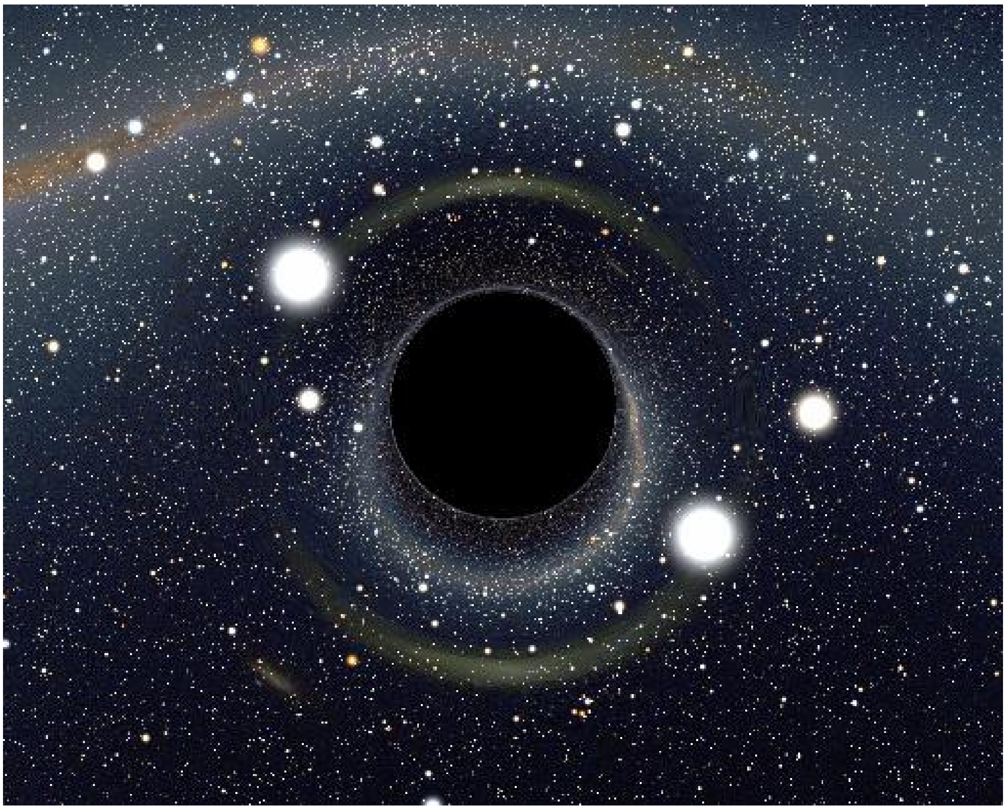
CHALMERS
UNIVERSITY OF TECHNOLOGY



Black Holes and Hidden Symmetries

Solution Generating Techniques with Dimensional Reduction and Group Theory

Bachelor of Science Thesis for the Engineering Physics Program

Marcus Aronsson, Axel Dahlberg, Linnea Hesslow,
Elin Romare, Erik Roos and Axel Widmark

Department of Fundamental Physics
Chalmers University of Technology
Gothenburg, Sweden 2014

Black Holes and Hidden Symmetries

Authors: Marcus Aronsson[a], Axel Dahlberg[b], Linnea Hesslow[c], Elin Romare[d], Erik Roos[e] and Axel Widmark[f]

Contact:

[a]mararons@student.chalmers.se
[b]axeld@student.chalmers.se
[c]hesslow@student.chalmers.se
[d]elinrom@student.chalmers.se
[e]erikro@student.chalmers.se
[f]waxel@student.chalmers.se

FUFX02 - Bachelor Thesis at Fundamental Physics
Bachelor Thesis N° FUFX02-14-06
Supervisor: Daniel Persson, Division of Mathematical Physics
Examiner: Christian Forssén, Division of Subatomic Physics

Department of Fundamental Physics
Chalmers University of Technology
SE-412 96 Gothemburg
Sweden
+ 46 (0) 31 772 1000

Printed by Chalmers reproservice
Göteborg, Sweden 2014

Cover: The figure on the front page is a computer generated image of the visual distortions around a black hole. Image: NASA.

REPORT NO. FUFX02-14-06

Black Holes and Hidden Symmetries

Solution Generating Techniques with Dimensional Reduction and Group Theory

MARCUS ARONSSON AXEL DAHLBERG LINNEA HESSLOW
ELIN ROMARE ERIK ROOS AXEL WIDMARK

Department of Fundamental Physics
Chalmers University of Technology
Gothenburg, Sweden 2014

Abstract

In this bachelor thesis we have studied a technique of generating solutions for the curvature of black holes. In order to understand the underlying theory of this method we have studied the fundamentals of general relativity and group theory. The technique utilizes dimensional reduction to expose hidden symmetries of black holes, which enables us to find new solutions. This type of solution-generating technique is currently subject to heavy research, with hope of exposing deeper symmetries of spacetime and black holes.

Sammandrag

I det här kandidatarbetet har vi studerat en metod att generera lösningar för rumtidens krökning kring ett svart hål. För att förstå den bakomliggande teorin har vi studerat allmän relativitet samt gruppteori. Metoden nyttjar dimensionell reduktion för att upptäcka gömda symmetrier hos svarta hål, vilket gör det möjligt att hitta nya lösningar. Den här sortens lösningsgenererande teknik är just nu ett stort forskningsområde, med vilken man hoppas kunna avkoda djupare symmetrier hos rumtid och svarta hål.

Acknowledgements

We would like to thank our mentor Daniel Persson, without whose knowledge and support we would never have understood anything. Additionally, we are very grateful to Henrik Gustafsson, Axel Radnäs, Erik Widén and Christian von Schultz for giving us further understanding of many difficult concepts.

The Authors, Gothenburg, June 9, 2014

Contents

Part I

A Swedish Summary of the Project

Bakgrund

När Albert Einstein i början av 1900-talet utvecklade sin relativitetsteori revolutionerade den fysiken. Den gav ett nytt sätt att beskriva världen vi lever i och öppnade upp ett enormt forskningsområde. Det mest revolutionerande med den allmänna relativitetsteorin var att den till skillnad från Newtons teori inte beskrev gravitation som en kraft, utan som en krökning av rumtiden. Tidigare hade man dessutom betraktat tiden som absolut, att tiden gick lika fort i alla delar av universum, men Einsteins teorier visade att så inte är fallet. Hur tid upplevs är helt enkelt relativt.

En astrofysiker vid namn Karl Schwarzschild satte sig in i Einsteins teorier och lyckades härleda en lösning till rörelseekvationerna för objekt i närheten av ett massivt sfäriskt objekt i vakuum, till exempel hur jorden roterar runt solen. Schwarzschildlösningen har egenskaperna att den är oladdad, sfäriskt symmetrisk och statisk, det vill säga tidsoberoende. Lösningen visade sig även ha en del märkliga egenskaper i vissa speciella fall. Om det massiva objektet har tillräckligt hög densitet övergår det till vad vi idag kallar för ett svart hål. Den höga densiteten innebär att all massa ligger innanför radien för händelsehorisonten, vilket resulterar i ett område där något som kommit in aldrig kan återvända ut.

Ett kraftfullt verktyg för att beskriva symmetrier är gruppteori. I gruppteori betraktar man grupper av element, som inom gruppen delar vissa egenskaper. Om detta sammanförs med en alternativ lösning till Schwarzschild kan detta resultera i en ny lösningsgång som kan utnyttjas till att bland annat härleda andra, liknande lösningar. För att härleda den alternativa lösningsgången till Schwarzschild utförs, i denna rapport, först en reduktion av den fyrdimensionella rumtiden till tre dimensioner. Genom att använda sig utav gruppteori på den tredimensionella verkan kan man upptäcka vissa tidigare dolda symmetrier. Denna kompaktifiering utvecklades av den tyske matematikern Theodore Kaluza och den svenske fysikern Oscar Klein. I stora drag går deras kompaktifiering ut på att man betraktar en dimension som obereonde för lösningen och på så sätt kan omformulera problemet till att ha en lägre dimension. Motsatsen, då man går till en högre dimension, kallas dekompaktifiering.

En stor anledning till varför det är intressant med svarta hål är bland annat att Einsteins teori bryter ihop i singulariteten i det svarta hålets mitt. Genom att studera detta beteende kan man förhoppningsvis få större förståelse för de extrema omständigheter som omger ett svart hål. Genom att kunna härleda lösningen, inte bara från ett fysikaliskt perspektiv som Schwarzschild, utan även ur ett matematiskt perspektiv med gruppteori, erhåller man ytterligare en infallsvinkel för en förståelse av svarta hål.

Syfte

Syftet med det här projektet är att härleda en alternativ lösningsgång för ett svart hål, samt att finna vidare användning av en sådan metod och hitta andra lösningar. Projektet delas därmed i två delmål:

- Bekanta oss med allmän relativitetsteori och gruppteori, för att få en basal förståelse inför nästa steg.

- Härleda Schwarzschildlösningen för ett svart hål genom att först utföra en dimensionell reduktion på den fyrdimensionella verkan i vakuum, och därefter beräkna de resulterande rörelseekvationerna med hjälp av bland annat gruppteori.

Detta innebär att man först studerar Schwarzschildlösningen ur ett fysikaliskt perspektiv, för att sedan ur ett matematiskt perspektiv härleda samma lösning genom en dimensionsreduktion. Resultatet, som redovisas i en rapport, är tänkt att kunna hjälpa studenter och intresserade som vill lära sig om allmän relativitetsteori och gruppteori, men framför allt hur man kan förena dem. Vi vill då visa hur dimensionell reduktion kan användas för att skapa familjer av svarta hål, härledda från en ursprungslösning.

Problem

Vanligen när man tar fram Schwarzschildlösningen så löser man Einsteins ekvationer utifrån en ansats. Genom att istället använda gruppteori och utnyttja lösningens symmetrier hittar man en alternativ lösningsgång, som resulterar i nya sätt att hitta mer allmänna lösningar. Fokus har därför legat mer på lösningsgången snarare än Schwarzschildlösningen i sig. Målet var att producera explicita beräkningar som, med utgångspunkt i verkansteori, når en lösning till Einsteins rörelseekvationer för ett svart hål. Vi avsåg även att erhålla tillräckligt god förståelse för att kunna förklara detta för någon med förkunskaper liknande de vi hade då projektet började.

För att uppnå detta var det nödvändigt att sätta sig in i grunderna för allmän relativitetsteori och gruppteori genom att börja projektet med litteraturstudier. Inom allmän relativitetsteori behövds förståelse för grundläggande begrepp, så som tensorer och krökta rum, samt mer avancerade begrepp som Einsteins ekvationer och hur dessa ger Schwarzschildlösningen. För gruppteorin behövdes förståelse för grundläggande begrepp såsom Liegrupper och Liealgebra.

Avgränsningar

Allmän relativitetsteori och gruppteori är två enorma ämnes- och forskningsområden, vilket gjorde att man med den knappa tiden var tvungen att hålla sig till vissa avgränsningar.

Vad gäller allmän relativitet låg intresset i specialfallet för rumtidens krökning i vakuum, vilket är en kraftig avgränsning och förenkling från det generella fallet. Då Schwarzschildmetriken var vad som skulle genereras krävdes det en kunskap om just detta: den sfäriskt symmetriska rumtidskrökningen kring en singularitet. Gruppteori är en väldigt omfattande gren inom matematiken, vilket gjorde att inriktningen var tvungen vara tydlig mot den teori som var relevant för projektet. Utöver en mycket grundläggande diskussion om gruppteori i allmänhet låg fokus på Liegrupper och Liealgebror. Inom Liegrupper begränsade vi oss även till de enklare symmetrigrupperna SL(2,$\mathbb{R}$), SO(2) och SO(1,1), då dessa hade störst koppling till vårt problem, samt att vi beskrev ett fåtal relevanta koncept för rapporten.

Metod

Alla medlemmar i kandidatgruppen hade innan arbetets början läst vår handledare Daniel Perssons kurs i speciell relativitetsteori, vilket innebar att alla hade en grundläggande förståelse för delar av rapportens ämne. För att utvidga vår kunskap tog vi del av litteratur på ämnet allmän relativitetsteori, vilka fanns tillgängliga för gruppen på vår projekthemsida.[1] Som komplettering till detta höll Daniel Persson ett flertal mindre föreläsningar, med fokus på de delar som var viktiga för just vår rapport. Daniel Persson höll även i kursen “Gravitation and Cosmology”, där ett antal föreläsningar var av relevans för vårt arbete, och som vi därför tog del av. Hälften av gruppen valde dessutom att läsa hela kursen utöver ordinarie poäng. För att befästa all inlärning fördes även diskussioner och höll föreläsningar inför varandra på valda ämnesområden, samtidigt som ett utkast till rapporten påbörjades. Parallellt med detta räknade vi på olika begrepp inom teorin, för att få bättre förståelse och för att på detta sätt komma fram till Schwarzschildlösningen på ett traditionellt vis. Räkningarna samlades i en TEX-fil på www.sharelatex.com så att alla kunde jämföra resultat och komma med kommentarer.

Vidare handlar arbetet mycket om användning av gruppteori inom fysiken. Därför tog vi oss an grunderna av gruppteori på liknande sätt som vi gjort med allmän relativitetsteori. Speciellt när det gällde gruppteori var det viktigt att vi avgränsade oss till just den information som skulle visa sig vara viktig för att utföra dimensionell reduktion på vårt problem.

[1] http://www.danper.se/Daniels_homepage/GroupTheory_2.html

Då större insikt i ämnet hade uppnåtts kunde Daniel Perssons exjobbare Axel Radnäs och Erik Widén hålla föreläsningar för oss om vad de höll på med, för att på så sätt se problemet från andra synvinklar. Detta blev en del av introduktionen till hur gruppteori och allmän relativitetsteori kan förenas. Då det finns lite litteratur som explicit går igenom de beräkningar fram till vår lösning låg det stor fokus på att vi själva utförde dessa beräkningar. Slutligen presenterades litteraturstudierna och beräkningarna i en sammanställd rapport.

Resultat

Målet att generera en lösningsmetod för Schwarzschildlösningen nåddes under arbetets gång. Detta användes sedan för att diskutera vidare möjliga lösningar och även specifikt för att lösa Reissner-Nordströms problem med ett laddad, statiskt och sfäriskt symmetriskt svart hål. Arbetet har resulterat i en rapport som är tänkt att kunna hjälpa studenter att förstå relativitetsteori, gruppteori och hur man för dem samman. Rapporten är därför uppdelad i tre huvudkapitel:

1. introduktion till allmän relativitetsteori,
2. introduktion till gruppteori,
3. en härledning av Schwarzschildlösningen ur ett gruppteoretiskt perspektiv, samt vidare applikation på bland annat Reissner-Nordströmlösningen.

Uppdelningen speglar inlärningsprocessen och upplevdes därför som passande för rapporten.

I kapitel 4 sammanfogas de två teorierna, vilket resulterade i en metod för att finna Schwarzschildlösningen utifrån dimensionell reduktion. Kärnan i detta var att metoden som använts sedan kunde modifieras på olika sätt för att hitta andra lösningar till svarta hål. Resultatet av en sådan modifikation kan ses i avsnitt 4.4.

Slutsats

Vår förhoppning är att rapporten kommer att kunna ge ett stöd för de som är intresserade av lösningsmodeller för olika typer av svarta hål. Utifrån den metod som presenterats i rapporten kan man finna vidare användningar för andra problem. Metodens bredd och allmänna tillämpningsmöjligheter är vad som gör den så kraftfull. Metoden presenterar både en inspiration till lösningsvägar för liknande rörelseekvationer med liknande variation, samt ett sätt att från Schwarzschild direkt generera andra lösningar i Schwarzschildfamiljen.

Part II

Report

1

Introduction

Albert Einstein's theory of relativity revolutionized the world of physics and created a new paradigm in terms of how we describe space and time. The theory of special relativity that Einstein published in 1905, building on the groundwork of many other physicists, unified space and time in the special case of no gravity. Due to the complicated nature of generalizing this special case to also describe gravity it would take another eleven years until he published his geometric theory of gravity, known as the theory of general relativity. The theory is geometric in the sense that gravity is not a force, rather it is an effect of a spacetime manifold curved by mass.

The theory of relativity has been proven empirically sound. Predictions such as the bending of light by gravity, time dilation for particles moving close to the speed of light, and the red shift from stars drifting away from us, have all been observed to correlate with the theory.

The most central equations in relativity are Einstein's equations. Einstein's equations describe motion in a curved spacetime. Using tensor formalism they can be written as

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}. \tag{1.0.1}$$

In four-dimensional spacetime this amounts to ten equations[1].

In 1931 an astrophysicist called Karl Schwarzschild derived a solution to Einstein's equations now known as the Schwarzschild solution. The solution describes the geometry of the space around a massive spherical object in vacuum. The Schwarzschild solution can be used to describe how the earth orbits the sun or how a satellite orbits the earth and is actually used to compensate for relativistic effects in GPS-satellites to make them precise in pinpointing a location.

[1] The equation and tensor formalism will be explained in chapter 2.

The Schwarzschild solution also describes the geometry of a more exotic massive object, that of a black hole. For black holes the Schwarzschild solution yields some interesting differences from other objects due to its dense mass. From the Schwarzschild solution we find the Schwarzschild radius at $r = 2GM$, and for most massive objects this radius falls within the objects perimeter. For a black hole however, all the mass lies within the Schwarzschild radius, which gives rise to the event horizon along the sphere of the Schwarzschild radius. The event horizon is what gives the blackness of a black hole, as essentially nothing can escape from within this horizon, not even light.

The Schwarzschild solution involves a lot of symmetries as it is both static and rotationally invariant. An important and powerful tool for describing symmetries in a compact and general way is group theory. A group is a set of transformations that obey certain axioms (see chapter 3). If we have something that is rotationally invariant and act on it with a transformation that is also invariant, the result is also rotationally invariant. This property is very useful when working with symmetries and is one of the strengths of group theory. When working with groups it is possible to derive very general and applicable solutions to problems involving some kind of symmetry. In chapter 3 some basics of group theory will be presented, with a focus on Lie groups and their applications.

Using group theory, one can derive the Schwarzschild solution in an alternative way to the more hands on physics approach. This approach is of interest as the solution is more general and the symmetries and characteristics of the solution are more apparent. This way of finding the Schwarzschild solution also presents a process that can be slightly altered to generate solutions to other types of black holes. For example the solution for a charged black hole, the Reissner-Nordström solution, can be found using the same techniques as for the Schwarzschild solution. In this text we will derive both the Schwarzschild and the Reissner-Nordström solutions. However, as the processes are analogous, the focus lies primarily on the simpler Schwarzschild solution.

An action is a functional and describes the characteristics of a system. The action principle says that by varying the action it is possible to derive the equations of motion for the system, as the action is assumed to be stationary. In this text we find the Schwarzschild solution by examining the action of gravity in four dimensions.

The black holes considered herein are static and spherically symmetric. This means that there is an amount of redundancy when treating it in four dimensions, as for instance nothing changes in time. The same hold true for the angular coordinates and one can reduce the dimension of the problem by mathematically removing one dimension. The dimensional reduction in this paper was developed by the German mathematician Theodore Kaluza and the Swedish physicist Oskar Klein. Reducing one dimension is not as easy as it may sound, and we do not discuss in detail how this works. The action in four dimensions is replaced by a corresponding action in three dimensions. Solving this problem and then performing a decompactification gives the solution in four dimensions.

The solution generating techniques described in this thesis are useful when developing a theory of quantum gravity, the combination of quantum mechanics and general relativity. To construct quantum gravity, it is necesary to understand the solutions predicted by gravity. Black holes are among the most interesting solutions as Einstein's theories break down into a singularity. To describe black holes in a completely satisfactory way, quantum gravity is needed. Black hole solutions are therefore of great interest, and dimensional reduction is a powerful tool when obtaining these solutions since hidden symmetries in four dimensions may revealed in three or two dimensions. Using these symmetries, it is possible to classify black holes and derive entire families of black holes from one solution.

Dimensional reduction is useful, not only to derive solutions of black holes, but also in constructing the theory of quantum gravity itself. These theories, such as supersymmetry and string theory, describe a world of ten or eleven dimensions and dimensional reduction is therefore necessary to describe our four-dimensional world.

In dimensional reduction, more terms are added to the action as the number of dimensions is reduced. Physical theories that seem different in four dimensions can be unified in ten or eleven dimensions. On the other hand, the different action obtained when four-dimensional Einstein gravity is reduced to three or two dimensions reveal hidden symmetries and can be analyzed in the framework of group theory to obtain the four-dimensional solution.

The purpose of this thesis was to show how it is possible to derive the Schwarzschild solution with the hidden symmetries of a black hole revealed with the dimensional reduction from four to three dimensions. To do this we first present a short introduction to general relativity and group theory and then combine the two to arrive at the Schwarzschild solution. We succeed in reaching our goal by looking at the action of the given system, then performing a Kaluza-Klein compactification on the four-dimensional spacetime to solve the problem in three dimensions. By performing a decompactification we then obtain the solution in four dimensions. We also reach beyond Schwarzschild and derive the solution of a charged black hole, the Reissner-Nordström solution.

To conclude, in this thesis we first study the Schwarzschild solution with Einstein's theory. Then we perform a dimensional reduction to three dimensions to derive the same solution using group theory. After that we go beyond the Schwarzschild solution and look at the Reissner-Nordström solution as well as other solutions in the Schwarzschild family.

1.1 Reading guide

The report is divided into three major parts. Chapter 2 and 3 are meant to work as an introduction to general relativity and group theory respectively. Some of the subjects to be discussed in chapter 2 are metrics, proper time, transformations, tensors, Einstein's equations and the Schwarzschild solution while chapter 3 treats, among other things, the definition of a group, Lie groups and Lie Algebras as well as some important concepts like the Iwasawa decomposition and the Killing form.

As the two first parts are supposed to serve as introductions the reader can, if already acquainted with general relativity and group theory, go directly to chapter 4. Chapter 4 discusses the dimensional reduction and uses what is discussed in earlier chapters to derive the Schwarzschild solution using group theory.

2

General Relativity

The intention of this chapter is to present an introduction of general relativity that is terse but adequate in reaching our prime objective of describing the spacetime singularity of a black hole. To achieve this we need to go over the basics of special relativity, discuss some technicalities of general relativity (such as metric, tensors, affine connection, covariant derivative, etc), derive Einstein's equations and finally find the Schwarzschild solution of a black hole, as well as the Reissner-Nordström solution.

The main sources of reference in this chapter are Carrol's "Lecture Notes on General Relativity" [1] and Weinberg's book "Gravitation and Cosmology" [2], which can be studied for further information on the subject.

2.1 Special Relativity and Minkowski Space

Before trying to describe how gravity works in Einstein's theory of general relativity we need talk about special relativity, which is the simple case without gravity, what one would call flat spacetime.

Newton's theory of motion, published in 1687, is in itself a theory of relativity, in the sense that motion is relative and what one considers to be at rest is arbitrarily chosen. In more technical terms, Newton's laws are invariant under so called Galilean transformations. Given two inertial systems that have a velocity v with respect to each other, we can relate the coordinates of one to the other according to the transformation

$$\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z \\ t' &= t. \end{aligned} \tag{2.1.1}$$

This is what we are used to in every day life. A passenger on a train does not feel that he is moving, if he closes his eyes he has the experience of standing still. A bystander on the other hand, that is standing still in his own inertial system, would say that the train is moving. With Galilean transformations one can have relative velocities in between inertial systems in any combination of the x-y-z-directions. One cannot, however, have different perceptions of time in different inertial systems. In Newton's theory of motion time is absolute. This might seem a superfluous statement, because how could it be otherwise? Well, we shall see.

Newtonian mechanics are invariant under these transformations. However, Maxwell's equations are not. The magnetic force depends on velocity, so that a particle that experiences a magnetic force in one inertial system does not experience any magnetic force at all in its own its rest frame, the inertial system in which it is standing still. This led Einstein to believe that Newton's theory of motion with its Galilean transformations could not be the whole truth.

Einstein based his theory of special relativity on two postulates:

1. the speed of light is the same in all inertial systems, and
2. the outcome of any experiment is independent of the inertial frame.

Given these postulates, Einstein started to explore what happened when things moves at a speed close to the speed of light. He found that we can no longer talk about a three-dimensional world with time as a separate parameter, but a four-dimensional spacetime where time and space intertwine. Galilean transformations were replaced by Lorentz transformations, formulated by Hendrik Antoon Lorentz.[1] A Lorentz transformation from one inertial system to another with a relative velocity v along their x-axes, known as a standard configuration, looks like

$$\begin{aligned} x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma(t - \frac{vx}{c^2}) \end{aligned} \tag{2.1.2}$$

[1] Lorentz, understandably, did not take his own equations literally enough. He merely stated that Maxwell's equation were invariant under these coordinate transformations without drawing the conclusion that space and time are intertwined.[3]

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{2.1.3}$$

called the **gamma factor**. The standard configuration transformation is a special case; Lorentz transformations include transformations with respect to velocities v in any spatial direction, spatial rotations, and any combination thereof.

We see that when moving close to the speed of light the transformations of x and t start to interfere, creating relativistic effects like time dilation and length contraction. Note that the idea of absolute time no longer holds, because the passage of time is dependent on the gamma factor.

A classic example of this is the **twin paradox**, which goes as follows. A pair of twins of the same age are separated. One stays on earth, while the other gets into a spaceship and travels to a nearby star and back again close to the speed of light. While the twin on earth ages many years, the twin in the spaceship does not experience a very long passage of time and hardly ages at all. In fact the time experienced by the traveling twin can be made arbitrarily short in our thought experiment if we do not limit the force by which he is accelerated.

None of this means that Newton's laws of motion are obsolete, they still hold for objects moving slowly relative to the speed of light. Comparing Galilean and Lorentz transformation, equations (2.1.1) and (2.1.2), they are asymptotically equal in the limit of low velocities, when

$$v \to 0 \Rightarrow \gamma(v) \to 1. \tag{2.1.4}$$

The geometry of special relativity, which is flat four-dimensional spacetime, is known as **Minkowski space**. The Minkowski space is flat for the same reason special relativity is special: we have no gravitation. Later, the introduction of gravity will produce the curved spacetime of general relativity. Before we get there we need to discuss a couple of other concepts in greater detail.

2.1.1 Metric

A **metric** is basically an abstract way to define distance within a geometry. For example, in Euclidean space, the infinitesimal distance ds between two points would be given by

$$ds^2 = dx^2 + dy^2 + dz^2. \tag{2.1.5}$$

A translation or rotation of the coordinate system (what would be a Galilean transformation) result in the same distance between the two points, given by the new coordinates according to

$$ds^2 = dx'^2 + dy'^2 + dz'^2 \tag{2.1.6}$$

Note that the way of calculating the distance are precisely the same in the two cases.

The metric of the Minkowski space is

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{2.1.7}$$

or

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2. \tag{2.1.8}$$

They describe the same Minkowski space but they are not the same. Choosing one or the other is merely a matter of preference. In this thesis we mostly use the one defined in equation (2.1.7), known as the $(-+++)$-convention. (We also make use of natural units where the speed of light is $c = 1$. Not using natural units would mean a factor c^2 in front of the first term dt^2.)

The $(-+++)$-metric can also be written as a matrix,

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.1.9}$$

Given a vector

$$l = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}, \tag{2.1.10}$$

with its length s_l given by the expression

$$s_l^2 = l^\top \eta l = -t^2 + x^2 + y^2 + z^2. \tag{2.1.11}$$

The distance between two points correspond to the **proper time**, which is the time experienced by a particle that travelled this distance. In the $(-+++)$-metric the proper time $\Delta\tau$ experienced along a straight line in Minkowski space is calculated as

$$\Delta\tau = \sqrt{-l^\top \eta l}. \tag{2.1.12}$$

Considering the proper time experienced by something traveling at the speed of light (remember we are using natural units where $c = 1$), we see that $\Delta\tau = 0$. It would seem that in the inertial system of something traveling at the speed of light, it arrives instantly and the distance travelled was zero. Considering even faster

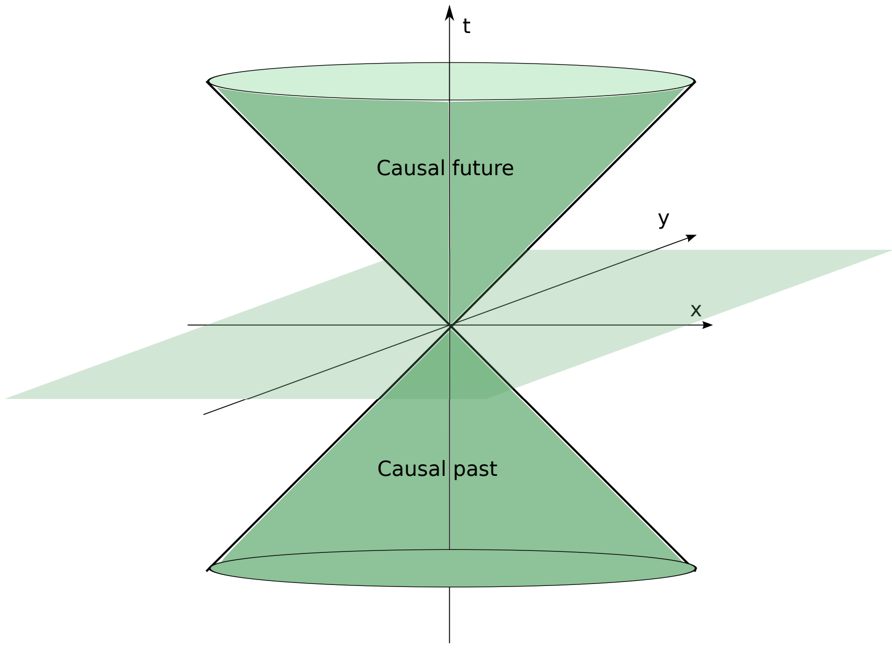


Figure 2.1: The light-cone stretching out into causal past and future from a point in spacetime.

speeds leads to this important point: something traveling faster than the speed of light, could in another inertial system (after Lorentz transformation) be said to travel the other way, backwards in time. In that case causality would be broken, because observers would disagree which one of the connected events happened first. In this sense the speed of light is a cosmic speed limit, the maximum speed at which information can travel. The speed of light defines the edges of a **light cone**, whose inside encapsulates events causally connected to that point (see figure 2.1).

2.2 Tensors

When talking about special and general relativity there is a convention of using tensor notation to describe the relations between spacetime coordinates, such as when calculating the proper time of a curve through spacetime. The tensor formalism enables us to describe mathematical relations in a compact and clear way that will prove indispensable for calculations in general relativity. Tensors can be scalars or vectors, but they can also represent matrices or objects of an arbitrary number of dimensions. For example, a coordinate vector is written x^μ, where μ

represent the four coordinates of spacetime so that

$$x^\mu = (t,x,y,z). \tag{2.2.1}$$

A **tensor** is defined as an object that transforms in a certain way during a general transformation from one coordinate system to another, $x^\alpha \to x'^\alpha$. For a tensor with an upper index V^α and a tensor with a lower V_α the coordinate change to V'^α and respectively V'_α is written

$$V'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta} V^\beta \tag{2.2.2}$$

$$V'_\alpha = \frac{\partial x^\beta}{\partial x'^\alpha} V_\beta \tag{2.2.3}$$

or more generally

$$T'^{\alpha_1 \cdots \alpha_k}{}_{\beta_1 \cdots \beta_l} = \frac{\partial x'^{\alpha_1}}{\partial x^{\lambda_1}} \cdots \frac{\partial x'^{\alpha_k}}{\partial x^{\lambda_k}} \frac{\partial x^{\zeta_1}}{\partial x'^{\beta_1}} \cdots \frac{\partial x^{\zeta_l}}{\partial x'^{\beta_l}} T^{\lambda_1 \cdots \lambda_k}{}_{\zeta_1 \cdots \zeta_l}\,. \tag{2.2.4}$$

In special relativity this coordinate transformation is always a Lorentz transformation.

The distinction between upper and lower indices, also known as **contravariant** and **covariant** indices, becomes important when we consider summation. When two indices are the same they are summed over according to

$$T^\alpha S_\alpha = \sum_\alpha T^\alpha S_\alpha, \tag{2.2.5}$$

just like in a normal scalar product of vectors. Summation over two upper indices, or two lower, is not allowed, so an expression such as $T_\alpha S_\alpha$ or $T^\alpha S^\alpha$ would be nonsense. An index that is not summed over is a **free index**. Respectively, an index that is summed over is a **contracted index** or a **dummy index**. The name of a dummy index can be arbitrarily changed, as it is "invisible" after the summation.

The contravariant and covariant forms are related through the metric. The Minkowski metric $\eta_{\alpha\beta}$ has components

$$\eta_{\alpha\beta} = \begin{cases} 0 & \alpha \neq \beta \\ -1 & \alpha = \beta = 0 \\ 1 & \alpha = \beta = 1,2,3. \end{cases} \tag{2.2.6}$$

A lower index is produced from an upper index according to

$$\eta_{\alpha\beta} T^\beta = T_\alpha, \tag{2.2.7}$$

and vice versa

$$\eta^{\alpha\beta}T_\beta = T^\alpha. \tag{2.2.8}$$

Because lowering an index and then raising it should leave a tensor unchanged, $\eta^{\alpha\beta}$ must be the inverse metric and fulfill that

$$\eta^{\alpha\gamma}\eta_{\gamma\beta} = \delta^\alpha_\beta, \tag{2.2.9}$$

δ^α_β is known as the Kronecker delta and is the tensor equivalent of the identity matrix. It is defined as

$$\delta^\alpha_\beta = \begin{cases} 0 & \alpha \neq \beta \\ 1 & \alpha = \beta. \end{cases} \tag{2.2.10}$$

Thus, when raising and then lowering the same index, the effect is only to change the name of that index.

The way we calculated the length of a four-vector in Minkowski spacetime, see equation (2.1.11), can now be written more concisely. The squared length of a vector l^β, where $\beta = t,x,y,z$, can now be expressed as

$$\eta_{\alpha\beta}l^\alpha l^\beta = l^\beta l_\beta, \tag{2.2.11}$$

which is a scalar because it has no free indices. A **scalar** is a tensor with no indices and is always invariant with respect to coordinate transformations.

Especially when a tensor has several indices this formalism is very concise. As an example we might have $T^{\alpha\beta}{}_{\lambda\zeta\epsilon}$ and $S_{\beta\xi}$ that could be contracted into another tensor $U^\alpha{}_{\lambda\zeta\epsilon\eta}$ according to

$$T^{\alpha\beta}{}_{\lambda\zeta\epsilon}S_{\beta\eta} = \sum_\beta T^{\alpha\beta}{}_{\lambda\zeta\epsilon}S_{\beta\eta} = U^\alpha{}_{\lambda\zeta\epsilon\eta}. \tag{2.2.12}$$

Note that a contraction causes two of the same indices to cancel each other out, leaving a tensor with two less indices. The number of free contravariant indices are added up in tensor multiplication, as does the number of free covariant indices. A sum of tensors is itself a tensor, if the terms added have the same upper and the same lower indices.

When referring to the symmetric or antisymmetric part of a tensor, there is the convention of writing $T_{(ij)}$ and $T_{[ij]}$ respectively. For example, if the tensor is strictly antisymmetric, we can write this as $T_{(ij)} = 0$.

2.3 Important Objects of General Relativity

The next step in approaching the description of a black hole, or the Schwarzschild solution in particular, is the introduction of a series of objects specific to the theory of general relativity. To approach the formulation of **Einstein's equations**, we will further expand the concept of the **metric**, to then move on to describe **affine connection**, **covariant derivative**, **parallel transport**, **Riemann tensor**, **Ricci tensor** and **Ricci scalar**.

There is also the concept of a **manifold**, which is briefly explained in appendix C. What is necessary to understand is simply this: a manifold is a geometry that is locally flat. Spacetime is a manifold and in every point it has a tangent space (a choice of coordinate system) that locally reduces to the Minkowski metric.

In special relativity the symmetries are independent of the coordinates. This is not the case for general relativity where these symmetries are only local, because any point in spacetime looks locally flat. What used to be a simple Lorentz transformation in flat Minkowski spacetime, is in general relativity some function of our general coordinates $\frac{\partial x'^{\mu}}{\partial x^{\nu}}$.

2.3.1 Metric

In special relativity the metric is constant, i.e. independent of coordinates. This is not the case of the curved spacetime of general relativity. To emphasize this difference the metric of general relativity is given a new symbol: $g_{\mu\nu}$. Analogous to the Minkowski metric, the infinitesimal distance is related to the metric and the generalized coordinates x^{μ} by

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu}. \tag{2.3.1}$$

The inverse metric $g^{\mu\nu}$, as described in section 2.2, is given by the relation

$$g^{\mu\nu}g_{\nu\sigma} = \delta^{\mu}_{\sigma}. \tag{2.3.2}$$

From this it also follows that the metric and its inverse, $g_{\mu\nu}$ and $g^{\mu\nu}$, are symmetric ($g_{\mu\nu} = g_{\nu\mu}$ or equivalently $g_{[\mu\nu]} = 0$).

The components of the metric vary with the coordinates but it is always possible to choose a coordinate system in which the metric locally takes the form of the Minkowski metric $\eta_{\alpha\beta}$ to a first order approximation. This is the meaning of spacetime being locally flat.

Example 2.1. The simplest possible example of a metric on a curved manifold would be the metric of a unit 2-sphere (meaning the two-dimensional surface of a sphere with radius $r = 1$). Unlike spacetime that has four coordinates, the 2-sphere only has two. From basic calculus we know that the squared infinitesimal length between two points on the unit 2-sphere is given by

$$\mathrm{d}s^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2. \tag{2.3.3}$$

Because the infinitesimal length is also given by $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$, we see that the components of the metric on a 2-sphere should be

$$\begin{aligned} g_{\theta\theta} &= 1 & g_{\theta\varphi} &= 0 \\ g_{\varphi\varphi} &= \sin^2\theta & g_{\varphi\theta} &= 0. \end{aligned} \tag{2.3.4}$$

Here we see that the metric is not a constant, rather it is a function of one of the coordinates, namely θ. The inverse metric, given by the relation $g^{\mu\nu}g_{\nu\rho} = \delta^\mu_\rho$, is simply the diagonal components inverted, according to

$$\begin{aligned} g^{\theta\theta} &= 1 & g^{\theta\varphi} &= 0 \\ g^{\varphi\varphi} &= \sin^{-2}\theta & g^{\varphi\theta} &= 0. \end{aligned} \tag{2.3.5}$$

The 2-sphere manifold has two coordinates, which resulted in the metric having four different components. Spacetime, on the other hand, has four coordinates and given the two indices of the metric we have $4 \cdot 4 = 16$ different components (although because of its symmetry there are actually ten different independent components).

We come back to this example as we discuss other mathematical objects related to a curved manifold.

2.3.2 Covariant Derivative and the Affine Connection

As has been mentioned before, not everything that looks like a tensor is a tensor. The partial derivative, for example, is not a tensor, because when subject to a coordinate transformation $x^\mu \to x'^\mu$ it becomes

$$\partial'_\nu V'^\mu = \frac{\partial V'^\mu}{\partial x'^\nu} = \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial V^\rho}{\partial x^\rho} + \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial^2 x'^\mu}{\partial x^\rho \partial x^\tau}V^\tau. \tag{2.3.6}$$

A detailed derivation of this result can be found in appendix A. This does not fulfill the definition of a correct tensor transformation, see equation (2.2.4), because of the additional term on the right hand side.

As a consequence, we define a derivative that is a tensor, the **covariant derivative**,

$$\nabla_\mu V^\nu \equiv \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda}V^\lambda, \tag{2.3.7}$$

where $\Gamma^\nu_{\mu\lambda}$ is the **affine connection**, defined as

$$\Gamma^{\rho}_{\mu\lambda} \equiv \frac{1}{2} g^{\rho\nu} \left(\partial_\mu g_{\nu\lambda} + \partial_\lambda g_{\mu\nu} - \partial_\nu g_{\lambda\mu} \right). \tag{2.3.8}$$

The affine connection is not itself a tensor, because it transforms according to

$$\Gamma'^{\mu}_{\nu\lambda} V'^{\lambda} = \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\sigma\kappa} V^{\kappa} - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\mu}}{\partial x^{\tau} \partial x^{\rho}} V^{\tau}. \tag{2.3.9}$$

Note that the right hand term is precisely the extra term produced when transforming the partial derivative (equation (2.3.6)), but with opposite sign. When added together to form the covariant derivative, see equation (2.3.7), we see that the sum of the two non-tensor elements does transform like a tensor.

A derivation of these results concerning the affine connection can be found in detail in appendix A.

For a tensor with a lower index, the covariant derivative is the same except for a change of the sign in front of the affine connection. For a general mixed tensor, this equates a covariant derivative as defined by

$$\begin{aligned}\nabla_\mu V^{\nu_1,\nu_2,\cdots\nu_k}{}_{\mu_1,\mu_2,\cdots\mu_l} \equiv& \partial_\mu V^{\nu_1,\nu_2,\cdots\nu_k}{}_{\mu_1,\mu_2,\cdots\mu_l} \\ &+ \Gamma^{\nu_1}_{\mu\lambda} V^{\lambda,\nu_2,\cdots\nu_k}{}_{\mu_1,\mu_2,\cdots\mu_l} + \cdots + \Gamma^{\nu_k}_{\mu\lambda} V^{\nu_1,\nu_2,\cdots\lambda}{}_{\mu_1,\mu_2,\cdots\mu_l} \\ &- \Gamma^{\lambda}_{\mu\mu_1} V^{\nu_1,\nu_2,\cdots\nu_k}{}_{\lambda,\mu_2,\cdots\mu_l} - \cdots - \Gamma^{\lambda}_{\mu\mu_l} V^{\nu_1,\nu_2,\cdots\nu_k}{}_{\mu_1,\mu_2,\cdots\lambda}\end{aligned} \tag{2.3.10}$$

In flat space the affine connection is zero and the covariant derivative reduces to the partial derivative. The covariant derivative of a scalar always reduces to the partial derivative of a scalar.

Example 2.2. Going back to the example of the unit 2-sphere, we can calculate the affine connection from the metric. As the only coordinate dependent component of the metric is $g_{\varphi\varphi}$ (and its inverse $g^{\varphi\varphi}$), the symmetries of the affine connection tells us that its non-zero components must have at least two indices which are φ. It turns out that the only non-zero components are $\Gamma^{\theta}_{\varphi\varphi}$, $\Gamma^{\varphi}_{\varphi\theta}$ and $\Gamma^{\varphi}_{\theta\varphi}$. They are

$$\begin{aligned}\Gamma^{\theta}_{\varphi\varphi} &= \frac{1}{2} g^{\theta\theta} (\partial_\varphi g_{\varphi\theta} + \partial_\varphi g_{\theta\varphi} - \partial_\theta g_{\varphi\varphi}) + \frac{1}{2} g^{\theta\varphi} (\partial_\varphi g_{\varphi\varphi} + \partial_\varphi g_{\varphi\varphi} - \partial_\varphi g_{\varphi\varphi}) \\ &= \frac{1}{2}(0 + 0 - 2\sin\theta\cos\theta) + \frac{1}{2} \cdot 0 \cdot (\ldots) \\ &= -\sin\theta\cos\theta\end{aligned} \tag{2.3.11}$$

$$\begin{aligned}\Gamma^{\varphi}_{\varphi\theta} &= \frac{1}{2} g^{\varphi\theta} (\partial_\varphi g_{\theta\theta} + \partial_\theta g_{\theta\varphi} - \partial_\theta g_{\varphi\theta}) + \frac{1}{2} g^{\varphi\varphi} (\partial_\varphi g_{\theta\varphi} + \partial_\theta g_{\varphi\varphi} - \partial_\varphi g_{\varphi\theta}) \\ &= \frac{1}{2} \cdot 0 \cdot (\ldots) + \frac{1}{2} \sin^{-2}\theta (0 + 2\sin\theta\cos\theta - 0) \\ &= \cot\theta\end{aligned} \tag{2.3.12}$$

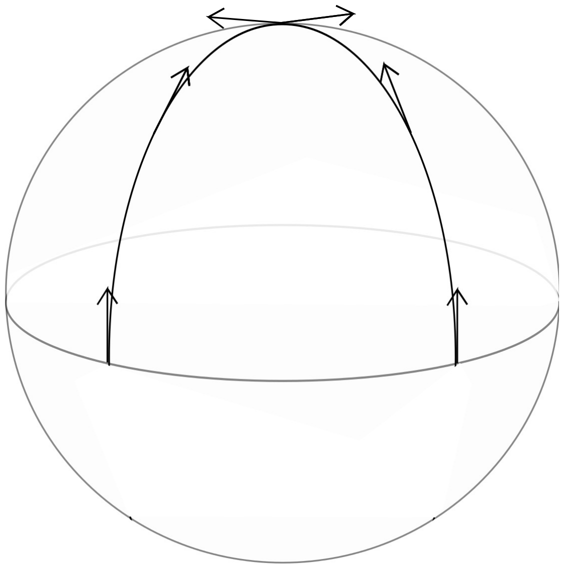


Figure 2.2: Two initially parallel vectors are parallel transported along a spherical manifold until they meet. This illustrates how parallel transport does not necessarily preserve the angles between vectors.

$$\begin{aligned} \Gamma^{\varphi}_{\theta\varphi} &= \{\text{symmetry of the lower indeces}\} \\ &= \cot\theta. \end{aligned} \tag{2.3.13}$$

2.3.3 Parallel Transport and Geodesics

In flat Minkowski spacetime a vector behaves in a uncomplicated manner. Moving a vector around in flat space will not change its size or direction, and vectors that are initially parallel will remain so during translation. A comparison of vectors at different points in space (such as when calculating a relative velocity) is trivial and uniquely defined. For a curved manifold, however, this is not the case. This is illustrated in figure 2.2.

To compare vectors at different points in a curved geometry, we define what is known as **parallel transport**. The equation of parallel transport describes what a tensor looks like after being transport along a chosen path, according to

$$\frac{dx^\mu}{d\lambda}\nabla_\mu V^\nu = 0. \tag{2.3.14}$$

Note that both the tangent vector $dx^\mu/d\lambda$ to the curve and the covariant derivate $\nabla_\mu V^\nu$ are tensors, which makes this a tensor equation. As we parallel transport a tensor along a chosen curve, the tensor will change according to the curvature of the manifold. Depending on the path chosen this outcome will vary. If we parallel transport two vectors V^μ and W^μ along the same curve, the product $g_{\mu\nu}V^\mu W^\mu$ is preserved.

We are now well equipped to introduce the important concept of a **geodesic**. A geodesic is the equivalent of a "straight line", it is the path of a freely moving particle (massive or massless) and a shortest distance between two points in space-time. The **geodesic equation** is a second order differential equation that traces a geodesic $x^\mu(\lambda)$, according to

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0. \tag{2.3.15}$$

If the geodesic is time-like, meaning that it describes a causal path, then the parameter λ is proportional to the proper time. The proper time is maximized along a geodesic. Going back to the twin paradox, it is the twin on earth that moves along a geodesic and therefore experiences the most proper time.

Example 2.3. The geodesics of the 2-sphere are quite simple. Equation (2.3.15) results in the following system of differential equations,

$$\begin{cases} \dfrac{\mathrm{d}^2\theta}{\mathrm{d}\lambda^2} + (-\sin\theta\cos\theta)\left(\dfrac{\mathrm{d}\varphi}{\mathrm{d}\lambda}\right)^2 = 0 \quad (\mu=\theta) \\ \dfrac{\mathrm{d}^2\varphi}{\mathrm{d}\lambda^2} + 2\cot\theta\left(\dfrac{\mathrm{d}\varphi}{\mathrm{d}\lambda}\dfrac{\mathrm{d}\theta}{\mathrm{d}\lambda}\right) = 0 \qquad (\mu=\varphi) \end{cases}. \tag{2.3.16}$$

They result in equator-like lines around the spherical surface, which means that two geodesics on this manifold always meet.

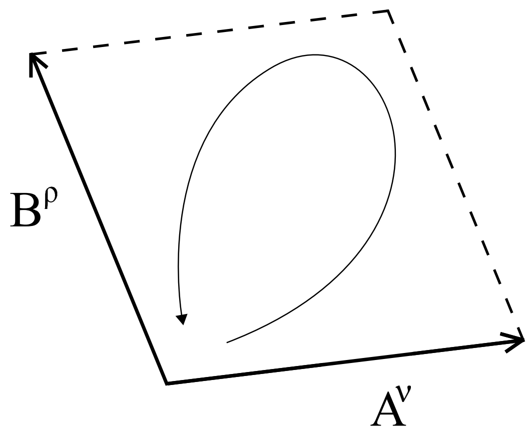


Figure 2.3: Above, the loop through which we parallel transport V^μ is shown, with A^ν and B^ρ as infinitesimally small vectors.

2.3.4 Riemann Tensor

The **Riemann tensor** is an important object in our geometrical theory of gravity. It can be described as follows: Consider parallel transporting a vector V^μ in an infinitesimally small loop along vectors A^ν and B^ρ, as illustrated in figure 2.3. The curvature will result in an infinitesimal change in V^μ, that we call δV^λ. The Riemann tensor $R^\lambda{}_{\mu\nu\rho}$ then relates, according to

$$\delta V^\lambda = A^\nu B^\rho R^\lambda{}_{\mu\nu\rho} V^\mu. \tag{2.3.17}$$

The Riemann tensor has four indices because it relates three spacetime vectors (A^ν, B^ρ and V^μ) to a fourth one (δV^λ).

The strict definition of the Riemann tensor, expressed in terms of the affine connection, is given by

$$R^\lambda{}_{\mu\nu\rho} \equiv \frac{\partial \Gamma^\lambda_{\mu\rho}}{\partial x^\nu} - \frac{\partial \Gamma^\lambda_{\mu\nu}}{\partial x^\rho} + \Gamma^\lambda_{\nu\kappa}\Gamma^\kappa_{\mu\rho} - \Gamma^\lambda_{\rho\kappa}\Gamma^\kappa_{\mu\nu}. \tag{2.3.18}$$

It follows from the definition that the Riemann tensor fulfills some symmetries, such that it is antisymmetric with respect to its two last indices, which equates to

$$R^\lambda{}_{\mu\nu\rho} = -R^\lambda{}_{\mu\rho\nu}. \tag{2.3.19}$$

The Riemann tensor, after lowering its first index, also fulfills symmetries

$$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu} \tag{2.3.20}$$

$$R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}. \tag{2.3.21}$$

If it is possible to find a coordinate system for which the metric is constant with respect to its coordinates, the Riemann tensor is zero, and vice versa. This is the case because they are both equivalent with the manifold being globally flat.

Example 2.4. Given the affine connection of the unit 2-sphere from example 2.2, we can calculate the Riemann tensor using equation (2.3.18). This gives us a component

$$\begin{aligned} R^{\theta}{}_{\varphi\theta\varphi} &= \partial_\theta \Gamma^{\theta}_{\varphi\varphi} - \partial_\varphi \Gamma^{\theta}_{\theta\varphi} + (\Gamma^{\theta}_{\theta\theta}\Gamma^{\theta}_{\varphi\varphi} + \Gamma^{\theta}_{\theta\varphi}\Gamma^{\varphi}_{\varphi\varphi}) - (\Gamma^{\theta}_{\varphi\theta}\Gamma^{\theta}_{\theta\varphi} + \Gamma^{\theta}_{\varphi\varphi}\Gamma^{\varphi}_{\theta\varphi}) \\ &= \partial_\theta(-\sin\theta\cos\theta) - 0 + (0+0) - (0 + (-\sin\theta\cos\theta)\cot\theta) \\ &= -\cos^2\theta + \sin^2\theta + \cos^2\theta \\ &= \sin^2\theta. \end{aligned} \tag{2.3.22}$$

The other three non-zero components of the Riemann tensor can be calculated in the same way from equation (2.3.18) or reached through the Riemann tensor's symmetries by lowering and raising indices. The components are

$$R^{\theta}{}_{\varphi\varphi\theta} = -\sin^2\theta \tag{2.3.23}$$

$$R^{\varphi}{}_{\theta\theta\varphi} = -1 \tag{2.3.24}$$

$$R^{\varphi}{}_{\theta\varphi\theta} = 1. \tag{2.3.25}$$

2.3.5 Ricci Tensor and Ricci Scalar

While the Riemann tensor itself will only be implicitly applied, its direct cousins, the **Ricci tensor** and **Ricci scalar**, are explicit component of Einstein's equations. The Ricci tensor $R_{\mu\nu}$ is found by contracting the first and third indices of the Riemann tensor, according to

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}. \tag{2.3.26}$$

The Ricci tensor is symmetric with respect to its two indices.

The Ricci scalar R is given by a further contraction, according to

$$R = R^{\mu}{}_{\mu} = g^{\mu\nu}R_{\mu\nu}. \tag{2.3.27}$$

The Ricci scalar describes the magnitude and type of curvature. If the Ricci scalar is positive the curvature is positively definite. If it is negative the curvature is negatively definite, which would correspond to a saddle point on the manifold.

Example 2.5. We now calculate the Ricci tensor and scalar of the unit 2-sphere from equation (2.3.26) and the Riemann tensor of the manifold, see example 2.4. This gives us

$$R_{\varphi\varphi} = R^{\lambda}{}_{\varphi\lambda\varphi} = R^{\theta}{}_{\varphi\theta\varphi} = \sin^2\theta \tag{2.3.28}$$

$$R_{\varphi\theta} = R^{\lambda}{}_{\varphi\lambda\theta} = 0 \tag{2.3.29}$$

$$R_{\theta\varphi} = R^{\lambda}{}_{\theta\lambda\varphi} = 0 \tag{2.3.30}$$

$$R_{\theta\theta} = R^{\lambda}{}_{\theta\lambda\theta} = R^{\varphi}{}_{\theta\varphi\theta} = 1. \tag{2.3.31}$$

We now calculate the Ricci scalar with equation (2.3.26), which results in

$$R = g^{\mu\nu}R_{\mu\nu} = g^{\varphi\varphi}R_{\varphi\varphi} + g^{\theta\theta}R_{\theta\theta} = \sin^{-2}\theta\sin^2\theta + 1\cdot 1 = 2. \tag{2.3.32}$$

The Ricci scalar is positive and independent of its coordinates, which signifies that we have a curvature that is positive definite everywhere on the manifold.

2.4 Einstein's Equations

Einstein's equations are the set of equations that describe how space is curved due to gravity. The solution to these equations will be a metric.

Einstein's equations are

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}, \tag{2.4.1}$$

where G is the gravitational constant and $T_{\mu\nu}$ is the stress-energy tensor, to be introduced in section 2.4.1.

One way to derive Einstein's equations is to make a qualified ansatz based on a combination of relevant physical quantities and an analogy to Maxwell's equations. Constants can be determined by taking the Newtonian limit; general relativity must be consistent with Newtonian gravity where the latter agrees with experimental results. How the Newtonian limit is used practically will be explained in better detail later on.

Another method to find Einstein's equations is based on the action principle and variational calculus. Deriving Einstein's equations from the action principle will be useful later on. Variational calculus is used to obtain the Schwarzschild solution in chapter 4, but is also an important tool when calculating geodesics as it is often easier than solving the geodesic equation.

2.4.1 Einstein's Equations From an Ansatz

In order to describe how spacetime is effected by matter, we first need to describe matter and spacetime. To describe spacetime, the Riemann tensor and the metric are natural ingredients. For matter, we first need to introduce the **stress-energy tensor**.

The stress-energy tensor is a tensor describing the properties of matter, such as density, flux and pressure. It is symmetric and of rank two (it has two indices). A special case, and commonly used in for example cosmology, is a perfect fluid. A perfect fluid can be completely described by its pressure p and its energy density ρ, thus having no heat conduction and no viscosity. Radiation and, in the large scale, dust (ordinary matter) are examples of perfect fluids. A perfect fluid has the stress tensor

$$T_{\mu\nu} = (p+\rho)U_\mu U_\nu + p g_{\mu\nu}, \tag{2.4.2}$$

where U^μ is the four-velocity of the matter. For dust in the local rest frame, the stress-energy tensor takes the form

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.4.3}$$

Now we have the ingredients for Einstein's equations. To get some more substance to the ansatz, an analogy to Maxwell theory is useful. General relativity and electromagnetic theory share several fundamental properties.

In Maxwell theory, there is the electromagnetic tensor $F^{\mu\nu}$. It describes the field from electric and magnetic charges and is composed of the components of the electric and magnetic fields. $F^{\mu\nu}$ fulfills the Bianchi identity, which means that

$$\partial_{[\mu} F_{\nu\rho]} = 0, \tag{2.4.4}$$

where $[\,]$ is the antisymmetric part of a tensor.

In Maxwell theory, there is a four-current J^μ which is a four-vector consisting of the current as well as the electric charge. The four-current is conserved, so that

$$\partial_\mu J^\mu = 0. \tag{2.4.5}$$

As for general relativity, the analogy to the electromagnetic tensor is the Riemann tensor, the tensor that describes the gravitational field and how space is affected by matter. The Bianchi identity for the Riemann tensor is

$$\nabla_{[\lambda} R_{\mu\nu]\rho\sigma} = 0. \tag{2.4.6}$$

The analogy to the four-current is the the stress-energy tensor $T^{\mu\nu}$. Just like J^μ, $T^{\mu\nu}$ is conserved:

$$\nabla_\mu T^{\mu\nu} = 0. \tag{2.4.7}$$

With so many similarities, there must certainly also be some similarities regarding these equations. As for electromagnetic fields, Maxwell's equations are equivalent to the Bianchi identity together with the following equation[2]

$$\partial_\nu F^{\mu\nu} = J^\mu. \tag{2.4.8}$$

Since the Ricci tensor is a tensor with the right rank, a first guess of Einstein's equations might be

$$R_{\mu\nu} \stackrel{?}{=} \kappa T_{\mu\nu} \quad \kappa\ const. \tag{2.4.9}$$

Both sides of this equation are symmetric, which is a good sign. Recalling covariant conservation however, we see that the right-hand side is conserved while the left-hand side is not. To find something that is conserved, the Bianchi identity can be rewritten as

$$\nabla^\mu \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) = 0. \tag{2.4.10}$$

A reasonable modification to the initial ansatz would therefore be

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa T_{\mu\nu}. \tag{2.4.11}$$

In order to determine κ it is convenient to have some reference where we know what the equations will look like. In our case, this reference is the Newtonian limit. We do not go through all the calculations here (a more detailed discussion is provided in appendix B) but the main idea is instructive. The key points are that the gravitational field is weak and static while velocities are low, implying that the rest energy component of the stress-energy tensor is much greater than the other components. We may therefore restrict our analysis to this component, T_{tt}. For a weak field, we expect only small deviations from the Minkowski metric, but we cannot set the metric equally to Minkowski as it would yield no new information. The most simple small deviation from Minkowski would be a constant:

$$g_{tt} = -1 + h_{tt}. \tag{2.4.12}$$

Neglecting all higher-order terms, we get (after some calculations)

$$R_{tt} \underset{\text{weak field}}{\longrightarrow} \nabla^2 h_{tt} = \kappa T_{tt}. \tag{2.4.13}$$

[2]Just plug in the definition of the electromagnetic stress tensor and write out the components; the equations will be identical to Maxwell's equations.

Comparing with Newtonian gravity, we have

$$\nabla^2\phi = 4\pi G\rho, \tag{2.4.14}$$

where ϕ is the gravitational potential. Taking the Newtonian limit yields

$$h_{tt} = -2\phi. \tag{2.4.15}$$

We sum everything up into **Einstein's equations:**

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}. \tag{2.4.16}$$

As elegant and comprehensive these might look, we should remember that the Ricci tensor in four dimensions has $4^2 = 16$ components, of which (from symmetry) 10 are independent. Einstein's equations are therefore 10 coupled differential equations and an analytic solution is not always possible to obtain.

2.4.2 The Action Principle

An action is a functional, meaning that it takes a function $\phi(x)$ and returns a scalar $S[\phi]$. The action principle states that the function $\phi(x)$ must give a stationary action. In variational calculus, stationary means that the action remains constant under a perturbation $\delta\phi$ to the first order. It can thus be used to derive equations of motion.

The action is typically an integral over time, for a system from one point in time to another,

$$S[x] = \int \mathrm{d}t\mathcal{L}. \tag{2.4.17}$$

with $\mathcal{L}$ being the Lagrangian. As the Lagrangian is an energy, the unit is found to be [energy]·[time]. When concerning fields, as is studied in this report, the integral includes the space variables as well, implying a replacement of the Lagrangian with a Lagrangian density.

When saying that an action is stationary we mean that its functional derivative, $\frac{\delta S}{\delta x(t)}$, is zero. Functional derivatives are much like partial derivatives, $\frac{\partial}{\partial x^i}$, except that instead of an index taking discrete values, there's a continuous "index" t. They can therefore be treated much like such derivatives; using the chain rule, product rule, etc.

It is also worth noting that much like

$$\frac{\partial x^j}{\partial x^i} = \delta^j_i \tag{2.4.18}$$

we have

$$\frac{\delta x(t')}{\delta x(t)} = \delta(t' - t), \tag{2.4.19}$$

an observation that will be useful when integrating by parts and evaluating integrals.

2.4.3 An Action Principle for Einstein's Equations

In this section, Einstein's equations will be derived from the action principle.

The solution to Einstein's equations is the metric, and therefore we will vary the action with respect to the metric. The action is written as

$$S[g_{\mu\nu}] = \frac{1}{8\pi G} \int \mathrm{d}^4x \sqrt{g} R, \tag{2.4.20}$$

where g is the absolute value of the determinant of $g^{\mu\nu}$. Neither $\sqrt{g}$ nor d^4x are tensors in and of themselves[3], but transform as a tensor together, thus making the volume element invariant.

The action being stationary with respect to small variations of the metric means that

$$0 = \delta S, \qquad g_{\mu\nu} \to g'_{\mu\nu} + \delta g_{\mu\nu}. \tag{2.4.21}$$

To find the variation of $\sqrt{g}R = \sqrt{g} g^{\mu\nu} R_{\mu\nu}$, we must describe $\delta\sqrt{g}$, $\delta g^{\mu\nu}$ and $\delta R_{\mu\nu}$.

With $g^{\mu\nu}$ being the inverse of $g_{\mu\nu}$, one gets

$$\begin{aligned} 0 = \delta(g^{\mu\rho} g_{\rho\nu}) &= \delta g^{\mu\rho} g_{\rho\nu} + g^{\mu\rho} \delta g_{\rho\nu} \\ &\Rightarrow \\ \delta g^{\mu\sigma} &= -g^{\mu\rho} \delta g_{\rho\nu} g^{\nu\sigma}, \end{aligned} \tag{2.4.22}$$

which holds for arbitrary tensors.

The variation of $\sqrt{g}$ can be found using the following statement which is proved in appendix D:

$$\delta(\mathrm{Det}\, A) = \mathrm{Det}\, A \, \mathrm{Tr}(A^{-1} \delta A). \tag{2.4.23}$$

From this statement and the chain rule we can finally obtain the variation of $\sqrt{g}$. Note how easily the trace of a product of two matrices is expressed in tensor formalism: after writing out the product the two indices are set equal to produce

[3]Strictly speaking, $\sqrt{g}$ and d^4x are actually tensor densities and transform as tensors except for a weight of proportionality.

a sum over all diagonal elements. Thereby,

$$\begin{aligned}\delta\sqrt{g} &= \frac{1}{2}\frac{1}{\sqrt{g}}\delta g \\ &= \frac{1}{2}\sqrt{g}g^{\mu\nu}\delta g_{\mu\nu}.\end{aligned} \tag{2.4.24}$$

Lastly for the variation of the Ricci tensor, we start from the Riemann tensor. Adding and subtracting the same connection yields

$$R^{\rho}{}_{\nu\mu\sigma} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\sigma \Gamma^\rho_{\mu\nu} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\sigma\lambda}\Gamma^\lambda_{\nu\mu} \tag{2.4.25}$$

$$\Rightarrow$$

$$\delta R^{\rho}{}_{\nu\mu\sigma} = \nabla_\mu \delta\Gamma^\rho_{\nu\sigma} - \nabla_\sigma \delta\Gamma^\rho_{\mu\nu}. \tag{2.4.26}$$

The variation of the Ricci tensor will be the contraction

$$\delta R_{\nu\sigma} = \nabla_\mu \delta\Gamma^\mu_{\nu\sigma} - \nabla_\sigma \delta\Gamma^\mu_{\mu\nu}. \tag{2.4.27}$$

If the action is given by equation (2.4.20), the variation will be:

$$\begin{aligned}\delta S &= \frac{1}{8\pi G}\int \mathrm{d}^4x \left(\delta\sqrt{g}g^{\mu\nu}R_{\mu\nu} + \sqrt{g}\delta g^{\mu\nu}R_{\mu\nu} + \sqrt{g}g^{\mu\nu}\delta R_{\mu\nu}\right) \\ &= \frac{1}{8\pi G}\int \mathrm{d}^4x \left(\frac{1}{2}\sqrt{g}g^{\rho\sigma}\delta g_{\rho\sigma}g^{\mu\nu}R_{\mu\nu} - \sqrt{g}g^{\mu\rho}\delta g_{\rho\sigma}g^{\sigma\nu}R_{\mu\nu} + \right. \\ &\quad \left. +\sqrt{g}g^{\mu\nu}\left(\nabla_\rho\delta\Gamma^\rho_{\mu\nu} - \nabla_\mu\delta\Gamma^\rho_{\rho\nu}\right)\right)\end{aligned} \tag{2.4.28}$$

The contribution of the $\delta\Gamma$-terms is zero, as will be shown below. Due to these terms beeing zero, the variation will then be

$$\delta S = -\frac{1}{8\pi G}\int \mathrm{d}^4x\sqrt{g}\left(R^{\rho\sigma} - \frac{1}{2}g^{\rho\sigma}R\right)\delta g_{\rho\sigma}. \tag{2.4.29}$$

To prove that the terms containing the affine connection vanish, we start with some simplification. As the difference between two connections is a tensor, we can write the integrand as the covariant derivative of some tensor V^ρ.

$$\int \mathrm{d}^4x\sqrt{g}g^{\mu\nu}\left(\nabla_\rho\delta\Gamma^\rho_{\mu\nu} - \nabla_\mu\delta\Gamma^\rho_{\rho\nu}\right) \equiv \int \mathrm{d}^4x\sqrt{g}\nabla_\rho V^\rho. \tag{2.4.30}$$

Expanding the covariant derivative, using equation (2.3.8) for the affine connection and noting that the results from variation, in particular (2.4.24), are valid for partial derivatives as well, yields

$$\begin{aligned}\int \mathrm{d}^4x\sqrt{g}\nabla_\rho V^\rho &= \int \mathrm{d}^4x\sqrt{g}\left(\partial_\rho V^\rho + \frac{1}{2}g^{\mu\lambda}\partial_\rho g_{\mu\lambda}V^\rho\right) \\ &= \int \mathrm{d}^4x\sqrt{g}\left(\partial_\rho V^\rho + \frac{1}{\sqrt{g}}\partial_\rho\sqrt{g}V^\rho\right) \\ &= \int \mathrm{d}^4x\sqrt{g}\left(\frac{1}{\sqrt{g}}\partial_\rho\left(\sqrt{g}V^\rho\right)\right) \\ &= \int \mathrm{d}^4x\partial_\rho\left(\sqrt{g}V^\rho\right).\end{aligned} \tag{2.4.31}$$

The divergence theorem and the fact that variations vanish at infinity makes this term zero, giving the variation of equation (2.4.29). This result emphasizes why the factor $\sqrt{g}$ must be added to make an invariant.

For the action to be stationary with respect to small changes in the metric, the integrand of equation (2.4.29) needs to be zero, that is

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = 0, \tag{2.4.32}$$

equal to Einstein's equations in vacuum.

For the vacuum case, Einstein's equations can be simplified by multiplying with the inverse metric. Since $g^{\mu\nu}g_{\mu\nu}$ is the trace of the identity matrix which is the number of dimensions, we get

$$\begin{aligned} g^{\mu\nu}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) &= 0 \\ &\Rightarrow \\ R - \frac{1}{2}g^{\mu\nu}g_{\mu\nu}R &= 0 \\ &\Rightarrow \\ R &= 0. \end{aligned} \tag{2.4.33}$$

and (2.4.32) can therefore be simplified to

$$R_{\mu\nu} = 0. \tag{2.4.34}$$

Einstein's equations not restricted to the case of vacuum can be obtained by adding a matter term to the action. This would yield

$$
\begin{aligned}
S &= \frac{1}{8\pi G}\int \mathrm{d}^4x\sqrt{g}(R+S_{matter}) \\
&\Rightarrow \\
\frac{\delta S}{\delta g_{\mu\nu}(x)} &= -\frac{\sqrt{g}}{8\pi G}\left(R^{\mu\nu}-\frac{1}{2}g^{\rho\sigma}R\right)+\frac{\delta S_{matter}}{\delta g_{\mu\nu}(x)} \\
&\Rightarrow \\
0 &= -\frac{1}{8\pi G}\left(R^{\mu\nu}-\frac{1}{2}g^{\rho\sigma}R\right)+\frac{1}{\sqrt{g}}\frac{\delta S_{matter}}{\delta g_{\mu\nu}(x)} \\
&\Rightarrow \\
R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R &= 8\pi G\frac{1}{\sqrt{g}}\frac{\delta S_{matter}}{\delta g_{\mu\nu}(x)} \\
&\Rightarrow \\
R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R &= 8\pi GT^{\mu\nu}, \qquad (2.4.35)
\end{aligned}
$$

where we have defined

$$T^{\mu\nu}(x)=\frac{1}{\sqrt{g}}\frac{\delta S_{matter}}{\delta g_{\mu\nu}(x)}. \qquad (2.4.36)$$

In the action of equation (2.4.20) we could also have added a constant. That would have yielded the full Einstein field equations including a cosmological constant:

$$R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R+g^{\mu\nu}\Lambda=8\pi GT^{\mu\nu}. \qquad (2.4.37)$$

The cosmological constant has an interesting history. It was initially introduced by Einstein to prevent the theory from describing an expanding universe. When observations showed that the universe is expanding, Einstein rejected the suggestion of the cosmological constant and called it the biggest mistake of his life [1]. However, decades later observations indicated the reintroduction of a cosmological constant, for example in terms of the accelerating universe. Nevertheless, observations show that it must be very small, actually much smaller than particle physics would predict[4]. Therefore, we assume the cosmological constant to be zero throughout this text.

2.5 Black Hole Solutions

The **Schwarzschild solution** or **Schwarzschild metric** describes the curved spacetime around a static, spherically symmetric, non-charged, non-rotating mass distribution in vacuum. The metric is given by

$$\mathrm{d}s^2 = -\left(1 - \frac{2GM}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2, \tag{2.5.1}$$

where G is the gravitational constant and M is the mass, or total energy of the mass distribution. The coordinates used are spherical, where $\mathrm{d}\Omega^2$ is short for the angular distance, such that

$$\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2. \tag{2.5.2}$$

Note that some interesting things are happening, especially at the radius $r_s = 2GM$. This is the famous **event horizon**, the point of no return that not even light can escape. The event horizon looks like it could be a singularity but is actually not, because we can choose a coordinate system in which the event horizon takes a non-singular expression. This is, on the other hand, not possible for the singularity in the center. Observed from the outside, a particle falling into a black hole never actually reaches the event horizon. What is observed is that the particle slows down, becomes red-shifted (due to Doppler effect) and fades away. From the particles perspective however, it does reach both the event horizon and the singularity at $r = 0$, and it does so quite quickly in terms of proper time. In this sense, from the perspective of a particle on the outside there is no past inside the event horizon. Likewise, from the perspective of a particle inside the event horizon there is no future on the outside.

The spacetime of the Schwarzschild solution satisfies the condition of vacuum, which is that the Ricci vector is zero, $R_{\mu\nu} = 0$. The most straight forward way to arrive at this solution is to make an ansatz and adjust it to fulfill the vacuum condition and Newtonian limit at infinity.

For a charged, static, spherically symmetric mass distribution the spacetime geometry in vacuum is described by the **Reissner-Nordström solution**. This is given by

$$\mathrm{d}s^2 = -\left(1 - \frac{2MG}{r} + \frac{Q^2}{r^2}\right)\mathrm{d}t^2 + \left(1 - \frac{2MG}{r} + \frac{Q^2}{r^2}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2, \tag{2.5.3}$$

where Q is the charge and when this is zero the metric is precisely the Schwarzschild metric.

There will be a event horizon where the radial component of the metric is singular. We can find where this is by solving the following equation,

$$1 - \frac{2MG}{r_e} + \frac{q^2}{r_e^2} = 0 \tag{2.5.4}$$

$$\Rightarrow$$

$$r_e^{\pm} = MG \pm \sqrt{(MG)^2 - Q^2}. \tag{2.5.5}$$

Thus we have two event horizons, both closer to the centrum of the black hole then the event horizon of the Schwarzschild solution at $2MG$, for $Q > 0$. For an outside observer the event horizon at r_e^+ has the same effect as the Schwarzschild event horizon. For the observer, nothing will ever reach $r = r_e^+$ as anything falling towards the black hole slows down and becomes all the more redshifted. For the object falling into the black hole the situation is the same at r_e^+ as for the Schwarzschild event horizon; after passing this radius the only possible future is towards the center of the black hole. Therefore there is no possibility of turning back again after passing $r = r_e^+$, as this would violate causality. Not until the object passes the second event horizon, at r_e^-, can it turn back, because causality does not force it towards the center anymore. This is contrary to the Schwarzschild solution, where the future of any object inside the event horizon is always towards $r = 0$. In theory an object inside r_e^- could return and escape the charged black hole. More discussion about this can be find in chapter 7 in Carroll [1].

For this to be a realistic solution we can assume that $MG > Q$. For $MG = Q$ the event horizons would coincide and this would give a so called extremal black hole. This would require that all the energy would come from charge and nothing from rest mass. Particles with this property, zero rest mass and non-zero net charge, have never been detected. The extremal black hole solution is, however, of much interest in theoretical work. Particularly, due to the cancellation of the gravitational and electromagnetic force at $r = MG$ where an object would not feel any pull from the black hole.

2.5.1 A Derivation of the Schwarzschild Metric

To derive the Schwarzschild metric we start off with the assumptions of spherical symmetry and that the metric is static, meaning that the metric is invariant with respect to a translation in time. Because of these symmetries, the components of the metric will depend on the radial component only.

Since the sphere is maximally symmetric we can use a theorem in chapter 13.5 in Weinberg's book [2] stating that it is possible to choose a metric without mixed terms containing angular coordinates, such as $\mathrm{d}r\mathrm{d}\theta$. The only possible mixed term is therefore $\mathrm{d}t\mathrm{d}r$. To have the same signature as for Minkowski metric $(-+++)$, a minus sign is added to the time term. The ansatz would be

$$\mathrm{d}s^2 = -a(r)\mathrm{d}t^2 + b(r)\mathrm{d}r^2 + c(r)\mathrm{d}t\mathrm{d}r + d(r)\mathrm{d}\Omega^2, \tag{2.5.6}$$

where $a(r)$, $b(r)$, $c(r)$ and $d(r)$ are arbitrary functions.

In Euclidean geometry, there is a factor of r^2 before the angular components. It would certainly facilitate the derivation if we could have this factor r^2. Another desirable simplification would be to get rid of the mixed term. In fact, both these features can be achieved by the same argument; to redefine the coordinates.

This leads to the final, simplest possible ansatz satisfying the assumption of spherical symmetry and stationarity:

$$\mathrm{d}s^2 = -A(r)\mathrm{d}t^2 + B(r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2. \tag{2.5.7}$$

So far, only geometry has been used. To begin with the physics, we want the solution in vacuum. The equation to solve is therefore, according to (2.4.34),

$$R_{\mu\nu} = 0. \tag{2.5.8}$$

To explicitly write out these equations, we have to calculate each term in the affine connection, then form the Riemann tensor and finally the Ricci tensor. All details will not be presented here since the methods are analogous to those of the examples of sections 2.3.2 through 2.3.5.

The Riemann tensor is easiest calculated using the covariant form so that all symmetries can be used. After some analysis, we find that the only possible non-zero components are

$$R_{rtrt}, \quad R_{r\theta r\theta}, \quad R_{r\varphi r\varphi}, \quad R_{\theta t\theta t}, \quad R_{t\varphi t\varphi}, \quad R_{\theta\varphi\theta\varphi} \quad \text{and} \quad R_{r\varphi\theta\varphi}. \tag{2.5.9}$$

After having calculated the Riemann tensor, it can be contracted to the Ricci tensor. The Ricci tensor will be diagonal and (2.4.34) will be the following coupled differential equations:

$$0 = R_{rr} = -\frac{A''}{2A} + \frac{1}{4}\left(\frac{A'}{A}\right)\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{1}{r}\left(\frac{B'}{B}\right), \tag{2.5.10}$$

$$0 = R_{\theta\theta} = 1 - \frac{r}{2B}\left(-\frac{A'}{A} + \frac{B'}{B}\right) - \frac{1}{B}, \tag{2.5.11}$$

$$0 = R_{\varphi\varphi} = \sin^2\theta R_{\theta\theta}, \tag{2.5.12}$$

$$0 = R_{tt} = \frac{A''}{2B} - \frac{1}{4}\left(\frac{A'}{B}\right)\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{1}{r}\left(\frac{A'}{B}\right). \tag{2.5.13}$$

We now form

$$0 = \frac{R_{rr}}{B} + \frac{R_{tt}}{A} = \frac{1}{rB}\left(\frac{B'}{B} + \frac{A'}{A}\right). \tag{2.5.14}$$

This implies that the factor AB must be constant since

$$\begin{aligned} \frac{B'}{B} &= -\frac{A'}{A} \\ \partial_r \ln B &= -\partial_r \ln A \\ AB &= const. \end{aligned} \tag{2.5.15}$$

The limit $r \to \infty$ should result in the Minkowski metric, so that

$$\lim_{r\to\infty} A = \lim_{r\to\infty} B = 1. \tag{2.5.16}$$

Therefore, the constant should be equal to one and A is the inverse of B. Reinserting into Einstein's equations yields

$$B = 1 + \frac{\gamma}{r}, \quad \gamma\, const. \tag{2.5.17}$$

To determine the constant γ, we require the solution to be consistent in the Newtonian limit, which is explained in more detail in appendix B. For large radii, g_{tt} must then satisfy

$$g_{tt} \to -1 - 2\phi, \tag{2.5.18}$$

where the gravitational potential is $\phi = -\frac{MG}{r}$. This gives $B(r) = \left(1 - \frac{2GM}{r}\right)$ and the Schwarzschild metric is

$$\mathrm{d}s^2 = -\left(1 - \frac{2GM}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2. \tag{2.5.19}$$

3

Group Theory and Symmetries

In the previous chapter we introduce tensors and use these to describe physics in a consistent way, regardless of the choice of coordinate system. The tensors are not invariant under coordinate transformation, but transform in a well-known manner. The next natural step is to find transformations that leave the physics, or rather the expressions derived, unchanged. If such transformations exists, we say that the system has certain symmetries. There are many different kinds of possible symmetries, but they can mainly be divided into geometrical symmetries and field symmetries. As an example of a geometrical symmetry, consider an equilateral triangle. If you rotate it $60°$ you end up with something identical to what you started with. As an example of a field symmetry there is Maxwell theory, where there is a degree of freedom in choosing the potential without affecting the electromagnetic fields at all.

An important and powerful tool for describing symmetries in a compact and general manner is group theory. In section 3.1 we go through the axioms of groups and some important definitions. Thereafter, in section 3.2 we study continuous and differentiable groups, called Lie groups, and their important algebras. After establishing the groundwork of group theory, we consider groups that can be represented with matrices which are important for the project. Finally, in section 3.4 we introduce some miscellaneous theorems of group theory that are also of importance in the next chapter.

This chapter is mainly based on the texts by Hall [5] and Cahn [6], which the interested reader may find useful to obtain a deeper knowledge in the large mathematical field that is group theory. Some specific definitions and properties have been gathered from a wide array of other texts [7–11].

3.1 Basic Theory and Definitions

Group theory is a large field of mathematics and impossible to cover in its entirety in a text such as this. We therefore instead focus on the parts needed for the outcome of this project. In this section we regard the axioms of a group, and basic properties of groups. We also define some relations between groups, such as isomorphisms and subgroups as well as means to generate other manifolds as cosets.

The core of group theory relies on four axioms which at a first glance may seem quite simple.

Definition 3.1. *A group $\boldsymbol{G}$ is a set of elements $\{g_i\}$ that together with an operation $*$ satisfies the following axioms:*

(i) ***Closure***
If g_i,$g_j \in \boldsymbol{G}$ there uniquely exists an element $g_k \in \boldsymbol{G}$ such that

$$g_i * g_j = g_k. \tag{3.1.1}$$

(ii) ***Unit element***
There exists a unit element $e \in \boldsymbol{G}$ which for all $g \in \boldsymbol{G}$ satisfies

$$e * g = g * e = g. \tag{3.1.2}$$

(iii) ***Inverse element***
For every $g \in \boldsymbol{G}$ there exists an inverse element denoted $g^{-1} \in \boldsymbol{G}$ where

$$g * g^{-1} = g^{-1} * g = e. \tag{3.1.3}$$

(iv) ***Associativity***
It holds true that for all sets of elements g_i,g_j,$g_k \in \boldsymbol{G}$,

$$g_i * (g_j * g_k) = (g_i * g_j) * g_k. \tag{3.1.4}$$

The operator symbol “$*$” is often omitted in writing. There is a number of basic properties of groups whose proofs will not be included here for brevity. Some of these are: the uniqueness of unit element, the uniqueness of inverses and that there cannot exist an inverse only from one side. To clarify these axioms we can use the following example.

Example 3.1. The real numbers form a group, $\mathbb{R}$, under addition. The unit element is 0 and the inverse element to any $x \in \mathbb{R}$ is $-x$. Let us show that this is the case:

(i) Closure: the sum of two real numbers is a real number.

(ii) Unit element: $0 + x = x + 0 = x$, and $0 \in \mathbb{R}$.
(iii) Inverse element: $x + (-x) = (-x) + x = 0$, and if $x \in \mathbb{R} \Rightarrow -x \in \mathbb{R}$.
(iv) Associativity: Addition is associative.

It is worth noting that the axioms do not require the operation to be commutative, although in this case it is. Groups where the elements commute under the operator are important and thus given their own name.

Definition 3.2. *An* ***abelian*** *or* ***commutative*** *group is a group where* g_i,$g_j \in \boldsymbol{G} \Rightarrow g_i * g_j = g_j * g_i$.

Sometimes it might be of interest to examine a smaller part of a group, which naturally leads to defining the **subgroups**.

Definition 3.3. $\boldsymbol{S}$ *is a* ***subgroup*** *to a group* $\boldsymbol{G}$ *if* $s \in \boldsymbol{S} \Rightarrow s \in \boldsymbol{G}$ *and* $\boldsymbol{S}$ *is a group under the same operator as* $\boldsymbol{G}$.

Although a group can have many subgroups, every group has at least two subgroups: the unit element and the group itself.

Example 3.2. The set of all integers form a subgroup $\mathbb{Z}$ of $\mathbb{R}$ in Example 3.1. The set of all natural (non-negative) numbers however do not.

Both the integers and natural numbers are subsets to $\mathbb{R}$. The unit element is in both sets, and closure is fulfilled. Associativity follows from the sets being subsets to a group, and left is only to show existence of an inverse. The inverse has to once again be $-x$, which is an integer if x is an integer. However if x is a natural number greater than zero, the inverse will be negative and not a natural number. We have thus proven our claim.

Note that the associativity axiom is always true for any subset of any group, as any three elements chosen from the subset lie in the group, and are associatively related from the group axiom.

Now let us look at an example where some of the usefulness of group theory regarding symmetries may be discerned in a simple fashion.

Example 3.3. An equilateral triangle has certain rotational symmetries which form a group **T**.

The elements in **T** are the rotations in its plane that keep the triangle looking the same, as shown in figure 3.1. So, as an example we can have three elements a, b, c in the group T. Here a represents no rotation at all, i.e $\theta = 0$, b is a

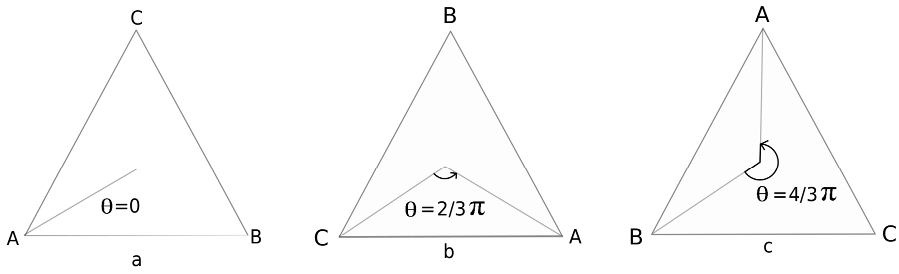


Figure 3.1: An equilateral triangle.

rotation of $\theta = 2\pi/3$ and c a rotation of $\theta = 4\pi/3$. The operator $*$ in this case is $g_i * g_j \Rightarrow \theta_i + \theta_j$. All of these elements fulfill the requirement of keeping the triangle looking as it is. The group axioms are fulfilled: any combination of the elements keep the triangle looking the same, the rotations are associative, there exists a unit element and there exists a complement to make a full turn rotation - an inverse.

The symmetries of the triangle, the transformations that leave it the same, thus form a group. Of course there are additional symmetries one could add, such as the reflection symmetries, but they would still together be a group. The proof of this is left to the reader.

The previous example was an example of a group with a **finite** number of elements and we have previously seen groups with **countably** or **uncountably infinite** elements. The first two typically represent **discrete symmetries** whereas the latter a **continuous symmetry**. Both these kinds of symmetries are of importance when describing nature, although only continuous symmetries will be discussed in this work. An example of a continuous symmetry could be the rotations of a sphere. These are described by the uncountably infinite group SO(3). SO(3) is a member of a large family of groups we will get to know better in section 3.3.

As we have so far only defined properties of groups, the next step is to act on groups to generate new manifolds and in the best scenarios even groups. There are a few ways to accomplish this, here we will consider the cosets and the quotient spaces. These are roughly “element-wise multiplication” and “element-wise modulus”.

Definition 3.4. *A group $\boldsymbol{G}$ and a subgroup $\boldsymbol{H}$, combined with an element $g \in \boldsymbol{G}$ can be used to form the left and right* **coset** *of $\boldsymbol{H}$. Here left and right respectively denoting which side of $\boldsymbol{H}$ the element g is applied:*

$$g\boldsymbol{H} = \{gh : h \in \boldsymbol{H}\}, \tag{3.1.5}$$

$$\boldsymbol{H}g = \{hg : h \in \boldsymbol{H}\}. \tag{3.1.6}$$

These, however, are generally not groups.

Definition 3.5. *A* ***quotient space*** *or* ***quotient set****, denoted* $\boldsymbol{G}/\boldsymbol{N}$*, where* $\boldsymbol{N}$ *is a subgroup to* $\boldsymbol{G}$*, is the set of all left cosets* $\{g\boldsymbol{N} : g \in \boldsymbol{G}\}$*. Similarly* $\boldsymbol{N}\backslash\boldsymbol{G}$ *the set of right cosets. Sometimes the quotient spaces are also called* ***coset spaces****.*

If $\mathbf{N}$ in the last definition is a **normal** subgroup, that is

$$gng^{-1} \in \mathbf{N}, \forall n \in \mathbf{N}, \forall g \in \mathbf{G}, \tag{3.1.7}$$

then the quotient space forms a group. This follows as for instance

$$(a\mathbf{N})(b\mathbf{N}) = a(\mathbf{N}b)\mathbf{N} = a(b\mathbf{N})\mathbf{N} = (ab)\mathbf{N} \tag{3.1.8}$$

ensures closure. Similar calculations ensure unit element, inverses and associativity, the details around those calculations are omitted for brevity.

From the definition, the resulting quotient group might seem obscure, but you may think of it almost as a modulus. If you can obtain an element of $\mathbf{G}$ from another by multiplying with an element of $\mathbf{N}$, then both elements can be regarded as equal under such an equivalence relation. This means that the quotient set is an equivalence class under this relation and every element of it is actually a set of equivalent elements. Each of these sets can then be identified with a single element: Suppose there exists a set of elements P, such that for any $g \in \mathbf{G}$ there exists unique elements $p \in P$ and $n \in \mathbf{N}$ such that $g = pn$. Then we can say that P is a quotient set, which can explain why we can think of it as a quota or modulus. Similarly we have $g = np$ when considering right cosets.

Example 3.4. We study the twelfth roots of unity. Together, these form a group R_{12} under multiplication. If we look at the fourth roots instead, we find that these form a normal subgroup R_4 to R_{12}. These form three different cosets, which can be seen as the three different colors in figure 3.2. The three cosets are the elements of R_{12}/R_4.

Lastly in this section, we introduce a relation between different groups.

Definition 3.6. *A map* $\phi : \boldsymbol{G} \mapsto \boldsymbol{H}$ *is called a* ***homomorphism*** *if*

$$\begin{gathered} g_i * g_j = g_k \\ \Rightarrow \\ \phi(g_i) * \phi(g_j) = \phi(g_k), \end{gathered} \tag{3.1.9}$$

with g_i,g_j,$g_k \in \boldsymbol{G}$*. If* ϕ *is bijective and a homomorphism it is an* ***isomorphism****, denoted as* $\boldsymbol{G} \cong \boldsymbol{H}$*.*

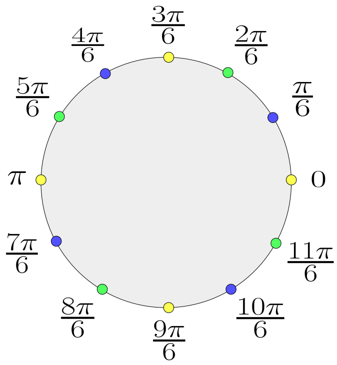


Figure 3.2: The cosets of the twelfth and fourth roots of unity.

If there exists an isomorphism between two groups, they can in practice be seen as equal, and are said to be **isomorphic**.

The expressions isomorphic and isomorphism are not restricted to groups only. As we see later, isomorphism is a concept that has significance for Lie algebra as well.

3.2 Lie Groups and Lie Algebra

After learning the basics of group theory we move on to a type of group that will be very helpful in our report. A Lie group is in short a continuous group that is also a differentiable manifold[1]. Many important Lie groups can be represented with matrices (then called **matrix Lie groups**) and so these are the focus of the next section. First, let us make a brief introduction to the more general Lie groups in this section.

Lie groups have the important property that they are generated by Lie algebras. Lie groups are manifolds with associated operators and similarly Lie algebras are vector spaces with associated bracket operators. A Lie algebra is a set of operations that represent infinitesimal transformations around the unit element of a Lie group. The Lie algebra takes the manifold of a group and expresses it as a tangent vector space around the unit element, and can in a sense be seen as a linearisation of the group. Exponentiating the elements of the Lie algebra then return elements of the Lie group representing finite transformations. However, the Lie algebra has the restriction that it only yields group elements that can be reached by infinitesimal transformations. Just like a group, there are some criteria that have to be satisfied

[1]It should also have a differentiable operator and inverse. The meaning of this is discussed briefly in Def 2.14 in Hall [5].

for an algebra to be a Lie algebra. First the effects of a Lie algebra will be discussed, followed by a strict, mathematical definition.

For clarity, let us make the following example:

Example 3.5. Let us study the Lie algebra of the Unitary group U(1). This continuous group is defined by

$$U(1) = \{e^{i\theta} \in \mathbb{C} | \theta \in \mathbb{R}\} \tag{3.2.1}$$

and might be recognized as isomorphic to S^1; that is a circle in the complex plane with a radius of 1. Note that the elements generate finite rotations around S^1.

Looking at this group around $\theta = 0$ we can expand the elements as

$$e^{i\theta} = 1 + i\theta + \mathcal{O}(\theta^2) \tag{3.2.2}$$

and find that $i\theta$ generates infinitesimal translations at the identity. From this, the Lie algebra $\mathfrak{u}(1)$ for this group can be defined

$$\mathfrak{u}(1) = \{i\theta \in \mathbb{C} | \theta \in \mathbb{R}\}. \tag{3.2.3}$$

In general a Lie group $\mathbf{G}$ and a Lie algebra $\mathfrak{g}$ can be related via the exponential equation

$$e^{tX} \in \mathbf{G} \;\Leftrightarrow\; X = \frac{d}{dt} e^{tX}|_{t=0} \in \mathfrak{g}. \tag{3.2.4}$$

With a base understanding of Lie algebra and how it works, a definition of why a Lie algebra works the way it does is now called for.

Definition 3.7. *A real or complex finite-dimensional* ***Lie algebra*** $\mathfrak{g}$ *is a real or complex finite-dimensional vector space with an associated map called the Lie bracket: [,],* $\mathfrak{g} \times \mathfrak{g} \mapsto \mathfrak{g}$*, that satisfies the following axioms:*

(i) ***Bilinearity***

$$[ax + by, z] = a[x,z] + b[y,z] \tag{3.2.5}$$

$$[z, ax + by] = a[z,x] + b[z,y] \tag{3.2.6}$$

$$x,y,z \in \mathfrak{g},\ \forall a,b \in \mathbb{R}$$

(ii) ***Anticommutativity***

$$[x,y] = -[y,x],\ \forall x,y \in \mathfrak{g} \tag{3.2.7}$$

(iii) ***Jacobi identity***

For $x,y,z \in \mathfrak{g}$

$$\big[x,[y,z]\big] + \big[y,[z,x]\big] + \big[z,[x,y]\big] = 0. \tag{3.2.8}$$

Additionally it is worth noting that from *(ii)* follows

$$[x,x] = 0, \ \forall x \in \mathfrak{g}. \tag{3.2.9}$$

In the case of when the elements of a Lie algebra can be represented with matrices, the Lie bracket is defined as the commutator, i.e $[X,Y] = XY - YX$ for two matrices X and Y. This comes in handy as many of the commonly used Lie groups and their algebras can be expressed with matrices.

You may have noticed that definition 3.7 does not mention any Lie groups. Lie algebras can be treated without connecting them to Lie groups, although traditionally that is not the case. The tangent space of a Lie group at the identity element forms a Lie algebra but proof of this is left out[2].

As with groups, there are a few useful definitions and properties to algebras, many of these mirrored for groups.

Definition 3.8. *$\mathfrak{s}$ is a* ***subalgebra*** *to an algebra $\mathfrak{g}$, if $\mathfrak{s}$ is a Lie algebra under the same bracket as $\mathfrak{g}$ and is a subspace to $\mathfrak{g}$.*

Definition 3.9. *Consider a Lie algebra $\mathfrak{g}$ and a subalgebra $\mathfrak{j}$. $\mathfrak{j}$ is called* ***ideal*** *or* ***invariant*** *if $\forall x \in \mathfrak{j}, y \in \mathfrak{g}, [x,y] \in \mathfrak{j}$.*

If for a Lie algebra there exists no non-trivial, ideal subalgebras, it is called **simple**.

Definition 3.10. *A Lie algebra is* ***abelian*** *if*

$$[x,y] = 0 \quad \forall x,y \in \mathfrak{g} \tag{3.2.10}$$

(compare this to the definition of abelian groups).

If a Lie algebra has no abelian, ideal subalgebras it is called **semi-simple**, a weaker condition compared to a simple Lie algebra. However, a semi-simple Lie algebra can be expressed as the direct sum of simple algebras.

Recalling the definition of **isomorphisms** for groups, there is a similar definition for Lie algebras. If there exists a linear, bijective map $\phi : \mathfrak{g} \mapsto \mathfrak{j}$, $x,y \in \mathfrak{g}$ it has to satisfy $\phi([x,y]) = [\phi(x)\phi(y)]$ for $\mathfrak{g}$ and $\mathfrak{j}$ to be isomorphic, once again denoted as $\mathfrak{g} \cong \mathfrak{j}$.

For some computations writing the Lie bracket as $[x,y]$ can become clumsy and arduous, so therefore there exists an alternative way to write it as

$$ad_x(y) = [x,y]. \tag{3.2.11}$$

[2]Such a proof can be found briefly described in H. Georgi's "Lie Algebras in Particle Physics" [7]

If we wanted to use the expression

$$[x,[x,[x,[x,y+z]]]] \tag{3.2.12}$$

it might be easy to at a glance miss how many brackets there are, while

$$ad_x(ad_x(ad_x(ad_x(y+z)))) = (ad_x)^4(y+z) \tag{3.2.13}$$

is more compact and easier to read. ad_x is called the **adjoint representation** or **adjoint action**.

Under the operator $ad_x(y)$ we see that all commuting elements will be zero. Other elements may not commute with each other, but can be combined with the adjoint action to become zero. So for example a, b and c may not commute but $ad_a ad_b(c)$ could be zero. From this concept we can define a type of algebra called **nilpotent**.

Definition 3.11. *A* ***nilpotent*** *Lie algebra $\mathfrak{g}$ is defined by the adjoint action as*

$$ad_{x_1} ad_{x_2} \cdots ad_{x_n}(y) = 0 \quad \forall x_i,y \ \in \mathfrak{g} \tag{3.2.14}$$

for a certain n for the algebra $\mathfrak{g}$.

3.3 Matrix Lie Groups and Algebras

As stated, many Lie groups can be represented with matrices. There are some major benefits of working with matrix Lie groups, most importantly the fact that so much of matrix algebra is familiar. When working with matrix Lie algebra, some of the abstract concepts can be expressed more explicitly.

With the group operator defined as matrix multiplication, we find that the sets must only consist of invertible matrices. This leaves what is called the **general linear group** (closure, unit element, inverse and associativity is given from elementary linear algebra) as the largest possible group. All matrix Lie groups are subgroups of the general linear group. This follows as all elements of a group need to be invertible and thus forms a subset and as the subset is a group, it is also a subgroup.

Definition 3.12. *The* ***general linear group****, GL(n) is the set of invertible $n \times n$-matrices under matrix multiplication. The group may be specified further by writing GL(n,$\mathbb{R}$) for real and GL(n,$\mathbb{C}$) for complex matrices.*

By representing groups with matrices, the Lie bracket becomes the commutator,

$$[X,Y] = XY - YX, \tag{3.3.1}$$

which fulfills the necessary axioms of definition 3.7.

For matrices, e^X can be defined by the Taylor expansion

$$\mathrm{e}^X = \sum_{k=0}^{\infty} \frac{X^k}{k!}. \tag{3.3.2}$$

Unfortunately e^X is not always easy to calculate. Using group theory however, many matrices can easily be written as exponentiated matrices, as we soon see.

The general linear group is called general, as it includes all matrices with non-zero determinants. The most intuitive matrix group to consider next is the **special linear group**, SL(n), which is a subgroup to GL(n) of matrices with determinant one. Closure is given from $\mathrm{Det}(AB) = \mathrm{Det}\, A\, \mathrm{Det}\, B$. The unit matrix has determinant one, and thus it follows that the inverse of a matrix with determinant one also has determinant one, which makes this a group. SL and its Lie algebra will be more extensively examined in the next section.

A subgroup to SL(n) is the **special orthogonal group**, SO(n), which consists of all orthogonal matrices with determinant one. These may be more commonly known as all rotational matrices. SO(n) is also a subgroup of O(n), the orthogonal matrices, which are rotational matrices with or without a reflection and thus have a determinant of ± 1.

Let us stop and study these groups for a while to showcase some of the previous definitions and make some new.

The elements of SO(2) are traditionally written on the form

$$A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \theta \in \mathbb{R}. \tag{3.3.3}$$

It is thus a continuous group. Every element can be uniquely determined with one continuous parameter, making it a **one parameter group**[3]. This implies that the Lie algebra is one-dimensional and spanned by one element. This element is given from (3.2.4), as

$$T = \frac{d}{d\theta} e^{\theta T}|_{\theta=0} = \frac{d}{d\theta} A(\theta)|_{\theta=0} = \begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}\Bigg|_{\theta=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.3.4}$$

The elements of O(2) can be written as either

$$A_+(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{3.3.5}$$

[3]The definition of a one-parameter group also requires $A(\alpha)A(\beta) = A(\alpha+\beta)$ which is more of a constraint on the parametrization than on the group.

or

$$A_-(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{3.3.6}$$

The second set of matrices is not however "close" to the unit element, no matter which θ is chosen. As the Lie algebra represents infinitesimal transformations at the unit element, the Lie algebra is given from A_+ and thus the same as for SO(2).

The first set, A_+, has determinant one and the second, A_- has determinant minus one. The determinant is continuous under continuous transformations, so the two sets must be disconnected. If a group is made from n disconnected sets, we say that it has n **components**. The Lie algebra is therefore always given from only one of these components, the component with the unit element.

There are a few equivalent properties of orthogonal matrices, with different implications. For instance we have that $AA^\top$ is the identity matrix. Another is that the bilinear form $< x,y >= x_1y_1 + x_2y_2 + \ldots + x_ny_n$ is preserved, that is: $< Ax,Ay >=< x,y >$. The second property is of particular interest. This is a preserved form in Euclidean space. If we recall the metric in Minkowski space, where the preserved form had alternating signs, this intuitively leads to defining new groups; the **generalized orthogonal groups** and their special equivalents. These are defined as the sets of matrices $B \in GL(m+k)$ that preserves the bilinear form $<>_{m+k}$ with m positive terms and k negative, called O(m,k) and SO(m,k). These bilinear forms can be written as

$$< x,y >_{m+k}= x^\top gy, \tag{3.3.7}$$

where g is a signature matrix; diagonal with the first m entries $= 1$ and the next k entries $= -1$. Specifically O(3,1) is also known as the Lorentz group, having the same signature as spacetime in special relativity, preserving the same form and with Lorentz-transformations as elements.

Example 3.6. SO(2) and SO(1,1): The elements of SO(2) can be written as in equation 3.3.3.

For SO(1,1) we look for matrices B that preserve the scalar product with signature "$+-$". Such matrices can be found by the same process as for ordinary orthogonal matrices:

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{3.3.8}$$

Generally for (S)O we have

$$
\begin{aligned}
(Bx)^\top gBy &= x^\top gy \\
&\Rightarrow \\
B^\top gB - g &= 0
\end{aligned}
\tag{3.3.9}
$$

where g is the signature matrix. In our case we have

$$
\begin{aligned}
\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &= 0 \\
&\Rightarrow \\
\begin{pmatrix} a^2 - c^2 & ab - cd \\ ab - cd & b^2 - d^2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &= 0 \\
&\Rightarrow \\
&\begin{cases} a^2 - c^2 = 1 \\ d^2 - b^2 = 1 \\ ab - cd = 0 \end{cases} .
\end{aligned}
\tag{3.3.10}
$$

Using the hyperbolic identity the first two yield that $a = \pm\cosh\theta$, $b = \sinh\varphi$, $c = \sinh\theta$ and $d = \pm\cosh\varphi$ with $\theta, \varphi \in \mathbb{R}$. Observe the $\pm$ in front of cosh, as it otherwise only takes positive values. If all entries are to be real, a and d cannot possibly have absolute values less than zero, thus we need no special case for $a \in (-1,1)$.

The last equation is equivalent to

$$
\begin{aligned}
\sinh(\theta \pm \varphi) &= 0 \\
&\Rightarrow \\
\theta &= \pm\varphi,
\end{aligned}
\tag{3.3.11}
$$

where we have a plus sign if a and d are of different signs. From this we find four

Table 3.1: Multiplicative relations between components of O(1,1).

B_*B_*	$+$	$-$	$\uparrow$	$\downarrow$
$+$	$+$	$-$	$\uparrow$	$\downarrow$
$-$	$-$	$+$	$\downarrow$	$\uparrow$
$\uparrow$	$\uparrow$	$\downarrow$	$+$	$-$
$\downarrow$	$\downarrow$	$\uparrow$	$-$	$+$

different cases,

$$\begin{aligned} B_+ &= \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} \\ B_- &= \begin{pmatrix} -\cosh\theta & \sinh\theta \\ \sinh\theta & -\cosh\theta \end{pmatrix} \\ B_\uparrow &= \begin{pmatrix} -\cosh\theta & -\sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} \\ B_\downarrow &= \begin{pmatrix} \cosh\theta & \sinh\theta \\ -\sinh\theta & -\cosh\theta \end{pmatrix}. \end{aligned} \tag{3.3.12}$$

These are disconnected and thus the group has four different components.

These sets of matrices fulfill the multiplicative relations of table 3.1. In this table we read out that an element from $B_\uparrow$ multiplied by an element from B_- ends up in $B_\downarrow$ and so on. Let us validate that these indeed make groups. We see closure and that the unit element lies in B^+. Matrix multiplication is associative, so left to do is to find the inverses. The latter two sets of matrices ($\uparrow\downarrow$) can be noted to have determinant minus one, meaning that they are not in SO(1,1) but rather O(1,1). The inverses of all these sets are given as

$$B_+^{-1}(\theta) = B_+^\tau(\theta) = \begin{pmatrix} \cosh\theta & -\sinh\theta \\ -\sinh\theta & \cosh\theta \end{pmatrix} = B_+(-\theta) \tag{3.3.13}$$

$$B_-^{-1}(\theta) = B_-^\tau(\theta) = \begin{pmatrix} -\cosh\theta & -\sinh\theta \\ -\sinh\theta & -\cosh\theta \end{pmatrix} = B_-(-\theta) \tag{3.3.14}$$

$$B_\uparrow^{-1}(\theta) = B_\uparrow^\tau(\theta) = \begin{pmatrix} -\cosh\theta & -\sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} = B_\uparrow(\theta) \tag{3.3.15}$$

$$B_{\downarrow}^{-1}(\theta) = B_{\downarrow}^{\mathcal{T}}(\theta) = \begin{pmatrix} \cosh\theta & \sinh\theta \\ -\sinh\theta & -\cosh\theta \end{pmatrix} = B_{\downarrow}(\theta). \tag{3.3.16}$$

We note that the inverses of a set lie in the same set. Thus B_+ and $B_+ \cup B_- = SO(1{,}1)$ form two subgroups of SL(2,ℝ) and $B_+ \cup B_- \cup B_\uparrow \cup B_\downarrow = O(1{,}1)$ form a subgroup of GL.

There is an important difference between O(2) and O(1,1). The entries of O(2), $\sin\theta$ and $\cos\theta$, are bounded, whereas the entries of O(1,1), $\sinh\theta$ and $\cosh\theta$, are unbounded.

Definition 3.13. *A group that is bounded, meaning that there exists a supremum and infimum that any matrix element of any matrix in the group lies between, is called* ***compact***[4]*. If a subgroup is the largest possible compact subgroup, it is called* ***maximally compact****.*

We defined the generalized orthogonal group from preserved bilinear forms, which is equivalent to having the transpose as inverse. Using that for O we have $B^\top gB = g$ (done generally in Example 3.6) we get

$$gB^\top gB = gg = I, \tag{3.3.17}$$

and a way to define a generalized transpose.

Definition 3.14. *The* ***generalized transpose*** *is given as*

$$B^{\mathcal{T}} = gB^\top g \tag{3.3.18}$$

where g is the signature matrix.

The generalized transpose has the neat property that it works much like the ordinary transpose unless you look at specific elements of a matrix. For instance we have

$$(AB)^{\mathcal{T}} = g(AB)^\top g = gB^\top A^\top g = gB^\top ggA^\top g = B^{\mathcal{T}}A^{\mathcal{T}}, \tag{3.3.19}$$

which is important for many calculations.

The complex version of orthogonal matrices are the unitary matrices U(n), with $A^*A = I$, where $A^*$is the conjugate transpose. The previous concepts also apply to these and we will for instance see groups such as SU(2,1) which are unitary matrices under signature "$++-$" and with determinant one.

[4]It is also required that any sequence of matrices in the group that converges, converges to an element in the group

3.4 Important Concepts of Lie Groups and Lie Algebra

After going through the basic idea of group theory, with a focus on Lie groups and algebras, there are certain concepts that can be of use to explore further. The special linear group SL(2,$\mathbb{R}$) plays a central role in group theory, as it can among other things be used to generate other groups. Therefore we will delve into SL(2,$\mathbb{R}$), as well as the more general group SL(n,$\mathbb{R}$). Additionally, there are certain operations on Lie groups and their algebras that can be used during the calculations presented herein. The **Iwasawa Decomposition** is used to split a Lie algebra into subalgebras, while the **Maurer-Cartan form** presents an alternative way to obtain a Lie algebra element from its respective Lie Group. Furthermore the **Killing Form** is presented as a way to describe a certain relationship between the elements of a Lie algebra. With these additional concepts we can acquire a broader spectrum of tools to make use of group theory.

3.4.1 SL(2,$\mathbb{R}$) and SL(n,$\mathbb{R}$)

To understand the group SL(2,$\mathbb{R}$) and its uses, we begin with the definition of its elements.

Definition 3.15. *The Lie group* ***SL***(2,$\mathbb{R}$) *is comprised of all the real* 2×2 *matrices with determinant one. Accordingly, for* $S \in SL(2,\mathbb{R})$

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad ad - bc = 1 \qquad a,b,c,d \in \mathbb{R}. \tag{3.4.1}$$

Groups of the type SL(n,$\mathbb{R}$) and their algebras have dimension $2n-1$, meaning that all elements of the group can be generated by $2n-1$ matrices and the vector space of the algebra can be spanned by $2n-1$ elements. Consequently, this means that for SL(2,$\mathbb{R}$) three matrices are needed to generate the group.

The Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ of the group SL(2,$\mathbb{R}$) is a simple and real algebra. As the elements in SL(2,$\mathbb{R}$) can be generated by three matrices with determinant one, the Lie algebra must be spanned by a set of three traceless matrices, to preserve the determinant of the group elements. These could be found from the matrix S

through differentiating by a, b and c. This gives three elements

$$h = \frac{d}{da} S|_{a,b,c=0} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$e = \frac{d}{db} S|_{a,b,c=0} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
$$f = \frac{d}{dc} S|_{a,b,c=0} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.4.2}$$

that span the algebra by the direct sum of the real vectorspaces $\mathbb{R}h$,$\mathbb{R}e$ and $\mathbb{R}f$ as follows,

$$\mathfrak{sl}(2,\mathbb{R}) = \mathbb{R}f \oplus \mathbb{R}h \oplus \mathbb{R}e. \tag{3.4.3}$$

However, an alternative way to arrive at these elements will be shown here, as it is of relevance for understanding how the groups SU(2), SL(2,$\mathbb{C}$) and SL(2,$\mathbb{R}$) relate. A visual representation of this process is shown in figure 3.3. These generators present the specific set h, e and f that can be used to further compose other groups, like SL(n,$\mathbb{R}$).

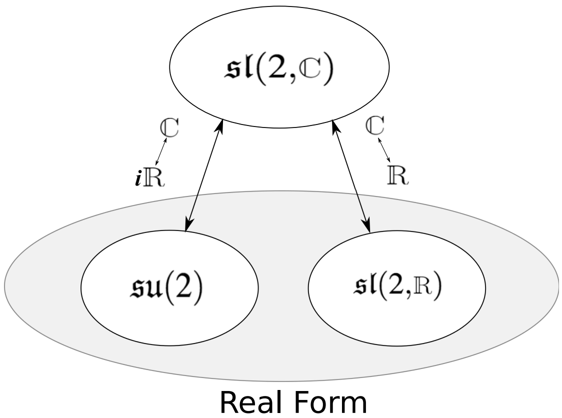


Figure 3.3: The two groups SU(2) and SL(2,$\mathbb{R}$) are real forms of the group SL(2,$\mathbb{C}$). Using the Lie algebra of $\mathfrak{su}(2)$ the algebra $\mathfrak{sl}(2,\mathbb{R})$ can be constructed via $\mathfrak{sl}(2,\mathbb{C})$.

To find a way to construct generators of $\mathfrak{sl}(2,\mathbb{R})$ that can express an infinitesimal transformation on the complex upper half-plane, we begin with the generators of $\mathfrak{su}(2)$. The group to this algebra is SU(2), representing all unitary 2×2 matrices with determinant one. These generators are in some ways easier to understand and construct, and from them the algebra $\mathfrak{sl}(2,\mathbb{R})$ can be found. The two groups SU(2) and SL(2,$\mathbb{R}$) are connected by the group SL(2,$\mathbb{C}$) as they are both real forms of this group, with their respective Lie algebras $\mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{su}(2)$ as real forms of $\mathfrak{sl}(2,\mathbb{C})$.

Finding the elements of the algebra, we begin with the group SU(2) and observe that its elements can additionally be seen as rotations in three dimensions. The elements in SU(2) can therefore be written as

$$\forall \, \mathrm{U} \in \mathrm{SU}(2), \quad a_i \in \mathbb{R} \;\Rightarrow\; \mathrm{U} = e^{ia_x\sigma_x + ia_y\sigma_y + ia_z\sigma_z}, \tag{3.4.4}$$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.4.5}$$

To obtain the elements of the Lie algebra $\mathfrak{su}(2)$, the principle of equation (3.2.4) can be used. This results in elements made up of direct sums of the vector spaces comprised $i\mathbb{R}$, times a Pauli matrix σ_i. The span of these elements can be written as

$$\mathfrak{su}(2) = \mathrm{span}_{\mathbb{R}}\{i\sigma_x, i\sigma_y, i\sigma_z\} = \mathbb{R}i\sigma_x \oplus \mathbb{R}i\sigma_y \oplus \mathbb{R}i\sigma_z. \tag{3.4.6}$$

From this span we take the next step to find the generators of $\mathfrak{sl}(2,\mathbb{C})$. This is found by expanding $\mathfrak{su}(2)$ to complex, and not only strictly imaginary combinations of the Pauli matrices. To make the next step slightly smoother, new generators are defined as

$$T_x = \frac{1}{2}i\sigma_x \quad T_y = \frac{1}{2}i\sigma_y \quad T_z = \frac{1}{2}i\sigma_z, \tag{3.4.7}$$

where every element in $\mathfrak{sl}(2,\mathbb{C})$ can now be constructed as a complex linear combination of these generators, written as

$$\mathfrak{sl}(2,\mathbb{C}) = \mathrm{span}_{\mathbb{C}}\{T_x, T_y, T_z\} = \mathbb{C}T_x \oplus \mathbb{C}T_y \oplus \mathbb{C}T_z. \tag{3.4.8}$$

This takes us one step closer to defining $\mathfrak{sl}(2,\mathbb{R})$. At this point, the generators e,f and h are defined from T_x,T_y,T_z by the relations

$$\begin{aligned} &e = T_2 - iT_1, \quad f = -(T_2 + iT_1), \quad h = -2iT_3 \\ &e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{3.4.9}$$

Redefining the generators like this may seem arbitrary, but it has the effect of creating two triangular matrices, and one diagonal matrix. This proves to be important when constructing other Lie algebras, as will be shown later. We can also recognize these as the same matrices that resulted from differentiating the matrix S in respect to a, b and c, as in equation (3.4.2). Since these are only linear combinations of the previous generators they also span $\mathfrak{sl}(2,\mathbb{C})$, which can be written as

$$\mathfrak{sl}(2,\mathbb{C}) = \mathbb{C}f \oplus \mathbb{C}h \oplus \mathbb{C}e. \tag{3.4.10}$$

Finally, to arrive at $\mathfrak{sl}(2,\mathbb{R})$ we replace every $\mathbb{C}$ with a $\mathbb{R}$, giving us a completely real Lie algebra, spanned as in equation (3.4.3), where the real combinations of h will span all the diagonal matrices. This abelian subalgebra can also be referred to as **Cartan subalgebra**. Commuting the generators e, f and h gives the following commutation relations

$$[e,f] = h \qquad [h,e] = 2e \qquad [h,f] = -2f. \tag{3.4.11}$$

As e is upper triangular, f lower triangular and h is diagonal, this way of constructing the span is referred to as a **triangular decomposition**.

The triangular decomposition of the group SL(2,$\mathbb{R}$) into upper triangular matrices, lower triangular and diagonal matrices is a trait that holds for $\mathfrak{sl}(n,\mathbb{R})$ for all n. The span for this algebra can similarly be written as

$$\mathfrak{sl}(n,\mathbb{R}) = \mathbb{R}\mathfrak{n}_- \oplus \mathbb{R}\mathfrak{h} \oplus \mathbb{R}\mathfrak{n}_+ \tag{3.4.12}$$

where the elements of $\mathfrak{n}_-$ are lower triangular matrices and the elements of $\mathfrak{n}_+$ are upper triangular matrices. Lastly $\mathfrak{h}$ has diagonal matrices as elements that represent the Cartan subalgebra spanned by the diagonal elements h_i, written as

$$\mathfrak{h} = \sum_i^{n-1} \mathbb{R}h_i. \tag{3.4.13}$$

The Lie algebra $\mathfrak{sl}(n,\mathbb{R})$ is spanned by $n-1$ triples of the generators (h_i,e_i,f_i) for $i = 1,...,n-1$, called the **Chevalley generators**. Every triple of these generators span $\mathfrak{sl}(2,\mathbb{R})_i$ and therefore satisfy the same criteria as the generators in the equation (3.4.11). $\mathfrak{sl}(n,\mathbb{R})$ is then constructed by taking all the algebras $\mathfrak{sl}(2,\mathbb{R})_i$ as a direct sum

$$\oplus_i^n \, \mathfrak{sl}(2,\mathbb{R})_i \tag{3.4.14}$$

and additionally linking them together by using the following rules, called the **Chevalley relations**,

$$[e_i,f_j] = \delta_{ij}h_j \tag{3.4.15}$$
$$[h_i,e_j] = A_{ij}e_j \tag{3.4.16}$$
$$[h_i,f_j] = -A_{ij}f_j \tag{3.4.17}$$
$$[h_i,h_j] = 0 \tag{3.4.18}$$
$$A = \begin{pmatrix} 2 & -1 & 0 & & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \vdots \\ \vdots & & & \ddots & & \vdots \\ 0 & & \cdots & & -1 & 2 \\ & & & & & \end{pmatrix}. \tag{3.4.19}$$

The matrix A is the **Cartan matrix**, and the diagonal values represent the generators from the same $\mathfrak{sl}(2,\mathbb{R})_i$ acting upon each other according to (3.4.11) . The values of -1 next to the diagonal show us that only elements from subsequent $\mathfrak{sl}(2,\mathbb{R})_i$ have a connection that results in something other than zero. All other combinations with the matrix A in (3.4.16), (3.4.17) will result in zero. However we have so far put no other constraints on the generators e_i with other e_j, nor f_i with f_j. Without further rules the generators e_i and f_i could, by commutation as follows

$$[e_i,e_j],\ [e_i,[e_j,e_k]], ... \quad [f_i,f_j],\ [f_i,[f_j,f_k]], ..., \tag{3.4.20}$$

generate an infinite amount of new elements. To see to it that the $n-1$ triples generate a finite Lie algebra, the following **Serre relations** are added

$$(ad_{e_i})^{1-A_{ij}}(e_j) = 0, \quad (ad_{f_i})^{1-A_{ij}}(f_j) = 0, \quad \forall i \neq j. \tag{3.4.21}$$

Finally, with these generators, relations and restraints we now arrive at all the generators for $\mathfrak{sl}(n,\mathbb{R})$.

Example 3.7. As an example of this way of building $\mathfrak{sl}(n,\mathbb{R})$ from $\mathfrak{sl}(2,\mathbb{R})_i$ we can look at $\mathfrak{sl}(3,\mathbb{R})$. Here the number of generators should in total be $2n-1=8$. We begin with $n-1$ triples, resulting in the pair (e_1,f_1,h_1) and (e_2,f_2,h_2) that produce the first six elements needed to span the Lie algebra. The remaining two elements can be found from the matrix A

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \tag{3.4.22}$$

and $1 - A_{12} = 1 - A_{21} = 2$. From equation (3.4.21) we find that $[e_1,[e_1,e_2]] = [f_1,[f_1,f_2]] = 0$. This gives us the only two possible new elements spanning the algebra $[e_1,e_2] = e_3$ and $[f_1,f_2] = f_3$, adding up to all eight elements needed for $\mathfrak{sl}(3,\mathbb{R})$.

It is worth noting that this way of constructing $\mathfrak{sl}(n,\mathbb{R})$ can be used when constructing other Lie algebras. By using $\mathfrak{sl}(2,\mathbb{R})_i$ with other values for A and the same Chevalley relations, another Lie algebra can be constructed. For $\mathfrak{sl}(n,\mathbb{R})$, specifically, A will be as in equation (3.4.19).

3.4.2 Iwasawa Decomposition

It may sometimes prove necessary to split a group into several different subgroups. One such decomposition is the **Iwasawa decomposition**. This decomposition makes the claim that every semi-simple Lie group **G** can be decomposed into three parts, each part having a specific requirement

$$\mathbf{G}{=}\mathbf{NAK}, \tag{3.4.23}$$

where **N, A** and **K** are Lie subgroups to **G**, with the respective Lie algebras $\mathfrak{n}$, $\mathfrak{a}$ and $\mathfrak{k}$. **K** will be the maximal compact group of **G**, as defined for matrix groups in definition 3.13. For **N** the matrix representation of the elements must be upper triangular matrices with determinant one, while for the subgroup **A**, the matrices are all diagonal matrices.

The proof of this claim will not be presented herein for a general group, but the Iwasawa decomposition for the Lie group SL(2,$\mathbb{R}$) will be shown in detail. As SL(2,$\mathbb{R}$) is a simple Lie algebra, and thereby also a semi-simple Lie algebra, there exists an Iwasawa decomposition where the maximal compact group **K** for SL(2,$\mathbb{R}$) is SO(2). This means that SL(2,$\mathbb{R}$) can be separated into the parts

$$\mathrm{SL}(2,\mathbb{R}) = \mathbf{NA}\,\mathrm{SO}(2) = \left\{\begin{pmatrix}1 & * \\ 0 & 1\end{pmatrix}\right\}\left\{\begin{pmatrix}* & 0 \\ 0 & 1/*\end{pmatrix}\right\}\mathrm{SO}(2) \tag{3.4.24}$$

where the matrices preserve the property of SL(2,$\mathbb{R}$) that the determinant is equal to one, and the diagonal of the matrices in **A** is restricted to only positive values. Now we have one subgroup to SL(2,$\mathbb{R}$) of elements that are represented by upper triangular matrices and one subgroup of diagonal matrices, with algebras that likewise are represented by upper triangular and traceless matrices respectively. To compare this to the definition of an element S of the group SL(2,$\mathbb{R}$), we create an element through the Iwasawa decomposition as follows, and relate (r,x,θ) to (a,b,c,d)

$$
\begin{aligned}
S &= \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\
&= \begin{pmatrix} r\cos\theta + x/r\sin\theta & x/r\cos\theta - r\sin\theta \\ 1/r\sin\theta & 1/r\cos\theta \end{pmatrix} = \begin{pmatrix} a & b \\ c & \frac{1+bc}{a} \end{pmatrix} \\
&\Rightarrow 1/r^2 = c^2 + d^2, \qquad x = \frac{a-d}{c} \qquad \cos\theta = \sqrt{\frac{1}{c^2/d^2+1}}
\end{aligned} \tag{3.4.25}
$$

$$
\mathfrak{sl}(2,\mathbb{R}) = \mathfrak{k} \oplus \mathfrak{h} \oplus \mathfrak{n} \tag{3.4.26}
$$

where $\mathfrak{k}$ represents the maximal compact subalgebra to the Lie algebra in question, which means $\mathfrak{k} = \mathfrak{so}(2)$, $\mathfrak{h}$ is all the non-compact elements of the Cartan subalgebra and lastly $\mathfrak{n}$ is a nilpotent subalgebra as defined in definition 3.11.

3.4.3 Maurer-Cartan Form

In section 3.2 it is shown how to find an element in a Lie algebra from an element of its respective Lie group with the equation (3.2.4). However, the Maurer-Cartan form presents an alternative way of finding the Lie algebra to a group. For an element U in a group $\mathbf{G}$, an element on the Cartan-Form of the Lie algebra $\mathfrak{g}$ can be found as

$$
U^{-1}\partial_t U. \tag{3.4.27}
$$

We can see that this yields the Lie algebra element $\partial_\mu X \in \mathfrak{g}$

$$
\begin{aligned}
U &= e^X \\
\Rightarrow\ U^{-1}\partial_\mu U &= e^{-X}e^X\partial_\mu X = \partial_\mu X,
\end{aligned} \tag{3.4.28}
$$

which normally is not the same as the Lie algebra element X, as found in equation (3.2.4).

3.4.4 Killing Form

For Lie algebras it would be beneficial to find something like a scalar product. The Killing form is a product between elements in an algebra that shares some similarities with a scalar product and help to compare elements.

Definition 3.16. ***Killing Form*** *is the product between the adjoint representations of* $a,b \in \mathfrak{g}$ *defined by*

$$K(a,b) = \mathrm{Tr}(ad_a ad_b) \tag{3.4.29}$$

with ad_a *as in equation* (3.2.11).

This product has the benefit of being bilinear and symmetric, that is

$$\begin{aligned} K(a,b) &= K(b,a) \\ K(a,b+c) &= K(a,b) + K(a,c). \end{aligned} \tag{3.4.30}$$

So far, the Killing form is a lot like a scalar product. Although, just like the metric described in section 2.1.1, it is an object that is similar to a scalar product, but is not necessarily positive definite.

Example 3.8. As an example of this, we examine the group SL(2,ℝ) and how the Killing form acts on the elements of its Lie algebra. Firstly we need to calculate what ad_e, ad_h and ad_f are. For $\mathfrak{sl}(2,\mathbb{R})$ any element will here be written as $v = v_1 e + v_2 h + v_3 f = (v_1,v_2,v_3)$, $\forall\, v \in \mathfrak{sl}(2,\mathbb{R})$, $v_i \in \mathbb{R}$. From their commutation relations in equation (3.4.11) we get

$$v = (v_1,v_2,v_3) \qquad u = (u_1,u_2,u_3)$$
$$[e,f] = h,\ \ [e,h] = -2e \qquad [h,e] = 2e,\ \ [h,f] = -2f \qquad [f,h] = 2f,\ \ [f,e] = -h$$
$$ad_e = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad ad_h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad ad_f = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$
$$\Rightarrow$$
$$\begin{aligned} K(e,e) &= \mathrm{Tr}(ad_e ad_e) = 0 \qquad & K(h,h) &= \mathrm{Tr}(ad_h ad_h) = 8 \\ K(f,f) &= \mathrm{Tr}(ad_f ad_f) = 0 \qquad & K(e,f) &= \mathrm{Tr}(ad_e ad_f) = 4 \\ K(h,e) &= \mathrm{Tr}(ad_h ad_e) = 0 \qquad & K(f,h) &= \mathrm{Tr}(ad_f ad_f) = 0 \end{aligned}$$
$$\Rightarrow$$
$$K(u,v) = \mathrm{Tr}(ad_u ad_v) = u^\top K v, \tag{3.4.31}$$

where K is the following matrix

$$K = \begin{pmatrix} K(e,e) & K(e,h) & K(e,f) \\ K(e,h) & K(h,h) & K(h,f) \\ K(e,f) & K(h,f) & K(h,f) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}. \tag{3.4.32}$$

From here we can see that depending on u and v, $K(u,v)$ can be either positive or negative.

As seen above, calculating the Killing form explicitly is not always quick and smooth. However, the Killing form has several handy traits that can be used during implicit calculation. Such a trait is that the Killing form is invariant during cyclic permutation, meaning that

$$K(a,[b,c]) = K([a,b],c). \tag{3.4.33}$$

This can be shown by firstly expressing the following relation

$$ad_{[ab]} = [ad_a,ad_b], \tag{3.4.34}$$

that can be proven by

$$\begin{aligned} ad_{[ab]}(c) &= \big[[a,b],c\big] \stackrel{?}{=} [ad_a,ad_b](c) = \big[a,[b,c]\big] - \big[b,[a,c]\big] \\ &\Leftrightarrow \\ 0 &\stackrel{?}{=} \big[a,[b,c]\big] + \big[b,[c,a]\big] + \big[c,[a,b]\big]. \end{aligned} \tag{3.4.35}$$

This means that the equivalency at $\stackrel{?}{=}$ holds true, as equation (3.4.35) is the third axiom, the Jacobi identity, in the definition 3.7 of a Lie algebra.

Another important step can also be shown with the Jacobi identity

$$\begin{aligned} \big[[a,b],c\big] &= -\Big(\big[b,[c,a]\big] + \big[c,[a,b]\big]\Big) \\ &= -(bca - bac - cab + acb + cab - cba - abc + bac) \\ &= abc - acb + cba - bca = a[b,c] - [b,c]a \\ &= \big[a,[b,c]\big]. \end{aligned} \tag{3.4.36}$$

These can now be used to show that

$$\begin{aligned} K([a,b],c) &= \mathrm{Tr}(ad_{[a,b]}ad_c) = \mathrm{Tr}\left([ad_a,ad_b]ad_c\right) \\ &= \mathrm{Tr}\left(ad_a[ad_b,ad_c]\right) = \mathrm{Tr}\left(\big[a,[b,c]\big]\right) && (3.4.37) \\ &= K(a,[b,c]). && (3.4.38) \end{aligned}$$

Expressing the Killing form with different representations of the elements may result in a scalar difference,

$$K(a,b) = C\,\mathrm{Tr}(ab), \tag{3.4.39}$$

where the constant C depends on the algebra and representation. An example of this constant for some common algebras and their fundamental representations can be seen in table 3.2. From this we find that for $\mathfrak{sl}(2,\mathbb{R})$ the Killing form can

Table 3.2: The table above presents three Lie algebras and how their Killing forms are expressed for their fundamental representation.

Algebra	C
$\mathfrak{sl}(n,\mathbb{R})$	$2n$
$\mathfrak{su}(n)$	$2n$
$\mathfrak{so}(n)$	$n-2$

also be written as $K(a,b) = 4\mathrm{Tr}(a,b)$. For $\mathfrak{sl}(n,\mathbb{R})$ the fundamental representation is comprised of elements of $n \times n$-matrices, which for $\mathfrak{sl}(2,\mathbb{R})$ is spanned by real combinations of the matrices e,f and h in (3.4.2). Comparing this to what was previously calculated in equation (3.4.32) we can see that C must indeed be 4, as

$$\begin{aligned} &4\,\mathrm{Tr}(ee) = 4\,\mathrm{Tr}(eh) = 4\,\mathrm{Tr}(hf) = 4\,\mathrm{Tr}(ff) = 4 \cdot 0 = 0 \\ &4\,\mathrm{Tr}(hh) = 4(1+1) = 8 \qquad 4\,\mathrm{Tr}(ef) = 4(1+0) = 4. \end{aligned} \tag{3.4.40}$$

4

Generating Black Hole Solutions From Group Theory

In this chapter we arrive at the Schwarzschild solution by a different method than in section 2.5.1. This different approach requires a bit more work, but the reward is that from this method we can find symmetries of the solution which can be hidden and very hard to see using the previous method. These symmetries can, with group theory, be used to generate new solutions with similar symmetries, effectively expanding a solution into a family of solutions. The method used in this chapter can also be used to find other families of solutions, in principle by using the same steps presented here. We go through these steps in detail for the Schwarzschild solution, and the Reissner-Nordström solution is found mostly by using what we did for Schwarzschild. The main line of this chapter will follow two articles of Breitenlohner et al. More information can be found in [12] and [13].

To do this we start by performing a dimensional reduction of a dimension that the solution is symmetric with respect to. This will give a new effective action in three dimensions originating from the original action in four dimensions. The dimensional reduction and the symmetries that manifest afterwards will be discussed in section 4.1. By use of group theory and the action principle we find the equations of motion for the three-dimensional action in section 4.2. These equations are solved in section 4.3 to find the Schwarzschild solution. Thereafter the Reissner-Nordström solution in section 4.4 will be derived using the same method as the Schwarzschild solution. Lastly, how to generate new solutions in the same solution family is discussed in section 4.5, along with some generalizations of the concepts and further uses of the method.

Throughout this whole chapter indices in four dimensions will be written as $\hat{\mu}$, $\hat{\nu}$,... and indices in three dimensions will be μ, ν,...

4.1 Dimensional Reduction

In four-dimensional spacetime the action from gravity in vacuum is given by

$$S_4[g_{\hat{\mu}\hat{\nu}}] = \int_{\mathbb{R}^4} \mathrm{d}^4x\sqrt{g}R, \tag{4.1.1}$$

where we have omitted a factor $\frac{1}{8\pi G}$ from the expression presented in section 2.4.3. Since the action in vacuum only has one term, this factor does not effect the stationary point of the action.

Many interesting solutions have symmetries so that the metric is independent of one or more dimensions[1]. Such a dimension is assumed to be compact, so that it can be seen as to parametrize a circle. If it in addition is small, the metric ansatz will only contain zero-order terms[2]. The ansatz for the four-dimensional metric can then be written be on the form

$$g_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} \mathrm{e}^{-\phi}g_{\mu\nu} + \mathrm{e}^{\phi}A_\mu A_\nu & \mathrm{e}^{\phi}A_\mu \\ \mathrm{e}^{\phi}A_\nu & \mathrm{e}^{\phi} \end{pmatrix}, \tag{4.1.2}$$

where $g_{\mu\nu}$ is a three-dimensional metric, A_μ is a vector field and ϕ is a scalar field. These three objects, $g_{\mu\nu}$, A_μ and ϕ are functions of the remaining dimensions. So for a timelike reduction, the three-dimensional metric is a function of the space coordinates. The parameterization of the ansatz in (4.1.2) is used because it turns out to be convenient during calculations. By convention, the coordinate reduced is placed to be the last index.

By integrating the action $S_4[g_{\hat{\mu}\hat{\nu}}]$ with respect to the reduced dimension, as well as calculating the three-dimensional Ricci scalar R and the determinant of the metric g, an effective three-dimensional action can be derived. This action contains the vector field A_μ, which can be simplified by dualizing to a scalar field χ:

$$\partial^\mu\chi = \epsilon^{\mu\nu\sigma}\partial_{[\nu}A_{\sigma]} \tag{4.1.3}$$

Dualization is a property that all Maxwell and gravitational theories share, and is linked to the Bianchi identity, see equation (2.4.6). By dualizing a field, both the original field and the dual fields will contain the same information. It should be noted that in the general case, the dual to a tensor could have a higher as well

[1] A mathematical description of symmetries allowing for dimensional reduction would be that there exist spacelike and timelike Killing vectors [12], objects which in themselves we do not treat in this text. Killing vectors are described in Weinberg [2]

[2] In fact, in the real world, the compactified dimension is not a microscopic; all four dimensions are indeed macroscopic and observable. Nevertheless, the method of dimensional reduction can be used to obtain solutions to Einsteins equations, even though the three-dimensional action is in some sense unphysical.

as lower rank. Which rank that is obtained, depends on the rank of the dualized tensor and the dimension, since the rank of the Levi-Civita tensor is equal to the number of dimensions. [14]

The processes of dimensional reduction and dualization are rather complicated. Therefore, the action will simply be stated without proof. More information about dimensional reduction can be found in Pope [15]. The action, when reduced from four to three dimensions, becomes

$$\int_{\mathbb{R}^4} \mathrm{d}^4 x \sqrt{g} R \tag{4.1.4}$$

$$\Downarrow$$

$$\int_{\mathbb{R}^3} \mathrm{d}^3 x \sqrt{g} R - \int_{\mathbb{R}^3} \mathrm{d}^3 x \sqrt{g} \mathcal{L}(\chi,\phi). \tag{4.1.5}$$

In the expression for the three-dimensional action the first term will be referred to as S_3 and the second as $S_{\mathbf{G}/\mathbf{K}}$. $\mathcal{L}(\chi,\phi)$ is the sigma model and is given by[3]

$$\mathcal{L}(\chi,\phi) = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi \pm \mathrm{e}^{2\phi} \partial_\mu \chi \partial^\mu \chi). \tag{4.1.6}$$

The plus sign in $\mathcal{L}$ is obtained when the reduction is done with respect to a space coordinate and the minus sign when it is done with respect to a time coordinate. As the Schwarzschild solution is both static and spherically symmetric, either a time coordinate or an angular coordinate can be reduced. The mathematical formalism will here be more thorough when reducing a spatial coordinate, but in order to later obtain the Schwarzschild solution it will be easier if the time coordinate is reduced. Since both spatial and timelike reduction are important for further use, both will be presented.

4.1.1 Invariance of $\mathcal{L}(\chi,\phi)$ Under SL(2,$\mathbb{R}$)

The aim of this section is to show that the sigma model is invariant under SL(2,$\mathbb{R}$). Using this invariance, it is possible to generate families of solutions from a single solution. This will be discussed further in section 4.5. Another important result from this section is that χ and ϕ parametrize SL(2,$\mathbb{R}$)/$\mathbf{K}$, where $\mathbf{K}$ depends on the type of reduction; for spacelike reduction the coset space will be SL(2,$\mathbb{R}$)/ SO(2) and for timelike it will be SL(2,$\mathbb{R}$)/ SO(1,1). The fact that χ and ϕ parametrize SL(2,$\mathbb{R}$)/$\mathbf{K}$ enables us to express the Lagrangian in a convenient way, to be used when deriving the Schwarzschild solution.

[3]See eq. (1.80) in Pope [15]. A discussion on signs depending on the reduced coordinate can be found in Persson [16].

We show this invariance by creating an object that is invariant and show that this is precisely our Lagrangian. To do this we use the Killing form, defined in section 3.4.4. The calculations for a timelike and a spacelike reduction are almost identical. The only difference will be that different generalized transposes are used, the choice of which depends on the group **K**, see section 3.3 for further disussion on generalized transpose. It will be pointed out where the differences arise.

From the Iwasawa decomposition, found in 3.4.2, we know that any element $g \in \mathrm{SL}(2,\mathbb{R})$ can be uniquely factorized into

$$g = nak = \mathrm{e}^{\chi e}\mathrm{e}^{-\phi/2h}k = \begin{pmatrix} 1 & \chi \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{-\phi/2} & 0 \\ 0 & \mathrm{e}^{\phi/2} \end{pmatrix} k, \tag{4.1.7}$$

where $k \in \mathrm{SO}(2)$. If $k \in \mathrm{SO}(1,1)$ the decomposition will not be unique. In addition, it will not work for all elements. Therefore the composition is not mathematically stringent, although it will work for all cases of physical interest.[4] The factorization can be used with k being an element of either $\mathrm{SO}(2)$ or $\mathrm{SO}(1,1)$, thus for both spacelike and timelike reductions.

Because of the factorization, elements in the coset space $\mathrm{SL}(2,\mathbb{R})/\mathbf{K}$ can be written on the form

$$\mathcal{V} = na = \begin{pmatrix} 1 & \chi \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{-\phi/2} & 0 \\ 0 & \mathrm{e}^{\phi/2} \end{pmatrix}. \tag{4.1.8}$$

As this is a left coset, $\mathcal{V}$ transforms with elements in $\mathrm{SL}(2,\mathbb{R})$ from the left and with elements in **K** from the right,

$$\mathcal{V} \to g\mathcal{V}k \quad g \in \mathrm{SL}(2,\mathbb{R}), \quad k \in \mathbf{K}. \tag{4.1.9}$$

By definition, any two elements in the coset space that just differ with an element $k \in \mathbf{K}$ from the right, are treated as equivalent elements. Thus one can always find a representation of an element in the coset space that is on the form of $\mathcal{V}$ in (4.1.8), by setting k in (4.1.7) to be identity. When $\mathcal{V}$ transforms with g from the left it can be quite difficult to find the right k that acts from the right to make the new $\mathcal{V}$ again on the form (4.1.8). Instead we create an object that does not transform at all under **K** by defining

$$M = \mathcal{V}\mathcal{V}^{\mathcal{T}}, \tag{4.1.10}$$

where ${}^{\mathcal{T}}$ is the generalized transpose, see definition 3.14, which is defined by requiring

$$k^{\mathcal{T}} = k^{-1}, \quad k \in \mathbf{K}. \tag{4.1.11}$$

[4]A detailed discussion can be found in section 5.2 in [17].

This requirement is useful since M will then transform as

$$M \to (g\mathcal{V}k)(g\mathcal{V}k)^{\mathcal{T}} = g\mathcal{V}kk^{\mathcal{T}}\mathcal{V}^{\mathcal{T}}g^{\mathcal{T}} = g\mathcal{V}kk^{-1}\mathcal{V}^{\mathcal{T}}g^{\mathcal{T}} = gMg^{\mathcal{T}}. \tag{4.1.12}$$

When **K**= SO(2) we can say that the generalized transpose $^{\mathcal{T}}$ is just the normal transpose $^{\top}$ since this then has the desired property of (4.1.11). If we instead would have a timelike reduction, k would be an element in SO(1,1) and for $^{\mathcal{T}}$ to have the property $k^{\mathcal{T}} = k^{-1}$ the normal transpose does not suffice. The generalized transpose of SO(1,1) is $g^{\mathcal{T}} = cg^{\top}c$ where $c = \mathrm{diag}\{1, -1\}$. This generalized transpose has the effect of switching the sign of the non-diagonal elements. M expressed in ϕ and χ will then be

$$\begin{aligned} M &= \mathcal{V}\mathcal{V}^{\mathcal{T}} \\ &= \begin{pmatrix} 1 & \chi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-\phi/2} & 0 \\ 0 & \mathrm{e}^{\phi/2} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-\phi/2} & 0 \\ 0 & \mathrm{e}^{\phi/2} \end{pmatrix}^{\top} \begin{pmatrix} 1 & \pm\chi \\ 0 & 1 \end{pmatrix}^{\top} \\ &= \begin{pmatrix} \mathrm{e}^{-\phi} \pm \chi^2\mathrm{e}^{\phi} & \chi\mathrm{e}^{\phi} \\ \pm\chi\mathrm{e}^{\phi} & \mathrm{e}^{\phi} \end{pmatrix}. \end{aligned} \tag{4.1.13}$$

To be able to calculate the Killing form we go to $\mathfrak{sl}(2,\mathbb{R})$, the Lie algebra of SL(2,$\mathbb{R}$), and we do this via the Maurer-Cartan form,

$$M^{-1}\partial_\mu M \in \mathfrak{sl}(2,\mathbb{R}), \tag{4.1.14}$$

which becomes

$$M^{-1}\partial_\mu M = \begin{pmatrix} -\partial_\mu\phi \pm \chi\partial_\mu\chi\mathrm{e}^{2\phi} & \partial_\mu\chi\mathrm{e}^{2\phi} \\ \pm 2\chi\partial_\mu\phi \pm \partial_\mu\chi - \chi^2\partial_\mu\chi\mathrm{e}^{2\phi} & \partial_\mu\phi \mp \chi\partial_\mu\chi\mathrm{e}^{2\phi} \end{pmatrix}. \tag{4.1.15}$$

Our Killing form will then be

$$4\,\mathrm{Tr}\Big((M^{-1}\partial_\mu M)(M^{-1}\partial^\mu M)\Big), \tag{4.1.16}$$

which is manifestly invariant under SL(2,$\mathbb{R}$). We show this, keeping in mind that $M \to gMg^{\top}$. By restriction g to be a constant, the derivative only acts on M. By expanding the parentheses and recalling from elementary algebra that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, we see that

$$\mathrm{Tr}\Big((M^{-1}\partial_\mu M)(M^{-1}\partial^\mu M)\Big) \to$$

$$\to \mathrm{Tr}\Big(\Big((gMg^{\top})^{-1}\partial_\mu(gMg^{\top})\Big)\Big((gMg^{\top})^{-1}\partial^\mu(gMg^{\top})\Big)\Big) \tag{4.1.17}$$

$$= \mathrm{Tr}\Big((M^{-1}\partial_\mu M)(M^{-1}\partial^\mu M)\Big). \tag{4.1.18}$$

Next we calculate the explicit expression of the Killing form. After some work we arrive at

$$4\,\mathrm{Tr}\left((M^{-1}\partial_\mu M)(M^{-1}\partial^\mu M)\right) = 8(\partial_\mu\phi\partial^\mu\phi \pm \partial_\mu\chi\partial^\mu\chi \mathrm{e}^{2\phi}). \tag{4.1.19}$$

We see that this is the same as the expression for the Lagrangian in equation (4.1.6) apart from a factor 16. The factor 16 does not make any difference since if the Killing form is invariant under SL(2,$\mathbb{R}$) then so is any multiple of it. Therefore we have proven that our Lagrangian is indeed invariant under SL(2,$\mathbb{R}$). Because of the restriction that g should be independent of the spacetime coordinates, transformations under SL(2,$\mathbb{R}$) will represent global symmetries. On the other hand there where no restrictions on k, so transformations under $\mathbf{K}$ represent local symmetries.

4.2 Equations of Motion on the Coset Space

In this section we derive the equations of motion by using variational calculus on the action in equation (4.1.4). We do this by varying the action, first with respect to the fundamental object $\mathcal{V}$ and then with respect to the metric $g_{\mu\nu}$. These two equations of motion will not describe motion in the spacetime, rather they are equations of motion in a more abstract sense on the coset space where coordinates are the parameters χ and ϕ. One of these equations can be expressed on an alternative form as a geodesic equation (this is further discussed in section 4.2.3).

These two equations of motion are used to derive the metric. To express one of the equations as a geodesic equation is not necessary when deriving the Schwarzschild solution, although in other cases it can be a central part of finding the solution.

4.2.1 Variation with Respect to $\mathcal{V}$

The action from equation (4.1.4) is

$$S_3 - S_{\mathbf{G}/\mathbf{K}} = \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}R - \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\mathcal{L}(\chi,\phi). \tag{4.2.1}$$

When we vary $\mathcal{V}$ this will only affect the second term, which can be written using the Killing form as

$$S_{\mathbf{G}/\mathbf{K}} = \frac{1}{4}\int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\,\mathrm{Tr}\left((M^{-1}\partial_\mu M)(M^{-1}\partial^\mu M)\right). \tag{4.2.2}$$

First we write the action on another form, which makes the calculations easier. The Cartan involution states that we can always expand any element in a Lie algebra in one symmetric part and one antisymmetric. Therefore we can expand the Maurer-Cartan form of the element $\mathcal{V}$, which lies in the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ as

$$\mathcal{V}^{-1}\partial_\mu\mathcal{V} = Q_\mu + P_\mu, \tag{4.2.3}$$

where Q_μ is antisymmetric and P_μ is symmetric with respect to the generalized transpose. When we have a matrix representation of an algebra this is always true due to the fact that you can always write a matrix A as the sum of its symmetric and antisymmetric part:

$$A = \frac{1}{2}(A - A^{\mathcal{T}}) + \frac{1}{2}(A + A^{\mathcal{T}}). \tag{4.2.4}$$

This gives expressions for Q_μ and P_μ:

$$Q_\mu = \frac{1}{2}\Big(\mathcal{V}^{-1}\partial_\mu\mathcal{V} - (\mathcal{V}^{-1}\partial_\mu\mathcal{V})^{\mathcal{T}}\Big) \tag{4.2.5}$$

$$P_\mu = \frac{1}{2}\Big(\mathcal{V}^{-1}\partial_\mu\mathcal{V} + (\mathcal{V}^{-1}\partial_\mu\mathcal{V})^{\mathcal{T}}\Big) \tag{4.2.6}$$

To be able to express the Killing form in terms of Q_μ and P_μ, the Maurer-Cartan form of $M = \mathcal{V}\mathcal{V}^{\mathcal{T}}$ is expanded. Note that the derivative will act element-wise on the matrix $\mathcal{V}$, and therefore the differentiation and transposition may change order, so that $\partial_\mu(\mathcal{V}^{\mathcal{T}}) = (\partial_\mu\mathcal{V})^{\mathcal{T}}$. When expanding, multiplying with $(\mathcal{V}^{\mathcal{T}})^{-1}\mathcal{V}^{\mathcal{T}}$ makes it possible to simplify the expression:

$$\begin{aligned} M^{-1}\partial_\mu M &= (\mathcal{V}\mathcal{V}^{\mathcal{T}})^{-1}\partial_\mu(\mathcal{V}\mathcal{V}^{\mathcal{T}}) \\ &= (\mathcal{V}^{-1})^{\mathcal{T}}\Big(\mathcal{V}^{-1}\partial_\mu\mathcal{V}\mathcal{V}^{\mathcal{T}} + \partial_\mu\mathcal{V}^{\mathcal{T}}((\mathcal{V}^{\mathcal{T}})^{-1}\mathcal{V}^{\mathcal{T}})\Big) \\ &= (\mathcal{V}^{-1})^{\mathcal{T}}(2P_\mu)\mathcal{V}^{\mathcal{T}}. \end{aligned} \tag{4.2.7}$$

The next step is to insert this expression into the Lagrangian of the action in (4.2.2). Expanding and utilizing that Tr(AB)=Tr(BA), the factors of $\mathcal{V}$ cancel, and the Lagrangian becomes

$$\begin{aligned} \mathcal{L}(\chi,\phi) &= \frac{1}{4}\operatorname{Tr}\bigg(\Big((\mathcal{V}^{-1})^{\mathcal{T}}2P_\mu\mathcal{V}^{\mathcal{T}}\Big)\Big((\mathcal{V}^{-1})^{\mathcal{T}}2P^\mu\mathcal{V}^{\mathcal{T}}\Big)\bigg) \\ &= \operatorname{Tr}(P_\mu P^\mu), \end{aligned} \tag{4.2.8}$$

which gives the action in equation (4.2.2) as

$$S_{\mathbf{G/K}} = \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\,\operatorname{Tr}(P_\mu P^\mu). \tag{4.2.9}$$

To arrive at the equations of motion this should be stationary with respect to a small perturbation of $\mathcal{V}$, thus $\delta_{\mathcal{V}} S_{\mathbf{G}/\mathbf{K}} = 0$. We begin by looking at the variation of P_μ. From equation (2.4.22), the variation of the inverse of $\mathcal{V}$ will be $\delta(\mathcal{V}^{-1}) = -\mathcal{V}^{-1}\delta\mathcal{V}\mathcal{V}^{-1}$. Utilizing this in the variation of P_μ results in

$$\begin{aligned}\delta_{\mathcal{V}} P_\mu = \frac{1}{2}\Big(&- (\mathcal{V}^{-1}\delta\mathcal{V}\mathcal{V}^{-1})\partial_\mu \mathcal{V} + \mathcal{V}^{-1}\partial_\mu \delta\mathcal{V} + \\ &+ \Big(- (\mathcal{V}^{-1}\delta\mathcal{V}\mathcal{V}^{-1})\partial_\mu \mathcal{V}\Big)^{\mathcal{T}} + (\mathcal{V}^{-1}\partial_\mu \delta\mathcal{V})^{\mathcal{T}}\Big). \end{aligned} \tag{4.2.10}$$

Because $\mathcal{V}^{-1}\delta\mathcal{V}$ is a Lie algebra element, we can expand it into one antisymmetric and one symmetric part,

$$\mathcal{V}^{-1}\delta\mathcal{V} = \Sigma + \Lambda. \tag{4.2.11}$$

The antisymmetric part Σ will not have any effect as the Lagrangian only depends on the symmetric part P_μ and not on Q_μ. Therefore we can freely set $\Sigma = 0$, which results in

$$\Lambda = \frac{1}{2}\Big(\mathcal{V}^{-1}\delta\mathcal{V} + (\mathcal{V}^{-1}\delta\mathcal{V})^{\mathcal{T}}\Big) = \mathcal{V}^{-1}\delta\mathcal{V}. \tag{4.2.12}$$

Thus $\mathcal{V}^{-1}\delta\mathcal{V}$ is symmetric, which will be useful in further calculations. To simplify equation (4.2.10) we start by calculating the derivative of Λ,

$$\begin{aligned}\partial_\mu \Lambda &= \frac{1}{2}\partial_\mu\Big(\mathcal{V}^{-1}\delta\mathcal{V} + (\mathcal{V}^{-1}\delta\mathcal{V})^{\mathcal{T}}\Big) \\ &= \frac{1}{2}\Big(\partial_\mu \mathcal{V}^{-1}\delta\mathcal{V} + (\partial_\mu \mathcal{V}^{-1}\delta\mathcal{V})^{\mathcal{T}} + \mathcal{V}^{-1}\partial_\mu\delta\mathcal{V} + (\mathcal{V}^{-1}\partial_\mu\delta\mathcal{V})^{\mathcal{T}}\Big). \end{aligned} \tag{4.2.13}$$

Comparing with equation (4.2.10) gives

$$\begin{aligned}\delta_{\mathcal{V}} P_\mu = \partial_\mu\Lambda - \frac{1}{2}\Big(&\partial_\mu \mathcal{V}^{-1}\delta\mathcal{V} + (\partial_\mu \mathcal{V}^{-1}\delta\mathcal{V})^{\mathcal{T}} + (\mathcal{V}^{-1}\delta\mathcal{V}\mathcal{V}^{-1})\partial_\mu \mathcal{V} + \\ &+ \Big((\mathcal{V}^{-1}\delta\mathcal{V}\mathcal{V}^{-1})\partial_\mu \mathcal{V}\Big)^{\mathcal{T}}\Big) \end{aligned} \tag{4.2.14}$$

Once again using equation (2.4.22), which holds for derivatives as well as variations, the variation is

$$\begin{aligned}\delta_{\mathcal{V}} P_\mu = \partial_\mu\Lambda - \frac{1}{2}\Big(&(-\mathcal{V}^{-1}\partial_\mu \mathcal{V}\mathcal{V}^{-1})\delta\mathcal{V} + \Big((-\mathcal{V}^{-1}\partial_\mu \mathcal{V}\mathcal{V}^{-1})\delta\mathcal{V}\Big)^{\mathcal{T}} + \\ &+ (\mathcal{V}^{-1}\delta\mathcal{V}\mathcal{V}^{-1})\partial_\mu \mathcal{V} + \Big((\mathcal{V}^{-1}\delta\mathcal{V}\mathcal{V}^{-1})\partial_\mu \mathcal{V}\Big)^{\mathcal{T}}\Big) \end{aligned} \tag{4.2.15}$$

Simplifying the expression shows that the last terms can be expressed with Q_μ and Λ:

$$\begin{aligned}\delta_{\mathcal{V}} P_\mu &= \partial_\mu \Lambda + Q_\mu \Lambda - \Lambda Q_\mu \\ &= \partial_\mu \Lambda + [Q_\mu, \Lambda]\end{aligned} \tag{4.2.16}$$

Because Λ is a scalar, its partial derivative and covariant derivative are equal, we can write $\nabla_\mu \Lambda$ instead of $\partial_\mu \Lambda$. Next we look at the variation of the action from equation (4.2.9) and set this to zero to arrive at our equations of motion. The action contains two elements depending on $\mathcal{V}$; P_μ and its contravariant form P^μ. From the product rule we get two terms that contribute to the variation. By raising and lowering the index of the terms, these terms are equal. The variation can therefore be written as

$$\delta_{\mathcal{V}} \int_{\mathbb{R}^3} \sqrt{g} \mathrm{d}^3 x \, \mathrm{Tr}(P_\mu P^\mu) = 0 \tag{4.2.17}$$

$$\Rightarrow$$

$$2 \int_{\mathbb{R}^3} \sqrt{g} \mathrm{d}^3 x \, \mathrm{Tr}(P^\mu \delta_{\mathcal{V}} P_\mu) = 0 \tag{4.2.18}$$

$$\Rightarrow$$

$$\int_{\mathbb{R}^3} \sqrt{g} \mathrm{d}^3 x \, \mathrm{Tr}\left(P^\mu (\nabla_\mu \Lambda + [Q_\mu, \Lambda])\right) = 0 \tag{4.2.19}$$

To continue we use the following relation, $\mathrm{Tr}(A[B{,}C]) = \mathrm{Tr}([A{,}B]C)$ which is true for arbitrary matrices. Also, we integrate the first term by parts,

$$0 - \int_{\mathbb{R}^3} \sqrt{g} \mathrm{d}^3 x \, \mathrm{Tr}(\nabla_\mu P^\mu \Lambda) + \int_{\mathbb{R}^3} \sqrt{g} \mathrm{d}^3 x \, \mathrm{Tr}([P^\mu{,}Q_\mu]\Lambda) = 0 \tag{4.2.20}$$

$$\Rightarrow$$

$$\int_{\mathbb{R}^3} \sqrt{g} \mathrm{d}^3 x \, \mathrm{Tr}\left((-\nabla_\mu P^\mu + [P^\mu{,}Q_\mu])\Lambda\right) = 0. \tag{4.2.21}$$

If this is going to be zero for an arbitrary perturbation $\delta\mathcal{V}$, thus an arbitrary Λ, the following must be true,

$$\nabla_\mu P^\mu - [P^\mu{,}Q_\mu] = 0 \tag{4.2.22}$$

which are then the equations of motion. We can also express this in terms of a different covariant derivative which we define as

$$D_\mu P_\nu \equiv \nabla_\mu P_\nu + [Q_\mu{,}P_\nu]. \tag{4.2.23}$$

Then the equations of motion can be expressed as

$$D^\mu P_\mu = 0. \tag{4.2.24}$$

4.2.2 Variation with Respect to the Metric $g_{\mu\nu}$

When varying the action with respect to the metric we must take both terms from the dimensional reduction into account,

$$S_3 - S_{\mathbf{G}/\mathbf{K}} = \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}R - \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\,\mathrm{Tr}(P_\mu P^\mu). \tag{4.2.25}$$

The variation will be done exactly as when we derived Einstein's equations in section 2.4.3, but here there will be another term from $\sqrt{g}\mathcal{L}(\chi,\phi)$. For details compare with the calculations in section 2.4.3. The result is

$$\delta_g(S_3 - S_{\mathbf{G}/\mathbf{K}}) = 0 \tag{4.2.26}$$

$$\Rightarrow$$

$$\int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\delta g_{\rho\sigma}\left(\frac{1}{2}g^{\rho\sigma}(R - \mathrm{Tr}(P_\mu P^\mu)) - (R^{\rho\sigma} - \mathrm{Tr}(P^\rho P^\sigma))\right) = 0. \tag{4.2.27}$$

Since the action must be stationary for an arbitrary variation of the metric, the expression inside the parentheses must be zero. Gathering terms with the Ricci tensor and Ricci scalar to the left, and terms containing P_μ to the right, the requirement for stationary action is

$$R^{\rho\sigma} - \frac{1}{2}g^{\rho\sigma}R = \mathrm{Tr}(P^\rho P^\sigma) - \frac{1}{2}g^{\rho\sigma}\,\mathrm{Tr}(P_\mu P^\mu) \tag{4.2.28}$$

To simplify this equation we multiply both sides with $g_{\rho\sigma}$ and use that the trace of this is three, since it is the three-dimensional metric,

$$R - \frac{1}{2}3R = \mathrm{Tr}(P_\mu P^\mu) - \frac{3}{2}\,\mathrm{Tr}(P_\mu P^\mu) \tag{4.2.29}$$

$$\Rightarrow$$

$$-\frac{1}{2}R = -\frac{1}{2}\,\mathrm{Tr}(P_\mu P^\mu). \tag{4.2.30}$$

This makes the terms with R and $\mathrm{Tr}(P_\mu P^\mu)$ cancel each other out in (4.2.28) and we arrive at the expression for the second equation of motion:

$$R^{\mu\nu} = \mathrm{Tr}(P^\mu P^\nu) \tag{4.2.31}$$

$$\Rightarrow$$

$$R_{\mu\nu} = \mathrm{Tr}(P_\mu P_\nu) \tag{4.2.32}$$

From Einsteins equations in section 2.4, the right hand side of (4.2.32) can be seen as a stress-tensor. So we see that these equations are not just three-dimensional general relativity in vacuum, but the dimension that was reduced will still have a noticeable effect on the solution.

4.2.3 Geodesic on the Coset Space

In this section we show that the equations of motion from the dimensional reduction actually give a geodesic equation on the coset space SL(2,$\mathbb{R}$)/**K**, where **K** is SO(2) for a spacelike reduction and SO(1,1) for a timelike reduction. Because we in the end want to arrive at the Schwarzschild solution, we here work with a timelike reduction as this simplifies the calculations. For a spacelike reduction the result is the same and the calculations are very similar. The Lagrangian $\mathcal{L}(\chi,\phi)$ from equation (4.1.6) can be rewritten as

$$\mathcal{L}(\Phi) = \gamma_{ij}(\Phi)\partial_\mu\Phi^i\partial^\mu\Phi^j, \tag{4.2.33}$$

where $\Phi = (\chi,\phi)$ are the coordinates and γ_{ij} is the metric of the coset space SL(2,$\mathbb{R}$)/ SO(1,1)[5]. The second term from the action in equation (4.1.4) is then given by

$$\begin{aligned}
S_{\mathbf{G}/\mathbf{K}} &= \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\mathcal{L}(\Phi) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\gamma_{ij}(\Phi)\partial_\mu\Phi^i\partial^\mu\Phi^j \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}g^{\mu\nu}\gamma_{ij}(\Phi)\partial_\mu\Phi^i\partial_\nu\Phi^j.
\end{aligned} \tag{4.2.34}$$

If our solution is going to be spherically symmetric we can make the following ansatz for the three-dimensional metric[6]:

$$\mathrm{d}s^2 = -\mathrm{d}r^2 - f^2(r)(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2) \tag{4.2.35}$$

$$\Rightarrow$$

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -f^2 & 0 \\ 0 & 0 & -f^2\sin^2\theta \end{pmatrix} \tag{4.2.36}$$

Using this in equation (4.2.34) with the fact that Φ only depends on r leads to

$$\begin{aligned}
S_{\mathbf{G}/\mathbf{K}} &= \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}g^{\mu\nu}\gamma_{ij}(\Phi)\partial_\mu\Phi^i\partial_\nu\Phi^j \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x f^2\sin\theta\gamma_{ij}(\Phi)\partial_r\Phi^i\partial_r\Phi^j \\
&= 4\pi\int_0^\infty \mathrm{d}r f^2\gamma_{ij}(\Phi)\partial_r\Phi^i\partial_r\Phi^j
\end{aligned} \tag{4.2.37}$$

[5]The components of the metric γ_{ij} can be identified directly from (4.1.6).
[6]The minus sign is needed to obtain the correct signature of the four-dimensional metric.

To get rid of the factor f^2 in the integral we make the substitution

$$\tau(r) = -\int_r^\infty \frac{1}{f^2(s)}\mathrm{d}s \tag{4.2.38}$$

$$\Rightarrow$$

$$\frac{\partial\tau}{\partial r} = \frac{1}{f^2(r).} \tag{4.2.39}$$

The action then becomes

$$\begin{aligned} S_{\mathbf{G}/\mathbf{K}} &= 4\pi\int_0^\infty \mathrm{d}r f^2\gamma_{ij}(\Phi)\partial_r\Phi^i\partial_r\Phi^j \\ &= 4\pi\int_{\tau(0)}^0 \frac{\partial r}{\partial\tau}\mathrm{d}\tau f^2\gamma_{ij}(\Phi)\partial_\tau\Phi^i\partial_\tau\Phi^j\left(\frac{\partial\tau}{\partial r}\right)^2 \\ &= 4\pi\int_{\tau(0)}^0 \mathrm{d}\tau\gamma_{ij}(\Phi)\partial_\tau\Phi^i\partial_\tau\Phi^j. \end{aligned} \tag{4.2.40}$$

Variating with respect to Φ gives

$$\delta_\Phi S_{\mathbf{G}/\mathbf{K}} = 4\pi\int_{\tau(0)}^0 \mathrm{d}\tau\bigg(\delta_\Phi\Big(\gamma_{ij}(\Phi)\Big)\partial_\tau\Phi^i\partial_\tau\Phi^j + 2\gamma_{ij}(\Phi)\partial_\tau\Phi^i\partial_\tau\delta\Phi^j\bigg). \tag{4.2.41}$$

Given that this is going to be zero for an arbitrary perturbation $\delta\Phi^i$ yields the equation

$$\partial_\tau^2\Phi^m + \Upsilon^m_{ij}\partial_\tau\Phi^i\partial_\tau\Phi^j = 0, \tag{4.2.42}$$

where Υ is the affine connection on the coset space. The specific calculations that give this result are given in appendix E.

Looking back at the geodesic equation that was introduced in the text about general relativity, equation (2.3.15), we see that this equation is really a geodesic on the coset space. Actually the geodesic equation of (4.2.42) expresses the same thing as equation (4.2.24), $D^\mu P_\mu = 0$, from the variation of $\mathcal{V}$. Therefore the equation $D^\mu P_\mu = 0$ can also be seen as a geodesic equation. We state this without proof.

In some cases, the geodesic can be used to obtain the solution for the metric. Otherwise, it will provide some interesting information about the solutions.

As geodesics in spacetime have been thoroughly analyzed, this theory can be used for geodesics on the coset space as well. A geodesic is described by the geodesic equation and two parameters; the starting point on the coset space and the initial velocity. In four-dimensional spacetime, the initial velocity corresponds to conserved charges [18]. The geodesic equation can thus be used to analyze general properties of possible solutions. This will be further discussed in section 4.5.

4.3 Solutions to the Equations of Motion

To obtain the Schwarzschild solution, we use the equations of motion (4.2.24) and (4.2.32). These are

$$\begin{cases} D^\mu P_\mu &= 0 \\ R_{\mu\nu} &= \mathrm{Tr}(P_\mu P_\nu). \end{cases} \tag{4.3.1}$$

From equation (4.2.8) we see that second equation of motion is equal to

$$R_{\mu\nu} = \frac{1}{4}\,\mathrm{Tr}\left((M^{-1}\partial_\mu M)(M^{-1}\partial_\nu M)\right), \tag{4.3.2}$$

while the first equation of motion can be rewritten as

$$\nabla^\mu(M^{-1}\partial_\mu M) = 0. \tag{4.3.3}$$

The proof of this is given in appendix F. Our equations of motion are then

$$\begin{cases} \nabla^\mu(M^{-1}\partial_\mu M) &= 0 \\ R_{\mu\nu} &= \frac{1}{4}\,\mathrm{Tr}((M^{-1}\partial_\mu M)(M^{-1}\partial_\nu M)). \end{cases} \tag{4.3.4}$$

This result holds both for a spacelike and timelike dimensional reduction, though we will in this chapter work only with a timelike reduction. To arrive at the Schwarzschild solution, which is spherically symmetric, we write the three-dimensional metric in the same way as in the ansatz of equation (4.2.36), which is the most general form for a spherically symmetric metric. This then gives the components of the affine connection and the Ricci tensor. The only non-zero components of the connection are

$$\begin{aligned} \Gamma^r_{\theta\theta}\sin^2\theta &= \Gamma^r_{\varphi\varphi} &&= -f'f\sin^2\theta \\ \Gamma^\theta_{(r\theta)} &= \Gamma^\varphi_{(r\varphi)} &&= \frac{f'}{f} \\ & \quad\; \Gamma^\theta_{\varphi\varphi} &&= -\ \sin\theta\cos\theta \\ & \quad\; \Gamma^\varphi_{(\theta\varphi)} &&= \cot\theta \end{aligned} \tag{4.3.5}$$

and the Ricci tensor becomes

$$R_{\mu\nu} = \begin{pmatrix} -2\dfrac{f''}{f} & 0 & 0 \\ 0 & \dfrac{\mathrm{d}}{\mathrm{d}r}\Big(ff'\Big) - 1 & 0 \\ 0 & 0 & \sin^2\theta\left(\dfrac{\mathrm{d}}{\mathrm{d}r}\Big(ff'\Big) - 1\right) \end{pmatrix} \tag{4.3.6}$$

From the spherical symmetry we also know that $\partial_\mu M$ is only non-zero when $\mu = r$. Because of this, we can simplify the equations of (4.3.4) further. Inserting the definition of the covariant derivative into the first of the equations of (4.3.4), we note that there are only three non-zero terms:

$$\nabla^\mu(M^{-1}\partial_\mu M) = g^{rr}\partial_r(M^{-1}\partial_r M) - g^{\theta\theta}\Gamma^r_{\theta\theta}M^{-1}\partial_r M - g^{\varphi\varphi}\Gamma^r_{\varphi\varphi}M^{-1}\partial_r M. \quad (4.3.7)$$

Inserting the affine connections and the metric components gives

$$0 = -\partial_r(M^{-1}\partial_r M) + \frac{1}{f^2}(-f'f)M^{-1}\partial_r M \quad (4.3.8)$$

$$+\frac{1}{f^2\sin^2\theta}(-f'f\sin^2\theta)M^{-1}\partial_r M \quad (4.3.9)$$

$$= -\partial_r(M^{-1}\partial_r M) - 2\frac{f'}{f}M^{-1}\partial_r M. \quad (4.3.10)$$

Multiplying by $-f^2$ we obtain the following equation:

$$f^2\partial_r(M^{-1}\partial_r M) + 2ff'M^{-1}\partial_r M = 0 \quad (4.3.11)$$

$$\Rightarrow$$

$$\partial_r(f^2M^{-1}\partial_r M) = 0 \quad (4.3.12)$$

If we once again introduce the parameter $\tau(r)$ from equation (4.2.38) we get

$$\partial_\tau(\hat{M}^{-1}\partial_\tau\hat{M}) = 0 \quad (4.3.13)$$

where $\hat{M}(\tau(r)) = M(r)$. Looking back at the equations in (4.3.4) we see that we have the four equations,

$$\begin{cases} \partial_\tau(\hat{M}^{-1}\partial_\tau\hat{M}) & = 0 \\ R_{rr} & = \frac{1}{4}\operatorname{Tr}\left((M^{-1}\partial_r M)^2\right) \\ R_{\theta\theta} & = 0 \\ R_{\varphi\varphi} & = 0 \end{cases} \quad (4.3.14)$$

The two last equations actually says the same thing as $R_{\varphi\varphi} = \sin^2\theta R_{\theta\theta}$. Now using what we know about the components of the Ricci tensor the equations of motion become

$$\begin{cases} \partial_\tau(\hat{M}^{-1}\partial_\tau\hat{M}) & = 0 \\ -2\frac{f''}{f} & = \frac{1}{4}\operatorname{Tr}\left((M^{-1}\partial_r M)^2\right) \\ \frac{\mathrm{d}}{\mathrm{d}r}(ff') - 1 & = 0 \end{cases} \quad (4.3.15)$$

The solution to the last equation will be

$$f^2 = (r+b)^2 + c. \tag{4.3.16}$$

where b and c are constants of integration. Let us then solve the first equation from (4.3.15):

$$\partial_\tau(\hat{M}^{-1}\partial_\tau \hat{M}) = 0 \tag{4.3.17}$$

$$\Rightarrow$$

$$\hat{M}^{-1}\partial_\tau \hat{M} = Q \tag{4.3.18}$$

$$\Rightarrow$$

$$\partial_\tau \hat{M} = \hat{M}Q \tag{4.3.19}$$

$$\Rightarrow$$

$$\hat{M} = M_0 \exp(\tau Q). \tag{4.3.20}$$

where Q and M_0 are some constant matrices yet to be determined. At infinity we want our solution to be flat, and when looking back at the original four-dimensional metric, equation (4.1.2), we see that this boundary conditions makes M go to identity when r goes to infinity. That r goes to infinity is the same as τ goes to zero. Therefore M_0 must be identity if $\hat{M}$ is to be identity when τ is zero.

From the spherical symmetry we also know that we can make our metric diagonal and therefore set A_μ to zero, which forces χ to be constant. This constant can be arbitrarily chosen since χ describes a potential. Setting χ to zero, makes M diagonal and from equation (4.3.20) we see that the same is true for Q. The definition of M states that $M = \mathcal{V}\mathcal{V}^\mathcal{T}$ and because $\mathcal{V} \in \mathrm{SL}(2,\mathbb{R})$ the determinant of M is one, therefore the trace of Q must be zero. If Q is diagonal and traceless we can write Q as a multiple of the generator h, which is one of the generators of $\mathfrak{sl}(2,\mathbb{R})$ and is given by $\mathrm{diag}\{1, -1\}$. Therefore we can write,

$$\hat{M} = \exp(\tau\alpha h). \tag{4.3.21}$$

We have one equation left: number two in (4.3.15). We solve this by using our previous results that $\hat{M}^{-1}\partial_\tau\hat{M} = Q = \alpha h$,

$$-2\frac{f''}{f} = \frac{1}{4}\,\mathrm{Tr}\left((M^{-1}\partial_r M)^2\right) \tag{4.3.22}$$

Making the same parameter substitution again we can simplify this to

$$-8\frac{f''}{f} = \frac{1}{f^4}\,\mathrm{Tr}\left((\hat{M}^{-1}\partial_\tau \hat{M})^2\right) \tag{4.3.23}$$

$$\Rightarrow$$

$$-8f^3 f'' = \mathrm{Tr}\left((\alpha h)^2\right) \tag{4.3.24}$$

And since $\mathrm{Tr}(h^2) = 2$,

$$-4f^3 f'' = \alpha^2. \tag{4.3.25}$$

In the previous calculations we found that $f^2 = (r+b)^2 + c$, see equation (4.3.16). From this we see that

$$f^3 = ((r+b)^2 + c)^{3/2} \tag{4.3.26}$$

$$f'' = \frac{c}{((r+b)^2 + c)^{3/2}} \tag{4.3.27}$$

$$\Rightarrow$$

$$f^3 f'' = c. \tag{4.3.28}$$

Equation (4.3.25) then relates the constant as $-4c = \alpha^2$. Since f is known we can calculate τ explicitly using partial fraction decomposition:

$$\tau(r) = -\int_r^\infty \frac{1}{f^2(s)} \mathrm{d}s \tag{4.3.29}$$

$$= \frac{1}{\alpha} \ln\left(\frac{r+b-\alpha/2}{r+b+\alpha/2}\right). \tag{4.3.30}$$

Putting this into equation (4.3.21) gives the final form of M,

$$M = \exp\left(\ln\left(\frac{r+b-\alpha/2}{r+b+\alpha/2}\right) h\right). \tag{4.3.31}$$

From the definition of M and the fact that $\chi = 0$ we can relate this to ϕ. We find that

$$\phi = -\ln\left(\frac{r+b-\alpha/2}{r+b+\alpha/2}\right) \tag{4.3.32}$$

Looking back at the metric that we started with from the dimensional reduction,

$$g_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} e^{-\phi} g_{\mu\nu} + e^{\phi} A_\mu A_\nu & e^{\phi} A_\mu \\ e^{\phi} A_\nu & e^{\phi} \end{pmatrix}, \tag{4.3.33}$$

where $g_{\mu\nu}$ is the three-dimensional metric, $\mathrm{diag}\{-1, -f^2, -f^2 \sin^2\theta\}$ and A_μ is now zero. When inserting this into the metric we arrive at

$$g_{\hat{\mu}\hat{\nu}} = \mathrm{diag}\Bigg\{ -\left(\frac{r+b-\alpha/2}{r+b+\alpha/2}\right), -(r+b-\alpha/2)^2, -(r+b-\alpha/2)^2 \sin^2\theta, \left(\frac{r+b+\alpha/2}{r+b-\alpha/2}\right) \Bigg\}. \tag{4.3.34}$$

In order to compare this to Schwarzschild, we want there to be a factor r^2 in front of the angular components. This can be done by setting $b = \alpha/2$, which then gives

$$g_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} -\left(\frac{r+\alpha}{r}\right)^{-1} & 0 & 0 & 0 \\ 0 & -r^2 & 0 & 0 \\ 0 & 0 & -r^2\sin^2\theta & 0 \\ 0 & 0 & 0 & \left(\frac{r+\alpha}{r}\right) \end{pmatrix}. \tag{4.3.35}$$

If this metric is going to fulfill the Newtonian limit when $r >> 1$ then α must be equal to $-2MG$. We end up at

$$g_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} -\left(1-\frac{2MG}{r}\right)^{-1} & 0 & 0 & 0 \\ 0 & -r^2 & 0 & 0 \\ 0 & 0 & -r^2\sin^2\theta & 0 \\ 0 & 0 & 0 & \left(1-\frac{2MG}{r}\right) \end{pmatrix}, \tag{4.3.36}$$

where the order of the coordinates are r,θ,ϕ,t. We see that this is really the Schwarzschild metric with signature $(---+)$. That the signature is different from the metric in section 2.5.1 has no physical implications, it is just a matter of choice which one you use.

4.4 Generalization to Reissner-Nordström Black Holes

In this section we go beyond the Schwarzschild solution and look at black holes that also carry charge. The metric that describes these black holes is called the Reissner-Nordström metric. Instead of starting with just the action as in the previous chapter,

$$S_4[g_{\hat{\mu}\hat{\nu}}] = \int_{\mathbb{R}^4} \mathrm{d}^4x\sqrt{g}R \tag{4.4.1}$$

we also require a component that is the stress-energy tensor from Maxwell theory,[7]

$$S_4[g_{\hat{\mu}\hat{\nu}}] + S_{Maxwell}[g_{\hat{\mu}\hat{\nu}}] = \int_{\mathbb{R}^4} \mathrm{d}^4x\sqrt{g}R - \frac{1}{4}\int_{\mathbb{R}^4} \mathrm{d}^4x\sqrt{g}F^{\hat{\mu}\hat{\nu}}F_{\hat{\mu}\hat{\nu}}, \tag{4.4.2}$$

where

$$F_{\hat{\mu}\hat{\nu}} = \partial_{[\hat{\mu}}\mathcal{A}_{\hat{\nu}]} \tag{4.4.3}$$

[7]This can be found in section 2.1 in [19].

and $\mathcal{A}_\mu$ is the 4-potential. We perform a dimensional reduction with respect to time and therefore we once again separate the time-component from the others, defining

$$\mathcal{A}_{\hat{\mu}} \equiv (\mathcal{A}_\mu, V). \tag{4.4.4}$$

Here V is the electric potential and $\mathcal{A}_\mu$ the three dimensional magnetic vector potential. After the dimensional reduction, $\mathcal{A}_\mu$ can just like A_μ in section 4.1, be dualized and represented with a scalar $\tilde{V}$, where $\partial^\mu \tilde{V} = \epsilon^{\mu\nu\sigma}(\partial_\nu \mathcal{A}_\sigma - \partial_\sigma \mathcal{A}_\nu)$. Together with the two previous scalars from the metric, we end up with a Lagrangian dependent on four scalar fields, ϕ, χ, $\tilde{V}$ and V:

$$\int_{\mathbb{R}^4} \mathrm{d}^4x\sqrt{g}R - \frac{1}{4}F^{\hat{\mu}\hat{\nu}}F_{\hat{\mu}\hat{\nu}} \tag{4.4.5}$$

$$\Downarrow$$

$$\int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}R - \int_{\mathbb{R}^3} \mathrm{d}^3x\sqrt{g}\mathcal{L}(\phi, \chi, \tilde{V}, V). \tag{4.4.6}$$

With the same reasoning as when we derived the Schwarzschild solution, we know that the metric is diagonizable, and thus we can again set $\chi = 0$. The scalar from the magnetic vector potential, $\tilde{V}$, can be set to zero if we assume that no magnetic monopoles exist and the solution is static. The four dimensional metric is then

$$g_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} g_{\mu\nu}\mathrm{e}^{2\phi} & 0 \\ 0 & \mathrm{e}^{-2\phi} \end{pmatrix} \tag{4.4.7}$$

and the explicit Lagrangian becomes:[8]

$$\mathcal{L}(\phi, V) = 2(\partial_\mu\phi\partial^\mu\phi - \mathrm{e}^{2\phi}\partial_\mu V\partial^\mu V). \tag{4.4.8}$$

We then make use of the same fundamental object M as in equation (4.1.13), with a minus sign since we are now working with a timelike reduction and V instead of χ. After inserting a spherically symmetric metric, the equations of motion will be similar to those in the solution for Schwarzschild, and from equation (4.3.15) we see that these are

$$\begin{cases} \partial_\tau(\hat{M}^{-1}\partial_\tau\hat{M}) &= 0 \\ -2\dfrac{f''}{f} &= \mathrm{Tr}\left((M^{-1}\partial_r M)^2\right) \\ \dfrac{\mathrm{d}}{\mathrm{d}r}\Big(ff'\Big) - 1 &= 0 \end{cases}. \tag{4.4.9}$$

[8] Equation (2.32) in [19]

Note the difference of a factor four in the second equation, which comes from that the Lagrangian in (4.4.8) is four times the Lagrangian we had when deriving the Schwarzschild solution. The first and last equation yields identical results as those in equation (4.3.16) and (4.3.20) respectively. Thus $M = M_0 \exp(\tau Q)$ and $f^2 = (r+b)^2 + c$, where M_0, Q, b and c are constants of integration. If the solution should be asymptotically flat at infinity, then ϕ and V must go to zero when r goes to infinity. When r goes to infinity τ goes to zero and from this we see that $M_0 = I$. As M is an element of SL(2,$\mathbb{R}$), Q can be written as a linear combination of the generators of SL(2,$\mathbb{R}$): h, e and f, defined in section 3.4.1. The explicit expression for $\exp\Big(\tau(\alpha h+\beta e+\gamma f)\Big)$ is derived in appendix G. For this to be equal to M in (4.1.13), that is antisymmetric in its off-diagonal elements, we see that $\gamma = -\beta$. M then becomes

$$M = \exp\Big(\tau(\alpha h + \beta e - \beta f)\Big) \tag{4.4.10}$$

$$= \begin{pmatrix} \cosh\left(\sqrt{E}\tau\right) + \frac{\alpha \sinh\left(\sqrt{E}\tau\right)}{\sqrt{E}} & \frac{\beta \sinh\left(\sqrt{E}\tau\right)}{\sqrt{E}} \\ -\frac{\beta \sinh\left(\sqrt{E}\tau\right)}{\sqrt{E}} & \cosh\left(\sqrt{E}\tau\right) - \frac{\alpha \sinh\left(\sqrt{E}\tau\right)}{\sqrt{E}} \end{pmatrix}, \tag{4.4.11}$$

where $E \equiv \alpha^2 - \beta^2$. Comparing with M expressed with ϕ and V we find that

$$\mathrm{e}^{\phi} = M_{22} = \cosh\left(\sqrt{E}\tau\right) - \frac{\alpha \sinh\left(\sqrt{E}\tau\right)}{\sqrt{E}}, \tag{4.4.12}$$

$$V = \frac{M_{12}}{M_{22}} = \frac{\beta \sinh\left(\sqrt{E}\tau\right)}{\sqrt{E}\cosh\left(\sqrt{E}\tau\right) - \alpha \sinh\left(\sqrt{E}\tau\right)}. \tag{4.4.13}$$

From the second equation in (4.4.9) we can determine the constant c in f^2. We do this precisely as in section 4.3 and find that

$$c = -\frac{1}{2}\operatorname{Tr}\left((M^{-1}\partial_\tau M)^2\right) \tag{4.4.14}$$

$$= -\frac{1}{2}\operatorname{Tr}(Q^2) \tag{4.4.15}$$

$$= -(\alpha^2 - \beta^2) \tag{4.4.16}$$

$$= -E. \tag{4.4.17}$$

This E is actually equal to the rest mass energy of the black hole, which we soon will see, thus we can say that $c < 0$.[9] As in section 4.3, τ can then be found to be

$$\tau = \frac{1}{2\sqrt{E}} \ln\left(\frac{r+b-\sqrt{E}}{r+b+\sqrt{E}}\right). \tag{4.4.18}$$

[9]More about this can be found in section 2.5

To find the constants α and β we use the newtonian limit with $\mathrm{e}^{-2\phi}$ as the gravitational potential and V as the electromagnetic potential. We therefore have the following conditions

$$\lim_{r\to\infty} \mathrm{e}^{-2\phi(r)} = 1 - \frac{2MG}{r}, \tag{4.4.19}$$

$$\lim_{r\to\infty} V(r) = \frac{Q}{r}. \tag{4.4.20}$$

To do this we calculate the derivative of these potentials with respect to $x \equiv \frac{1}{r}$:

$$\frac{\partial}{\partial x}\left(\mathrm{e}^{-2\phi(x)}\right)\bigg|_{x=0} = \frac{\partial}{\partial \tau}\left(\mathrm{e}^{-2\phi(\tau)}\right)\bigg|_{\tau=0} \frac{\partial \tau}{\partial x}\bigg|_{x=0} = (2\alpha)(-1), \tag{4.4.21}$$

$$\frac{\partial V(x)}{\partial x}\bigg|_{x=0} = \frac{\partial V(\tau)}{\partial \tau}\bigg|_{\tau=0} \frac{\partial \tau}{\partial x}\bigg|_{x=0} = (\beta)(-1). \tag{4.4.22}$$

From this we can directly say that $\alpha = MG$ and $\beta = -Q$ by using the Newtonian limit. Also we see that $E = (MG)^2 - Q^2$ and that this truly represents the rest mass energy. What is left is now to insert τ into $\mathrm{e}^{-2\phi}$ and use this to find the metric in (4.4.7). After some work we find that,

$$\mathrm{e}^{-2\phi} = \left(\cosh\left(\sqrt{E}\tau\right) - \frac{\alpha \sinh\left(\sqrt{E}\tau\right)}{\sqrt{E}}\right)^{-2} \tag{4.4.23}$$

$$= 4\left(\frac{\left(\frac{r+b-\sqrt{E}}{r+b+\sqrt{E}}\right)+1}{\sqrt{\frac{r+b-\sqrt{E}}{r+b+\sqrt{E}}}} - \frac{\alpha}{\sqrt{E}}\frac{\left(\frac{r+b-\sqrt{E}}{r+b+\sqrt{E}}\right)-1}{\sqrt{\frac{r+b-\sqrt{E}}{r+b+\sqrt{E}}}}\right)^{-2} \tag{4.4.24}$$

$$= \frac{(r+b)^2 - E}{(r+b+\alpha)^2}. \tag{4.4.25}$$

The numerator is precisely f^2 in (4.3.16) and if the angular components in the metric are going to be r^2 and $r^2 \sin^2\theta$ the denominator must be r^2, thus $b = -\alpha$. Finally we arrive at,

$$\mathrm{e}^{-2\phi} = \frac{(r-\alpha)^2 - E}{r^2} \tag{4.4.26}$$

$$= \frac{r^2 - 2MGr + Q^2}{r^2}. \tag{4.4.27}$$

Consequently f^2 is

$$f^2 = r^2 - 2MGr + Q^2. \tag{4.4.28}$$

Inserting this into the original four-dimensional metric,

$$g_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} g_{\mu\nu}\mathrm{e}^{2\phi} & 0 \\ 0 & \mathrm{e}^{-2\phi} \end{pmatrix} \tag{4.4.29}$$

then gives the Reissner-Nordström metric,

$$g_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} -\left(1 - \frac{2MG}{r} + \frac{Q^2}{r^2}\right)^{-1} & 0 & 0 & 0 \\ 0 & -r^2 & 0 & 0 \\ 0 & 0 & -r^2\sin^2\theta & 0 \\ 0 & 0 & 0 & \left(1 - \frac{2MG}{r} + \frac{Q^2}{r^2}\right) \end{pmatrix}, \tag{4.4.30}$$

where the order of the coordinates are r,θ,ϕ,t. A discussion on the Reissner-Nordström metric is found in section 2.5.

4.5 Further Use of Dimensional Reduction

In the previous sections, we have shown how a sigma model can be used to obtain the Schwarzschild and the Reissner-Nordström black hole solution. Compared to the simple derivation of the metric of section 2.5.1, the method presented prior in this chapter required significantly more work, even though many important steps were only mentioned briefly. However, the sigma model is very powerful as it provides a general tool to derive black hole solutions. In addition, it provides a framework to obtain entire classes of black hole solutions from a single solution [12]. Sigma models also play an important role in the discussion of uniqueness of black hole solutions [20]. In the following paragraphs, three main paths of developing different types of solutions will be discussed.

Firstly, other kinds of black hole solutions can be obtained using the the same method, but a different action. Alternatively, the dimensional reduction can be performed on another dimension. The more general the action, the larger the class of solutions will be. This will affect the Lagrangian after the dimensional reduction, which results in other quotient groups than SL(2,$\mathbb{R}$)/SO(2) or SL(2,$\mathbb{R}$)/ SO(1,1), as used in the Schwarzschild example. Otherwise, the method is similar. After obtaining the Schwarzschild and Reissner-Nordstöm solutions, a natural next step is to find the solutions to the axially symmetric, rotating Kerr (uncharged) or Kerr-Newman (charged) black holes.

Another path to follow from the Schwarzschild solution is to generate other solutions of the Schwarzschild family. Here, a few words about the geodesic equation obtained in section 4.2.3 is appropriate. All spherically symmetric black holes have geodesic equations of the coset space $\mathbf{G}/\mathbf{K}$.[10] Apart from sometimes being useful for finding the solution, the geodesic equation allows for an analysis of the general properties of the metric without explicitly finding it[12]. Of the two characteristics of a specific geodesic, the starting point and the initial velocity, the initial velocity corresponds to conserved charges in four dimensions. Since $\mathbf{K}$ is the subgroup of $\mathbf{G}$ that corresponds to a fixed starting point, the set of conserved charges is generated by transformations of $\mathbf{K}$[18]. All spherically symmetric non-extremal[11] black holes can be generated using $\mathbf{K}$-transformation starting from Schwarzschild.[18] [12]. The principle of generating new solutions is as follows: starting with the parameter $\mathcal{V}_{seed}$ defined as in equation (4.1.8), corresponding to one solution, we can form

$$M_{seed} = \mathcal{V}_{seed}\mathcal{V}_{seed}^{\mathcal{T}}. \tag{4.5.1}$$

M_{seed} will then, according to (4.1.12), transform as

$$M_{seed} \to M_g = gM_{seed}g^{\mathcal{T}}, \qquad g \in \mathbf{G} \tag{4.5.2}$$

If we can factorize M_g :

$$M_g = \mathcal{V}_g\mathcal{V}_g^{\mathcal{T}}, \tag{4.5.3}$$

the new solution $\mathcal{V}_g$ is obtained. Though simple in principle, solutions are not generated too easily [21]. Especially the factorization is often very complicated.

A third possibility to continue exploring dimensional reduction is to reduce on one more dimension, to the two-dimensional case. This path is however much more complicated than the previous two, both regarding dimensional reduction and factorization when classifying solutions. In return, it is very rewarding as it is possible to obtain all solutions with two independent spacetime symmetries from Minkowski metric. This technique is therefore still object to intense research. [14, 21]

[10] This is however not true for other symmetries; for example an axially symmetric black hole will not have such a geodesic.

[11] Extremal black holes have a nilopotent conserved charge, see [18].

5

Conclusion

In this thesis, we have described a general method of dimensional reduction from four to three dimensions that reveals hidden symmetries of black hole solutions. Using this technique, we have obtained the simplest possible form of a black hole, the Schwarzschild solution. Using the same technique, we have also presented how to derive the Reissner-Nordström black hole and thus demonstrated the method's power of versatility. Thereby, the aim of the project was achieved. By this manner of coding and utilizing symmetries we have paved a way to finding more complicated and exotic metrics that could perhaps not have been found with a less sophisticated method.

A next step to continue the work of this thesis is to generate other solutions belonging to the Schwarzschild or Reissner-Nordström families, as discussed in section 4.5. Alternatively, dimensional reduction can be performed on a more general action resulting in an larger class of solutions. These solutions can then be generated by acting with elements belonging to the coset group of the relevant action.

Although dimensional reduction was used to derive the Schwarzschild and Reissner-Nordström solutions decades ago, it is still subject to intense research. Among the topics of recent research we find the use of sigma models to investigate black hole solutions in supersymmetric theories of gravity [18, 22]. Another current field of research is on five-dimensional supergravity, see for example [18], and studies of dimensional reduction to two dimensions [21].

When performing a dimensional reduction on multiple dimensions or a more complicated action, the resulting Lagrangian becomes all the more complicated and contains all the more terms. The merit of this is exposing even deeper symmetries. There is hope of reducing the eleven-dimensional super-gravity to two dimensions, which would expose the full set of symmetries.[1] Thereby one could generate any black hole solution simply from the Minkowski metric.

Hopefully, this paper has provided a first glance into the fascinating theory of dimensional reduction and given some indication as to its strengths, its implications and its possibilities.

[1] This set is infinite-dimensional and is described by a Kac-Moody algebra, which is a Lie algebra defined by generators and relations of a generalized Cartan matrix. [16]

Bibliography

[1] Carroll, S.M. (1997) *Lecture Notes on General Relativity.* Institute for Theoretical Physics, University of California, Santa Barbara, USA. arXiv:gr-qc/9712019.

[2] Weinberg, S. (1972) *Gravitation and Cosmology.* John Wiley & Sons, Inc.

[3] Rothman, T. (2006) *Lost in Einstein's Shadow.* *American Scientist*, volume 94. .

[4] Carroll, S.M. (1999) *The Cosmological Constant.* Enrico Fermi Institute and Department of Physics, University of Chicago, Chicago, USA. arXiv:astro-ph/0004075.

[5] Hall, B.C. (2000) *An Elementary Introduction to Groups and Representations.* University of Notre Dame, Department of Mathematics, Notre Dame IN 46556 USA.

[6] Cahn, R.N. (1984) *Semi-Simple Lie Algebras and Their Representations.* Menlo Park, California: The Benjamin/Cummings Publishing Company.

[7] Georgi, H. (1999) *Lie Algebras in Particle Physics.* Westview Press.

[8] Igusa, K. (2011) *Semisimple Lie algebras and the Killing form.* Department of Mathematics, Brandeis University, Waltham, USA. `http://people.brandeis.edu/~igusa/Math223aF11/Notes223a05.pdf`.

[9] Conrad, K. *Decomposing* $SL_2(\mathbb{R})$. Math Departement, University of Conneticut,, Storrs, USA. `http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/SL(2,R).pdf`.

[10] Grebović, S., Radnäs, A., Ranjbar, A., Renneby, M., Toft, C. and Widén, E. (2012) *Group Theory and Symmetries in Particle Physics*. Department of Fundamental Physics, Chalmers Univeristy of Technology.

[11] Gustafsson, H. (2013). *Eisenstein Series and Instantons in String Theory*.

[12] Beritenlohner, P., Maison, D. and Gibbons, G. *4-dimensional black holes from Kaluza-Klein theory*. *Commun. Math. Phys.*, volume 120, 2, pp. 295–333.

[13] Beritenlohner, P. and Maison, D. (1986) *Solitons in Kaluza-Klein theories*. Max-Planck-Institut für Physik und Astrophysik, Werner-Heisenberg-Institut für Physik.

[14] Nicolai, H. (2005) *Gravitational billiards, dualities and hidden symmetries*. *100 Years Of Relativity : space-time structure: Einstein and beyond*, pages 39–75. arXiv:gr-qc/0506031.

[15] Pope, C. *Lectures on Kaluza-Klein Theory*. Department of Physics and Astronomy, Texas A&M University. `http://people.physics.tamu.edu/pope/ihplec.pdf`.

[16] Persson, D. (2005) *Field Realization of the Geroch Group*. Physique Théorique et Mathématique, Université Libre de Bruxelles & International Solvay Institutes. `http://danper.se/Daniels_homepage/Expository_Notes.html`.

[17] Bossard, G., Nicolai, H. and Stelle, K. (2009) *Universal BPS structure of stationary supergravity solutions*. Max-Planck-Institut für Gravitationsphysik Am Mühlenberg & Theoretical Physics group, Imperial College London. arXiv:0902.4438.

[18] Kim, S.S., Lindman Hörnlund, J., Palmkvist, J. and Virmani, A. (2010) *Extremal solutions of the S3 model and nilpotent orbits of G2(2)*. *Journal of High Energy Physics*, volume 2010, 8, 1–43. arXiv:hep-th/1004.5242.

[19] Houart, L., Kleinschmidt, A., Hörnlund, J., Persson, D. and Tabti, N. (2009) *Finite and infinite-dimensional symmetries of pure* $\mathcal{N} = 2$ *supergravity in* $D = 4$. Physique Théorique et Mathématique, Université Libre de Bruxelles & International Solvay Institutes. hep-th/arXiv:0905.4651.

[20] Mazur, P.O. (2000) *Black Hole Uniqueness Theorems*. Department of Physics and Astronomy, University of South Carolina. arXiv:hep-th/0101012.

[21] Katsimpouri, D., Kleinschmidt, A. and Virmani, A. (2013) *Inverse scattering and the Geroch group*. *Journal of High Energy Physics*, volume 2013, 2, 1–29. hep-th/arXiv:1211.3044 .

[22] Hörnlund, J.L. (2011) *Sigma-models and Lie group symmetries in theories of gravity.* Ph.D. thesis, l'Université Libre de Bruxelles.

A

Transformations

For the sake of brevity, this appendix adheres to the convention of denoting coordinates of flat Minkowski space as ξ^α with indices α, β, γ. General coordinates of curved spacetime are written as x^μ with indices μ, ν, σ, ρ...

A.1 Transformation of the Partial Derivative

In chapter 2.3.2 the covariant derivative is presented, as the partial derivative ∂_μ does not transform lika a tensor. The transformation of $\partial_\nu V^\mu$ can be calculated by use of the chain rule, according to

$$\begin{aligned}\partial'_\nu V'^\mu &= \\ \frac{\partial V'^\mu}{\partial x'^\nu} &= \\ \frac{\partial}{\partial x'^\nu}\left(\frac{\partial x'^\mu}{\partial x^\tau}V^\tau\right) &= \\ \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial}{\partial x^\rho}\left(\frac{\partial x'^\mu}{\partial x^\tau}V^\tau\right) &= \\ \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial V^\rho}{\partial x^\rho} + \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial^2 x'^\mu}{\partial x^\rho \partial x^\tau}V^\tau&.\end{aligned} \tag{A.1.1}$$

A.2 Transformation of the Affine Connection

Here we make use of an alternative and equivalent definition of the affine connection compared to the previous definition of (2.3.8), which can also be written as

$$\Gamma^{\nu}_{\mu\lambda} \equiv \frac{\partial x^{\nu}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\mu} \partial x^{\lambda}}. \tag{A.2.1}$$

It transforms according to

$$\begin{aligned}
\Gamma'^{\mu}_{\nu\lambda} &= \\
\frac{\partial x'^{\mu}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{\alpha}}{\partial x'^{\nu} \partial x'^{\lambda}} &= \\
\left(\frac{\partial x'^{\mu}}{\partial \xi^{\alpha}}\right) \frac{\partial}{\partial x'^{\nu}} \left(\frac{\partial \xi^{\alpha}}{\partial x'^{\lambda}}\right) &= \\
\left(\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}}\right) \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial}{\partial x^{\sigma}} \left(\frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \frac{\partial \xi^{\alpha}}{\partial x^{\kappa}}\right) &= \\
\left(\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}}\right) \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \left(\frac{\partial^2 x^{\kappa}}{\partial x^{\sigma} \partial x'^{\lambda}} \frac{\partial \xi^{\alpha}}{\partial x^{\kappa}} + \frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\sigma} \partial x^{\kappa}}\right) &= \\
\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial^2 x^{\kappa}}{\partial x^{\sigma} \partial x'^{\lambda}} \frac{\partial \xi^{\alpha}}{\partial x^{\kappa}} + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\sigma} \partial x^{\kappa}} &= \\
\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial}{\partial x^{\sigma}} \left(\frac{\partial x^{\kappa}}{\partial x'^{\lambda}}\right) \left(\frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \frac{\partial \xi^{\alpha}}{\partial x^{\kappa}}\right) + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \left(\frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\sigma} \partial x^{\kappa}}\right) &= \\
\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial}{\partial x'^{\nu}} \left(\frac{\partial x^{\kappa}}{\partial x'^{\lambda}}\right) \delta^{\rho}_{\kappa} + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \Gamma^{\rho}_{\sigma\kappa} &= \\
\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\lambda}} + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \Gamma^{\rho}_{\sigma\kappa}&.
\end{aligned} \tag{A.2.2}$$

When coupled with a vector, the affine connection transforms according to

$$\begin{aligned}
\Gamma'^{\mu}_{\nu\lambda} V'^{\lambda} &= \\
\left(\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\lambda}} + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \Gamma^{\rho}_{\sigma\kappa}\right) \left(\frac{\partial x'^{\lambda}}{\partial x^{\tau}} V^{\tau}\right) &= \\
\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\lambda}} \frac{\partial x'^{\lambda}}{\partial x^{\tau}} V^{\tau} + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\sigma\kappa} V^{\tau} \frac{\partial x^{\kappa}}{\partial x'^{\lambda}} \frac{\partial x'^{\lambda}}{\partial x^{\tau}} &= \\
\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\lambda}} \frac{\partial x'^{\lambda}}{\partial x^{\tau}} V^{\tau} + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\sigma\kappa} V^{\tau} \delta^{\kappa}_{\tau} &= \\
\frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\lambda}} \frac{\partial x'^{\lambda}}{\partial x^{\tau}} V^{\tau} + \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\sigma\kappa} V^{\kappa}&.
\end{aligned} \tag{A.2.3}$$

Using the following relation,

$$\frac{\partial}{\partial x^\rho}(\delta^\rho_\tau) = 0$$
$$\Rightarrow$$
$$\frac{\partial}{\partial x^\tau}\left(\frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial x'^\mu}{\partial x^\rho}\right) = 0$$
$$\Rightarrow$$
$$\frac{\partial x'^\lambda}{\partial x^\tau}\frac{\partial^2 x^\rho}{\partial x'^\lambda \partial x'^\nu}\frac{\partial x'^\mu}{\partial x^\rho} + \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial^2 x'^\mu}{\partial x^\tau \partial x^\rho} = 0$$
$$\Rightarrow$$
$$\frac{\partial x'^\lambda}{\partial x^\tau}\frac{\partial^2 x^\rho}{\partial x'^\lambda \partial x'^\nu}\frac{\partial x'^\mu}{\partial x^\rho} = -\frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial^2 x'^\mu}{\partial x^\tau \partial x^\rho} \tag{A.2.4}$$

we arrive at

$$\Gamma'^\mu_{\nu\lambda}V'^\lambda = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}\Gamma^\rho_{\sigma\kappa}V^\kappa - \frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial^2 x'^\mu}{\partial x^\tau \partial x^\rho}V^\tau \tag{A.2.5}$$

where the last term is exactly the same as in the transformation for $\partial'_\nu V'^\mu$.

A.3 Deriving the Affine Connection From the Metric

As discussed in chapter 2.3.2, the affine connection does not transform as a tensor, but coordinate transformation result as in equation (2.3.6). To show this we begin by observing that a derivate of the metric becomes

$$\partial_\mu g_{\nu\lambda} = \frac{\partial}{\partial x^\mu}\left(\frac{\partial \xi^\alpha}{\partial x^\nu}\frac{\partial \xi^\beta}{\partial x^\lambda}\eta_{\alpha\beta}\right) = \frac{\partial^2 \xi^\alpha}{\partial x^\nu \partial x^\mu}\frac{\partial \xi^\beta}{\partial x^\lambda}\eta_{\alpha\beta} + \frac{\partial \xi^\alpha}{\partial x^\nu}\frac{\partial^2 \xi^\beta}{\partial x^\lambda \partial x^\mu}\eta_{\alpha\beta}. \tag{A.3.1}$$

This shows some similarities with the affine connection, with some modification. One thing to note is the following relation

$$\frac{\partial \xi^\beta}{\partial x^\mu}\Gamma^\mu_{\nu\lambda} = \frac{\partial \xi^\beta}{\partial x^\mu}\frac{\partial x^\mu}{\partial \xi^\alpha}\frac{\partial^2 \xi^\alpha}{\partial x^\nu \partial x^\lambda} = \delta^\beta_\alpha \frac{\partial^2 \xi^\alpha}{\partial x^\nu \partial x^\lambda} = \frac{\partial^2 \xi^\beta}{\partial x^\nu \partial x^\lambda}, \tag{A.3.2}$$

where, by replacing these terms in equation (A.3.1) for ξ^β and ξ^α respectively, the relation can be written as

$$\partial_\mu g_{\nu\lambda} = \frac{\partial \xi^\alpha}{\partial x^\rho}\Gamma^\rho_{\nu\mu}\frac{\partial \xi^\beta}{\partial x^\lambda}\eta_{\alpha\beta} + \frac{\partial \xi^\beta}{\partial x^\rho}\Gamma^\rho_{\lambda\mu}\frac{\partial \xi^\alpha}{\partial x^\nu}\eta_{\alpha\beta} = \Gamma^\rho_{\nu\mu}g_{\rho\lambda} + \Gamma^\rho_{\lambda\mu}g_{\rho\nu}. \tag{A.3.3}$$

From here terms are taken as they are in equation (A.3.3), then added with the indices switched $\mu \leftrightarrow \lambda$, to lastly be subtracted where the indices have been switched like $\mu \leftrightarrow \nu$. This, with the affine connection's symmetric properties, gives

$$\begin{aligned}
&\partial_\mu g_{\nu\lambda} + \partial_\lambda g_{\mu\nu} - \partial_\lambda g_{\nu\mu} \\
&= \\
\Gamma^\rho_{\nu\mu}g_{\rho\lambda} + \Gamma^\rho_{\lambda\mu}g_{\rho\nu} + \Gamma^\rho_{\nu\lambda}g_{\rho\mu} + \Gamma^\rho_{\mu\lambda}g_{\rho\nu} - \Gamma^\rho_{\mu\nu}g_{\rho\lambda} - \Gamma^\rho_{\lambda\nu}g_{\rho\mu}& \\
&= \{\Gamma^\rho_{\nu\mu} = \Gamma^\rho_{\mu\nu}\} = \\
2\Gamma^\rho_{\mu\lambda}g_{\rho\nu}.&
\end{aligned} \tag{A.3.4}$$

This leads us to this very important results, which relates the affine connection to the metric according to:

$$\Gamma^\rho_{\mu\lambda} = \frac{1}{2}g^{\rho\nu}\left(\partial_\mu g_{\nu\lambda} + \partial_\lambda g_{\mu\nu} - \partial_\nu g_{\lambda\mu}\right). \tag{A.3.5}$$

A.4 Geodesics

Here the concept of a geodesic, as presented in section 2.3.3, will be further discussed, as well as explicitly calculated. Initially we observe what would constitute a "straight line" on a curved manifold. In flat space a straight line fulfills the relation

$$\frac{\mathrm{d}^2\xi^\alpha}{\mathrm{d}\tau^2} = 0, \tag{A.4.1}$$

where ξ^α are the Cartesian coordinates of the curve, dependent on a parameter τ. Expressed in the general coordinates of a differentiable curved manifold, we have that

$$\begin{aligned}
\frac{\mathrm{d}^2\xi^\alpha}{\mathrm{d}\tau^2} &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}\xi^\alpha}{\mathrm{d}\tau}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right) \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial\xi^\alpha}{\partial x^\mu}\right)\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} + \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} \\
&= \left(\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial\xi^\alpha}{\partial x^\nu\partial x^\mu}\right)\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} + \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}.
\end{aligned} \tag{A.4.2}$$

Analogous to equation (A.4.1) in flat space, this expression should be equal to zero. By multiplying both sides with $\frac{\partial x^\sigma}{\partial \xi^\alpha}$ and using the fact that $\frac{\partial x^\sigma}{\partial \xi^\alpha}\frac{\partial \xi^\alpha}{\partial x^\mu} = \delta^\sigma_\mu$, we have that

$$\begin{aligned}
\left(\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial \xi^\alpha}{\partial x^\nu \partial x^\mu}\right)\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} + \frac{\partial \xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} &= 0 \\
&\Rightarrow \\
\frac{\partial x^\sigma}{\partial \xi^\alpha}\left(\left(\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial \xi^\alpha}{\partial x^\nu \partial x^\mu}\right)\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} + \frac{\partial \xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2}\right) &= 0 \\
&\Rightarrow \\
\frac{\partial x^\sigma}{\partial \xi^\alpha}\frac{\partial \xi^\alpha}{\partial x^\nu \partial x^\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} + \delta^\sigma_\mu \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} &= 0 \\
&\Rightarrow \\
\frac{\mathrm{d}^2 x^\sigma}{\mathrm{d}\tau^2} + \frac{\partial x^\sigma}{\partial \xi^\alpha}\frac{\partial \xi^\alpha}{\partial x^\nu \partial x^\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} &= 0 \\
&\Rightarrow \\
\frac{\mathrm{d}^2 x^\sigma}{\mathrm{d}\tau^2} + \Gamma^\sigma_{\nu\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} &= 0. \qquad \text{(A.4.3)}
\end{aligned}$$

B

Newtonian Limit Approximation

The Newtonian limit approximation states what requirements the metric must have far from the source to not disagree with Newtonian gravity. In the Newtonian limit the particle that is affected by the field must be moving slowly, the field must be weak and it also has to be static with respect to time. These three properties will be strictly defined later on and are used in sections 2.4, 2.5 and throughout chapter 4.

To start with, we know that for a free particle that is not affected by gravity, the equations of motion is just as for a straight line,

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = 0. \tag{B.0.1}$$

In the presence of gravity, space will be curved and the equations of motion for a free particle will be that of a straight line in curved spacetime. This is then represented by the geodesic equation (2.3.15),

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\rho\sigma} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} = 0. \tag{B.0.2}$$

As the particle is supposed to be moving slowly the space coordinates must be changing much slower than the time coordinate, therefore,

$$\left| \frac{\mathrm{d}x^i}{\mathrm{d}\tau} \right| \ll \left| \frac{\mathrm{d}t}{\mathrm{d}\tau} \right| \tag{B.0.3}$$

where x^i are the space coordinates. From this the geodesic equation can be approximated by,

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{tt} \left(\frac{\mathrm{d}t}{\mathrm{d}\tau} \right)^2 = 0. \tag{B.0.4}$$

The relevant components of the connection can be found from,

$$\Gamma^{\mu}_{tt} = \frac{1}{2} g^{\mu\nu} (\partial_t g_{\nu t} + \partial_t g_{t\nu} - \partial_\nu g_{tt}). \tag{B.0.5}$$

Since the field is static the metric cannot be time dependent and therefore the terms with time derivatives in the connection disappear. These components of the connection result in,

$$\Gamma^{\mu}_{tt} = -\frac{1}{2} g^{\mu\nu} \partial_\nu g_{tt}. \tag{B.0.6}$$

The last requirement for the Newtonian limit was that the field would be weak. This is the same as saying that the metric is equal to the flat metric, plus a small perturbation,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad |h_{\mu\nu}| \ll 1. \tag{B.0.7}$$

From the fact that $g_{\mu\nu} g^{\mu\sigma} = \delta^{\sigma}_{\nu}$, see equation (2.3.2), we then find that $g^{\mu\nu} = \eta^{\mu\nu} - h_{\sigma\rho}\eta^{\mu\sigma}\eta^{\nu\rho}$. Inserting this into equation (B.0.6) and only keeping the first order terms of the perturbation $h_{\mu\nu}$ we find that,

$$\Gamma^{\mu}_{tt} = -\frac{1}{2} \eta^{\mu\nu} \partial_\nu h_{tt}. \tag{B.0.8}$$

Continuing with the equations of motion, (B.0.4), these become

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = \frac{1}{2} \eta^{\mu\nu} \partial_\nu h_{tt} \left(\frac{\mathrm{d}t}{\mathrm{d}\tau} \right)^2. \tag{B.0.9}$$

To be able to compare with Newtonian gravity, we look at the equations for the time and space coordinates separately. These are then,

$$\begin{cases} \dfrac{\mathrm{d}^2 t}{\mathrm{d}\tau^2} &= 0 \\ \dfrac{\mathrm{d}^2 x^i}{\mathrm{d}\tau^2} &= \dfrac{1}{2} \partial_i h_{tt} \left(\dfrac{\mathrm{d}t}{\mathrm{d}\tau} \right)^2 \end{cases} \tag{B.0.10}$$

The first equation, which states that $\frac{\mathrm{d}t}{\mathrm{d}\tau}$ is constant, is consistent with Newton. Dividing the second equation with $\left(\dfrac{\mathrm{d}t}{\mathrm{d}\tau} \right)^2$ leeds to,

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = \frac{1}{2} \partial_i h_{tt} \tag{B.0.11}$$

and we see that this very much looks like the Newtonian equation of motion for gravity, which says that the acceleration of a particle is equal to the divergence of the gravitational potential, $\vec{a} = -\nabla V$. To arrive at this, the time component of the small perturbation must be $h_{tt} = -2V$ and therefore the time component of the metric will be $g_{tt} = -(1 + 2V)$. The gravitational potential in Newtonian theory is $V = -\frac{GM}{r}$ so in the Newtonian limit g_{tt} must asymptotically approach $-(1 - \frac{2GM}{r})$ when $r \to \infty$.

C

Manifolds

Manifolds are mathematical constructs used to describe the curved spacetime of general relativity, as presented in chapter 2. In short, manifolds describe complicated geometries by a collection of overlapping coordinate systems. To give a somewhat more detailed definition we need to start with some basic concepts: **map**, **chart**, **atlas**.

Map

A map is a relationship between two sets for which the each element in one set M is assigned one element in another set N. The map ϕ is a function on the form

$$\phi : M \rightarrow N. \tag{C.0.1}$$

A map has an inversion ϕ^{-1} if it is both **onto** and **one-to-one**, meaning that the map assigns each element in N and does so for only one element in M. This is analogous to an invertible function in one-dimensional calculus.

The two sets are **diffeomorphic** if there exists a map that satisfies that it and its inversion are infinitely differentiable, also known as smooth.

Chart

A chart or coordinate system ϕ is an invertible map between an open subset U of a set M to an open set of coordinates in $\mathbb{R}^n$.

$$\phi : U \rightarrow \mathbb{R}^n. \tag{C.0.2}$$

The trouble with complicated geometries, such as spacetime in the vicinity of a black hole, is that it is not always possible to describe the whole geometry, the whole set, with a single coordinate system. This necessitates the use of overlapping charts to describe the whole geometry. A simple one-dimensional example would be describing the geometry of a circle. Using a single chart, where each point on the circle is assigned an angle coordinate, one wouldn't be able to represent all of the points on the circle without having an overlap (given that the chart's coordinates are an open set) and thus losing the criterion of inversion.

Atlas

An atlas is an indexed collection of charts ϕ_i on subsets U_i of the set M that fulfills two criterion.

1. Firstly, the charts cover the totality of M, meaning that the subgroups U_i must overlap.

2. Secondly, the charts are smoothly sown together, meaning the overlaps behave in a reasonable manner. More specifically, if two subsets U_a and U_b overlap in a subset U_{ab}, then the chart $\phi_b \circ \phi_a^{-1}$ takes a point from $\phi_a(U_{ab})$ to $\phi_b(U_{ab})$.

Manifold

Finally, a manifold is a set M with a maximal atlas, meaning one that contains every possible chart. The point of defining it as with a maximal atlas is so that equivalent spaces described with different coordinate systems don't count as different manifolds.

An important consequence of the definition is that a manifold always looks flat locally. This means that geometries that cannot be regarded as locally flat everywhere are not manifolds, such as a non-smooth curve.

The precise definition is not very important in this matter. The important thing to remember from this section is as follows. The description of spacetime of general relativity is a manifold. Sometimes one cannot describe the complete geometry of a curved spacetime with a single coordinate system, instead one has to use a collection of overlapping coordinate systems.

D

Variation of a Determinant

The following calculations är used for the variation of the Schwarzschild action in section 2.4.3. The variation of a determinant is given by

$$\delta(\mathrm{Det}\, A) = \mathrm{Det}\, A\, \mathrm{Tr}(A^{-1}\delta A). \tag{D.0.1}$$

To show this, we start by writing out the variation:

$$\begin{aligned}
\delta(\ln \mathrm{Det}\, A) &= \ln \mathrm{Det}(A + \delta A) - \ln \mathrm{Det}\, A \\
&= \ln \frac{\mathrm{Det}(A + \delta A)}{\mathrm{Det}\, A} \\
&= \ln \mathrm{Det}\, A^{-1}(A + \delta A) \\
&= \ln \mathrm{Det}(1 + A^{-1}\delta A).
\end{aligned} \tag{D.0.2}$$

Recall from linear algebra that the trace of a matrix equals the sum of it's eigenvalues, the determinant of a matrix equals the product of it's eigenvalues, and that the exponential of a matrix has the exponentials of the eigenvalues as eigenvalues, thus

$$\mathrm{Det}(\exp A) = \exp(\mathrm{Tr}\, A). \tag{D.0.3}$$

We can also note that $(1 + A^{-1}\delta A)$ is the first order approximation of the expression $\exp(A^{-1}\delta A)$, so (D.0.2) can be written as

$$\begin{aligned}
\delta(\ln \mathrm{Det}\, A) &\to \ln \mathrm{Det} \exp(A^{-1}\delta A) \\
&= \ln \exp(\mathrm{Tr}(A^{-1}\delta A)) \\
&= \mathrm{Tr}(A^{-1}\delta A).
\end{aligned} \tag{D.0.4}$$

By the chain rule one also acquire

$$\delta(\ln \mathrm{Det}\, A) = \frac{1}{\mathrm{Det}\, A} \delta\, \mathrm{Det}\, A. \tag{D.0.5}$$

Together the result is

$$\delta(\mathrm{Det}\, A) = \mathrm{Det}\, A\, \mathrm{Tr}(A^{-1} \delta A), \tag{D.0.6}$$

which trivially holds also for the absolute value of the determinant.

E

Geodesic Equation

We want to derive the equation (4.2.42) from the variation of the action,

$$S'_G = 4\pi \int_{\tau(0)}^{0} \mathrm{d}\tau \gamma_{ij}(\Phi)\partial_\tau \Phi^i \partial_\tau \Phi^j. \tag{E.0.1}$$

Variating with respect to Φ results in,

$$\delta_\Phi S'_G = 4\pi \int_{\tau(0)}^{0} \mathrm{d}\tau(\delta_\Phi(\gamma_{ij}(\Phi))\partial_\tau \Phi^i \partial_\tau \Phi^j + 2\gamma_{ij}(\Phi)\partial_\tau \Phi^i \partial_\tau \delta\Phi^j), \tag{E.0.2}$$

see equation (4.2.41). For the first term we use the fact that the derivative of the metric with respect to coordinates is given as,

$$\begin{aligned} \partial_\mu g_{\nu\lambda} &= \Gamma^\rho_{\nu\mu} g_{\rho\lambda} + \Gamma^\rho_{\lambda\mu} g_{\rho\nu} \\ &\Rightarrow \\ \delta g_{\nu\lambda} &= (\Gamma^\rho_{\nu\mu} g_{\rho\lambda} + \Gamma^\rho_{\lambda\mu} g_{\rho\nu})\delta x^\mu. \end{aligned} \tag{E.0.3}$$

Here we remember that the coordinates, metric and affine connection are in regards to the coset space and therefore we write this as,

$$\delta_\Phi \gamma_{ij} = (\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\delta\Phi^k, \tag{E.0.4}$$

where Υ^k_{ij} is the affine connection for the group. Using this in equation (E.0.2), along with partial integration, the second term equates,

$$\begin{aligned} \delta_\Phi S'_G &= 4\pi \int_{\tau(0)}^{0} \mathrm{d}\tau(\delta_\Phi(\gamma_{ij}(\Phi))\partial_\tau \Phi^i \partial_\tau \Phi^j + 2\gamma_{ij}(\Phi)\partial_\tau \Phi^i \partial_\tau \delta\Phi^j) \\ &= 4\pi \int_{\tau(0)}^{0} \mathrm{d}\tau((\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\delta\Phi^k \partial_\tau \Phi^i \partial_\tau \Phi^j - 2\partial_\tau(\gamma_{ij}(\Phi)\partial_\tau \Phi^i)\delta\Phi^j) \end{aligned} \tag{E.0.5}$$

Because the indices k and j for the first and second term are only dummy indices we can change these and pull out the factor $\delta\Phi^k$ from both terms,

$$\delta S'_G = 4\pi \int_{\tau(0)}^{0} \mathrm{d}\tau \delta\Phi^k ((\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\partial_\tau\Phi^i\partial_\tau\Phi^j - 2\partial_\tau(\gamma_{ik}(\Phi)\partial_\tau\Phi^i)) \tag{E.0.6}$$

If this is going to be zero for all $\delta\Phi^k$ then we must have that

$$\begin{aligned}
(\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\partial_\tau\Phi^i\partial_\tau\Phi^j - 2\partial_\tau(\gamma_{ik}(\Phi)\partial_\tau\Phi^i) &= 0 \\
&\Rightarrow \\
2\partial_\tau(\gamma_{ik}(\Phi)\partial_\tau\Phi^i) - (\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0 \\
&\Rightarrow \\
2\partial_\tau\gamma_{ik}(\Phi)\partial_\tau\Phi^i + 2\gamma_{ik}(\Phi)\partial_\tau^2\Phi^i - (\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0 \\
&\Rightarrow \\
2\left(\frac{\partial\gamma_{ik}}{\partial\Phi^j}\partial_\tau\Phi^j\right)\partial_\tau\Phi^i + 2\gamma_{ik}(\Phi)\partial_\tau^2\Phi^i - (\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0 \\
&\Rightarrow \\
2(\Upsilon^m_{ij}\gamma_{mk} + \Upsilon^m_{kj}\gamma_{mi})\partial_\tau\Phi^j\partial_\tau\Phi^i + 2\gamma_{ik}(\Phi)\partial_\tau^2\Phi^i - (\Upsilon^m_{ik}\gamma_{mj} + \Upsilon^m_{jk}\gamma_{mi})\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0 \\
&\Rightarrow \\
2\gamma_{ik}(\Phi)\partial_\tau^2\Phi^i + (2\Upsilon^m_{ij}\gamma_{mk} + 2\Upsilon^m_{kj}\gamma_{mi} - \Upsilon^m_{ik}\gamma_{mj} - \Upsilon^m_{jk}\gamma_{mi})\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0
\end{aligned} \tag{E.0.7}$$

To calculate the sum of the affine connections we first notice that it is symmetric in its two lower indices, which can be used to sum the second and fourth term. This means that,

$$2\gamma_{ik}(\Phi)\partial_\tau^2\Phi^i + (2\Upsilon^m_{ij}\gamma_{mk} + \Upsilon^m_{kj}\gamma_{mi} - \Upsilon^m_{ik}\gamma_{mj})\partial_\tau\Phi^i\partial_\tau\Phi^j = 0 \tag{E.0.8}$$

The sum of the two last connections are antisymmetric in i and j, so they will not contribute because they are contracted with the symmetric tensor $\partial_\tau\Phi^i\partial_\tau\Phi^j$. This gives us,

$$\begin{aligned}
2\gamma_{ik}(\Phi)\partial_\tau^2\Phi^i + (2\Upsilon^m_{ij}\gamma_{mk} + \Upsilon^m_{kj}\gamma_{mi} - \Upsilon^m_{ik}\gamma_{mj})\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0 \\
&\Rightarrow \\
2\gamma_{ik}(\Phi)\partial_\tau^2\Phi^i + 2\Upsilon^m_{ij}\gamma_{mk}\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0 \\
&\Rightarrow \\
\gamma_{mk}(\Phi)(\partial_\tau^2\Phi^m + \Upsilon^m_{ij}\partial_\tau\Phi^i\partial_\tau\Phi^j) &= 0 \\
&\Rightarrow \\
\partial_\tau^2\Phi^m + \Upsilon^m_{ij}\partial_\tau\Phi^i\partial_\tau\Phi^j &= 0
\end{aligned} \tag{E.0.9}$$

This is then the geodesic equation on the coset space.

F

Equivalencies of Equations

We want to prove the following:

$$D^\mu P_\mu = 0 \Rightarrow \nabla^\mu(M^{-1}\partial_\mu M), \tag{F.0.1}$$

from section 4.3, where $D^\mu P_\mu = \nabla^\mu + [Q^\mu, P_\mu]$. We start by expanding the commutator and writing A_μ for $\mathcal{V}^{-1}\partial_\mu\mathcal{V}$,

$$\begin{aligned}
&g^{\mu\nu}[P_\nu, Q_\mu] = \qquad\qquad (\text{F.0.2})\\
=&g^{\mu\nu}(P_\nu Q_\mu - Q_\mu P_\nu)\\
=&\frac{1}{4}g^{\mu\nu}\left((A_\nu + A_\nu^{\mathcal{T}})(A_\mu - A_\mu^{\mathcal{T}}) - (A_\mu - A_\mu^{\mathcal{T}})(A_\nu + A_\nu^{\mathcal{T}})\right)\\
=&\frac{1}{4}g^{\mu\nu}\left((A_\nu + A_\nu^{\mathcal{T}})(A_\mu - A_\mu^{\mathcal{T}}) - (A_\nu - A_\nu^{\mathcal{T}})(A_\mu + A_\mu^{\mathcal{T}})\right)\\
=&\frac{1}{4}g^{\mu\nu}\left(A_\nu A_\mu + A_\nu^{\mathcal{T}} A_\mu - A_\nu A_\mu^{\mathcal{T}} - A_\nu^{\mathcal{T}} A_\mu^{\mathcal{T}} - A_\nu A_\mu - A_\nu A_\mu^{\mathcal{T}} + A_\nu^{\mathcal{T}} A_\mu + A_\nu^{\mathcal{T}} A_\mu^{\mathcal{T}}\right)\\
=&\frac{1}{2}g^{\mu\nu}\left(A_\nu^{\mathcal{T}} A_\mu - A_\nu A_\mu^{\mathcal{T}}\right) \qquad\qquad (\text{F.0.3})
\end{aligned}$$

Using this in the original equation leads to

$$\begin{aligned}
0 &= \nabla^\mu P_\mu + [Q^\mu, P_\mu]\\
&= \nabla_\mu P^\mu - [P^\mu, Q_\mu]\\
&= \nabla_\mu (g^{\mu\nu} P_\nu) - g^{\mu\nu}[P_\nu, Q_\mu]\\
&= g^{\mu\nu}\partial_\mu P_\nu - \Gamma^\nu_{\mu\lambda} g^{\mu\lambda} P_\nu - g^{\mu\nu}[P_\nu, Q_\mu]\\
&= g^{\mu\nu}\partial_\mu P_\nu - \Gamma^\nu_{\mu\lambda} g^{\mu\lambda} P_\nu - \frac{1}{2}g^{\mu\nu}\left(A_\nu^{\mathcal{T}} A_\mu - A_\nu A_\mu^{\mathcal{T}}\right).
\end{aligned} \tag{F.0.4}$$

Comparing this to the other end of the relation that we wish to prove this gives

$$\begin{aligned}
\nabla^\mu(M^{-1}\partial_\mu M) &= \nabla_\mu(g^{\mu\nu}M^{-1}\partial_\nu M) \\
&= \nabla_\mu\left(g^{\mu\nu}(\mathcal{V}^{-1})^{\mathcal{T}}2P_\nu\mathcal{V}^{\mathcal{T}}\right) \\
&= g^{\mu\nu}\partial_\mu\left((\mathcal{V}^{-1})^{\mathcal{T}}2P_\nu\mathcal{V}^{\mathcal{T}}\right) - \Gamma^\nu_{\mu\lambda}g^{\mu\lambda}\left((\mathcal{V}^{-1})^{\mathcal{T}}2P_\nu\mathcal{V}^{\mathcal{T}}\right) \\
&= g^{\mu\nu}(-\mathcal{V}^{-1}\partial_\mu\mathcal{V}\mathcal{V}^{-1})^{\mathcal{T}}2P_\nu\mathcal{V}^{\mathcal{T}} + g^{\mu\nu}(\mathcal{V}^{-1})^{\mathcal{T}}2\partial_\mu P_\nu\mathcal{V}^{\mathcal{T}} \\
&+ g^{\mu\nu}(\mathcal{V}^{-1})^{\mathcal{T}}2P_\nu\partial_\mu\mathcal{V}^{\mathcal{T}} - \Gamma^\nu_{\mu\lambda}g^{\mu\lambda}\left((\mathcal{V}^{-1})^{\mathcal{T}}2P_\nu\mathcal{V}^{\mathcal{T}}\right). \qquad \text{(F.0.5)}
\end{aligned}$$

Multiply by $(\mathcal{V}^{-1})^{\mathcal{T}}$ from the right and $(\mathcal{V}^{-1})^{\mathcal{T}}$ from the left, and divide by 2

$$\begin{aligned}
&\frac{1}{2}(\mathcal{V}^{-1})^{\mathcal{T}}\nabla^\mu(M^{-1}\partial_\mu M)(\mathcal{V}^{-1})^{\mathcal{T}} = \\
&= -g^{\mu\nu}(\mathcal{V}^{-1}\partial_\mu\mathcal{V})^{\mathcal{T}}P_\nu + g^{\mu\nu}\partial_\mu P_\nu + g^{\mu\nu}P_\nu\partial_\mu\mathcal{V}^{\mathcal{T}}(\mathcal{V}^{-1})^{\mathcal{T}} - \Gamma^\nu_{\mu\lambda}g^{\mu\lambda}P_\nu \\
&= g^{\mu\nu}\partial_\mu P_\nu - \Gamma^\nu_{\mu\lambda}g^{\mu\lambda}P_\nu + g^{\mu\nu}\left(-(\mathcal{V}^{-1}\partial_\mu\mathcal{V})^{\mathcal{T}}P_\nu + P_\nu(\mathcal{V}^{-1}\partial_\mu\mathcal{V})^{\mathcal{T}}\right) \\
&= g^{\mu\nu}\partial_\mu P_\nu - \Gamma^\nu_{\mu\lambda}g^{\mu\lambda}P_\nu + \frac{1}{2}g^{\mu\nu}\left(-A_\mu^{\mathcal{T}}(A_\nu + A_\nu^{\mathcal{T}}) + (A_\nu + A_\nu^{\mathcal{T}})A_\mu^{\mathcal{T}}\right) \\
&= g^{\mu\nu}\partial_\mu P_\nu - \Gamma^\nu_{\mu\lambda}g^{\mu\lambda}P_\nu + \frac{1}{2}g^{\mu\nu}\left(-A_\nu^{\mathcal{T}}(A_\mu + A_\mu^{\mathcal{T}}) + (A_\nu + A_\nu^{\mathcal{T}})A_\mu^{\mathcal{T}}\right) \\
&= g^{\mu\nu}\partial_\mu P_\nu - \Gamma^\nu_{\mu\lambda}g^{\mu\lambda}P_\nu - \frac{1}{2}g^{\mu\nu}\left(A_\nu^{\mathcal{T}}A_\mu - A_\nu A_\mu^{\mathcal{T}}\right). \qquad \text{(F.0.6)}
\end{aligned}$$

Comparing these to results we arrive at

$$\nabla^\mu(M^{-1}\partial_\mu M) = 2\mathcal{V}^{\mathcal{T}}(D_\mu P^\mu)(\mathcal{V}^{-1})^{\mathcal{T}} \qquad \text{(F.0.7)}$$

and therefore

$$D^\mu P_\mu = 0 \Rightarrow \nabla^\mu(M^{-1}\partial_\mu M). \qquad \text{(F.0.8)}$$

G

Exponential of the Generators of SL(2,$\mathbb{R}$)

In this appendix we calculate

$$\exp\Big(\tau(\alpha h + \beta e + \gamma f)\Big), \tag{G.0.1}$$

for arbitrary α, β and γ and where h, e and f are given as

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{G.0.2}$$

The matrix in the exponent is therefore,

$$A = \tau(\alpha h + \beta e + \gamma f) = \begin{pmatrix} \alpha\tau & \beta\tau \\ \gamma\tau & -\alpha\tau \end{pmatrix}. \tag{G.0.3}$$

To be able to perform the exponential of this matrix we want to diagonalize it and for this we need its eigenvalues and eigenvectors. These are

$$\lambda_{1,2} = \pm\sqrt{\alpha^2 + \gamma\beta}\,\tau \quad v_{1,2} = \begin{pmatrix} \frac{\alpha\pm\sqrt{\alpha^2+\gamma\beta}}{\gamma} \\ 1 \end{pmatrix}. \tag{G.0.4}$$

The matrix is diagonalizable for $E = \alpha^2 + \gamma\beta > 0$ and as this quantity is found to be the rest mass energy of the black hole in section 4.4, it is valid to assume that this is greater than zero. With the eigenvalues and eigenvectors the matrix can be written as,

$$\begin{aligned} A &= \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} v_1 & v_2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{\alpha+\sqrt{E}}{\gamma} & \frac{\alpha-\sqrt{E}}{\gamma} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{E}\tau & 0 \\ 0 & -\sqrt{E}\tau \end{pmatrix} \begin{pmatrix} \frac{\gamma}{2\sqrt{E}} & \frac{\sqrt{E}-\alpha}{2\sqrt{E}} \\ -\frac{\gamma}{2\sqrt{E}} & \frac{\sqrt{E}+\alpha}{2\sqrt{E}} \end{pmatrix}. \end{aligned} \tag{G.0.5}$$

The exponential of the matrix then becomes,

$$\begin{aligned} \exp(A) &= \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} \exp(\lambda_1) & 0 \\ 0 & \exp(\lambda_2) \end{pmatrix} \begin{pmatrix} v_1 & v_2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{\alpha+\sqrt{E}}{\gamma} & \frac{\alpha-\sqrt{E}}{\gamma} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \exp(\sqrt{E}\tau) & 0 \\ 0 & \exp(-\sqrt{E}\tau) \end{pmatrix} \begin{pmatrix} \frac{\gamma}{2\sqrt{E}} & \frac{\sqrt{E}-\alpha}{2\sqrt{E}} \\ -\frac{\gamma}{2\sqrt{E}} & \frac{\sqrt{E}+\alpha}{2\sqrt{E}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\alpha+\sqrt{E}}{\gamma} & \frac{\alpha-\sqrt{E}}{\gamma} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{\gamma}{2\sqrt{E}}\exp(\sqrt{E}\tau) & \frac{\sqrt{E}-\alpha}{2\sqrt{E}}\exp(\sqrt{E}\tau) \\ -\frac{\gamma}{2\sqrt{E}}\exp(-\sqrt{E}\tau) & \frac{\sqrt{E}+\alpha}{2\sqrt{E}}\exp(-\sqrt{E}\tau) \end{pmatrix} \\ &= \begin{pmatrix} \cosh(\sqrt{E}\tau) + \frac{\alpha}{\sqrt{E}}\sinh(\sqrt{E}\tau) & \frac{\beta}{\sqrt{E}}\sinh(\sqrt{E}\tau) \\ \frac{\gamma}{\sqrt{E}}\sinh(\sqrt{E}\tau) & \cosh(\sqrt{E}\tau) - \frac{\alpha}{\sqrt{E}}\sinh(\sqrt{E}\tau) \end{pmatrix}. \end{aligned} \tag{G.0.6}$$