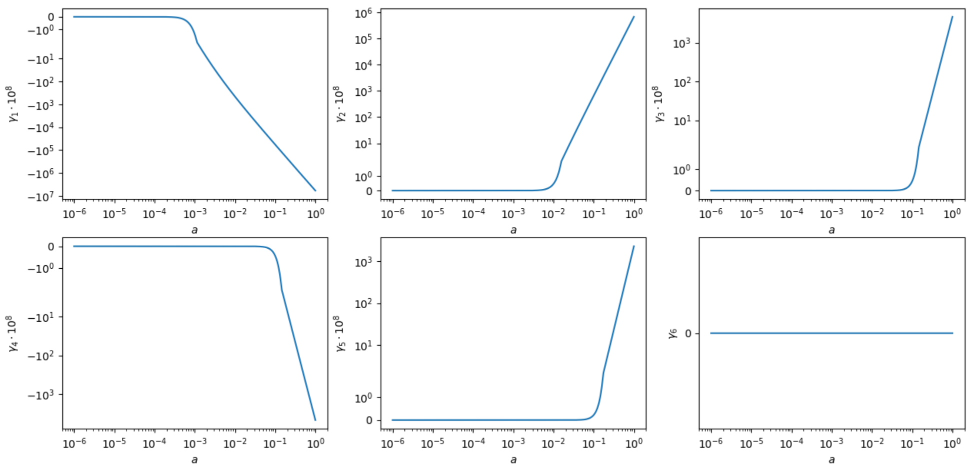

Figure 5.3.4: *The dimensionless EFT functions corresponding to the solution of the model given by the simulation in Figure 5.3.1 as a function of the scale factor in the quartic Covariant Galileon Model using parameter values as described above.*

To look at the sensitivity of the model to the parameter values chosen two parameters are varied. Here the present day matter energy density is varied in three steps between ±5% of its original value and the initial value of the (dimensionless) Hubble constant is varied in three steps between ±0,01% of its original value. This gives for the solutions to the model:

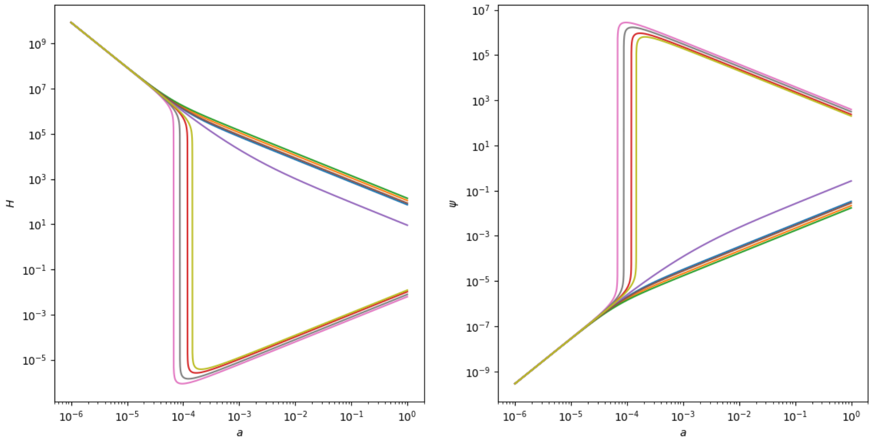

Figure 5.3.5: *The solutions, with on the left the dimensionless Hubble constant and on the right the dimensionless time derivative of the Galileon field, as a function of the scale factor with parameter values as above and with the present day matter energy density Ω_{m0} and the initial value of the Hubble constant $\bar{H}_i$ varied resulting in nine graphs. The offshoots to below on the left and to the top on the right are present in the solutions in which Ω_{m0} is bigger than its original value, or $\bar{H}_i$ is smaller than its original value and Ω_{m0} is at its original value.*

This gives as EFT functions:

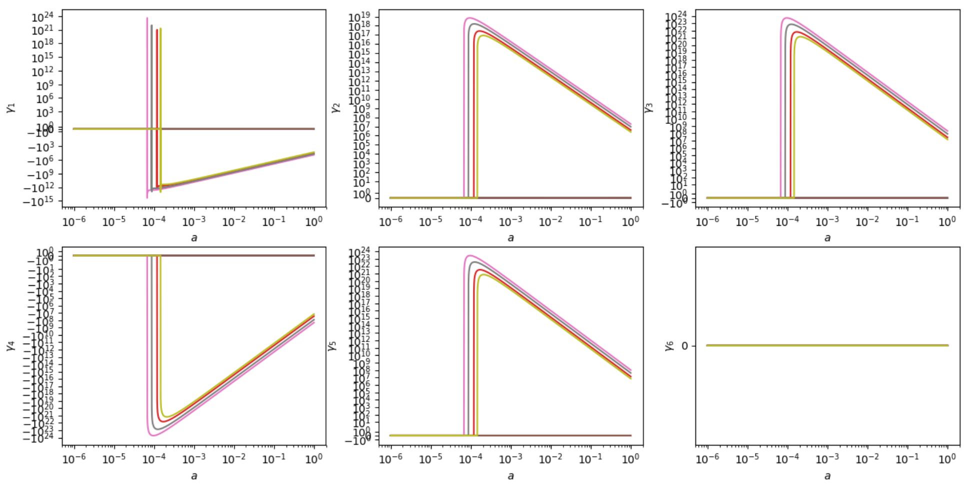

Figure 5.3.6: *The EFT functions corresponding to the solutions shown in the figure above.*

For the other parameter value set the results are very much the same. The parameter values now are:

$$c_3 = 0,1 \qquad \Omega_{m0} = 0,315$$
$$c_4 = -2,778 \cdot 10^{-3} \qquad \Omega_{r0} = 10^{-4}$$
$$\xi = 2$$

The solution of the model is now found to be:

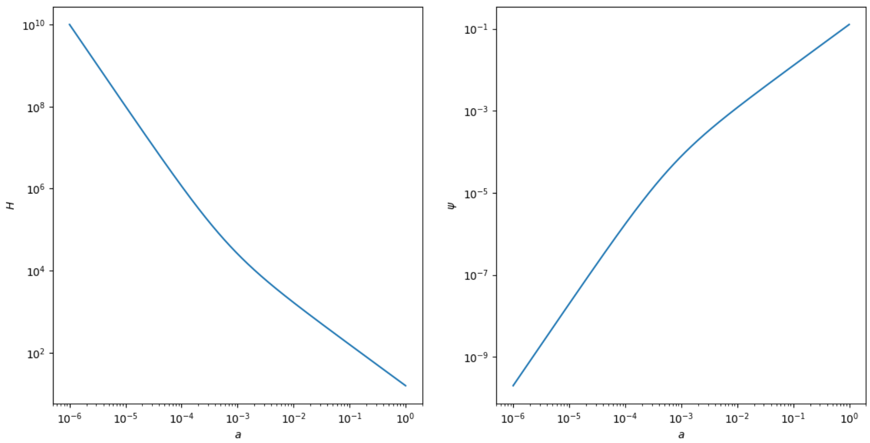

Figure 5.3.7: *The evolution of the dimensionless Hubble constant on the left and the dimensionless time derivative of the Galileon field on the right as a function of the scale factor in the quartic Covariant Galileon Model using parameter values as described directly above the figure.*

For this solution the simulated present day dimensionless Hubble constant is approximately 16, which thus given a present day Hubble constant that is of by a factor of 16. Furthermore for the constraint equation it holds $\bar{H}^2(1) - \bar{\rho}(1) = 248$.

Varying the parameters again, now the present day matter energy density is varied in three steps between $\pm 5\%$ of its original value and the initial value of the (dimensionless) Hubble constant is varied in three steps between $\pm 0,005\%$ of its original value, gives:

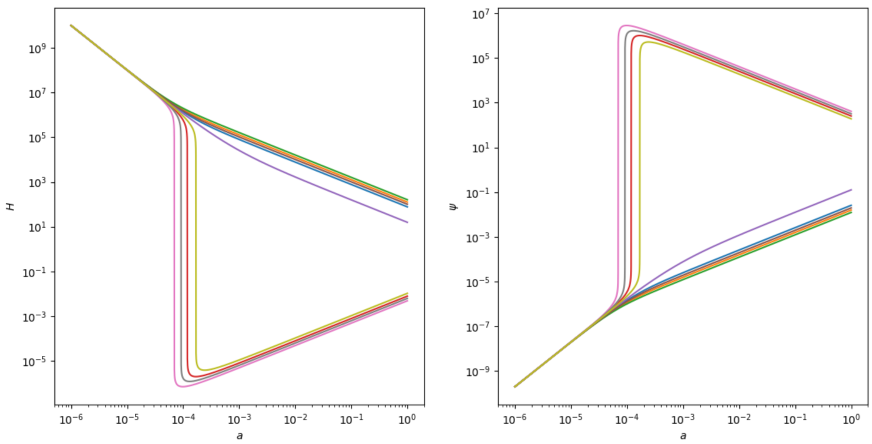

Figure 5.3.8: *The solutions, with on the left the dimensionless Hubble constant and on the right the dimensionless time derivative of the Galileon field, as a function of the scale factor with parameter values as above and with the present day matter energy density Ω_{m0} and the initial value of the Hubble constant $\bar{H}_i$varied, resulting in the nine graphs. The offshoots to below on the left and to the top on the right are present in the solutions in which Ω_{m0} is bigger than its original value, or $\bar{H}_i$ is smaller than its original value and Ω_{m0} is at its original value.*

and the corresponding EFT functions:

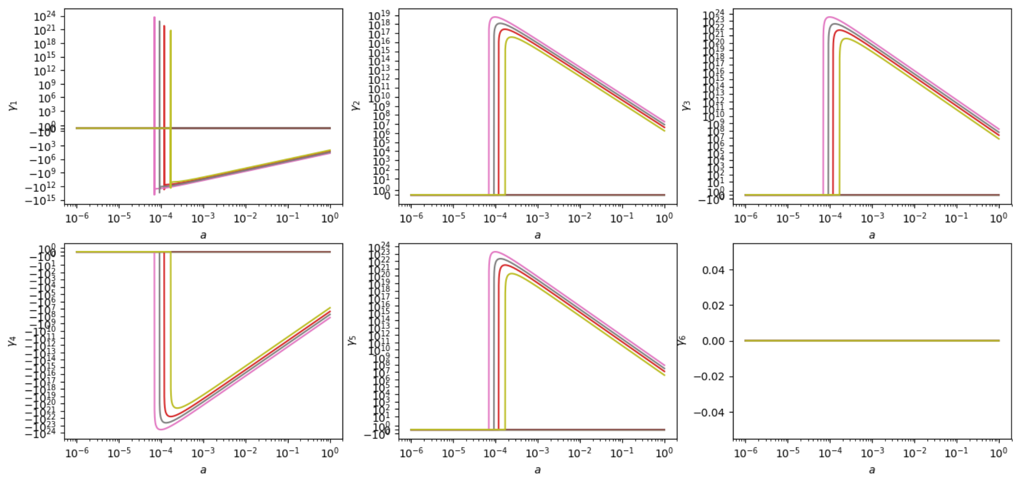

Figure 5.3.9: *The EFT functions corresponding to the solutions shown in the figure above.*

Chapter 6

Conclusion

Building on the differential geometric setting from General Relativity, the Covariant Galileon Model gives a description of the Universe. It posits in the same way as General Relativity that the Universe can be described by a four-dimensional manifold consisting of the three spatial dimensions and the one temporal dimension we experience. To this manifold a metric and a connection are associated and gravity is then described by the curvature resulting from those two. Moreover a dynamical scalar field, the Galileon field, is defined on the manifold, which interacts with gravity, extending the General Relativity model. Applying the Principle of Least Action to the Lagrangian of the model, the equations of motion can be obtained. After having done this, the model was analysed by numerically finding solutions to the equations of motion. This has been done for two sets of parameters.

Firstly, by looking at the equation of state of the Galileon field, something can be said about the nature of the introduced Galileon field. For normal non-relativistic matter the equation of state is $w = 0$ and for relativistic matter such as radiation it is $w = \frac{1}{3}$. The Galileon field is thus clearly a substance of a whole different kind with a negative equation of state.

When comparing the simulations of the two sets of parameters, there is no big qualitative difference between the results of the simulations for the different parameters both for the actual solutions to the model and the EFT functions of the model. However direct comparison of the found solutions with the tracker solution shows that there is a mostly increasing difference between the found solutions and the tracker solution. The tracker solution however should attract solutions. This suggests that the numerical integration of the equations of motion was not implemented correctly and thus behaved badly.

When considering the varying of the present day matter density and the Hubble constant at the start of the simulation, what immediately strikes the eye are the spikes arising in the simulations of the model for both sets of parameters. This suggests that the model is highly sensitive to parameter changes. On the other hand the spikes only occurred when the parameters were varied in a specific direction, they didn't occur in the other direction. This suggests that the high sensitivity only occurs for some parameter values and that there is a region in parameter space for which the solutions are relatively stable under the variation of the present day matter density and the Hubble constant at the start of the simulation. The tracker solution however should stabilize the behaviour of the solutions for all parameter values. This contradiction reinforces the idea that the integration of the equations of motion was not implemented correctly.

Furthermore the spikes arise around the $a = 10^{-4}$ mark in the simulation. It is at the moment unclear why this happens at this specific value of the scale factor. Maybe it relates to the fact that the equation of state of the Galileon field begins to change at that moment. Furthermore, when performing the variations of the present day matter density and the Hubble constant at the start of the simulation in this analysis, the other parameters were kept fixed. This however will cause the solutions to in general not fit the pre-imposed conditions that determine the other parameters. In particular this means that the assumed flatness of spacetime at the start and end of the simulation will not in general hold for the solutions with the varied parameters.

The current response of the solutions and EFT functions of the Covariant Galileon Model to the variations of the present day matter density and the Hubble constant at the start of the simulation is thus contradictory to the expectations obtained from the existence of the tracker solution. If the contradictions in the predictions of the model can be resolved however, then the viability of the model can be tested by comparing observations concerning the EFT functions with the predicted EFT functions and hence ruling out possible parameter values of the model. Hence to be able to test this model, more research is needed in the analysis of this model, especially concerning the implementation of the numerical integration of the model.

Bibliography

[1] S. Kobayashi and K. Nomizu, *Foundations of differential geometry. Vol I*, Interscience Publishers, a division of John Wiley & Sons, New York-Lond on, 1963.

[2] M. Lübke, *Introduction to manifolds*, Part of the course Differential Manifolds 2 (2016-2017) of Universiteit Leiden, 2014.

[3] L. W. Tu, *An Introduction to Manifolds*, Springer, New York, NY, 2008.

[4] S. M. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*, Pearson Education, 2013, Google-Books-ID: m75XngEACAAJ.

[5] R. M. Wald, *General Relativity*, University Of Chicago Press, first edition edition, 1984.

[6] A. Kriegl and P. W. Michor, *The Convenient Setting of Global Analysis*, American Mathematical Society, 1997, Google-Books-ID: s7fPYRqhXEUC.

[7] A. Barreira, B. Li, C. Baugh, and S. Pascoli, *Linear perturbations in Galileon gravity models*, Physical Review D **86** (2012), arXiv: 1208.0600.

[8] E. Gourgoulhon, *3+1 Formalism and Bases of Numerical Relativity*, arXiv:gr-qc/0703035 (2007), arXiv: gr-qc/0703035.

[9] N. Frusciante, G. Papadomanolakis, and A. Silvestri, *An Extended action for the effective field theory of dark energy: a stability analysis and a complete guide to the mapping at the basis of EFTCAMB*, Journal of Cosmology and Astroparticle Physics **2016**, 018 (2016), arXiv: 1601.04064.

[10] A. Barreira, B. Li, C. Baugh, and S. Pascoli, *The observational status of Galileon gravity after Planck*, Journal of Cosmology and Astroparticle Physics **2014**, 059 (2014), arXiv: 1406.0485.