university of groningen / faculty of science and engineering

Master Project Mathematics and Physics

Chaotic scattering in relativistic N-center problems

Martijn Kluitenberg
s2745542

Abstract. We study relativistic systems with N static black holes in the Einstein-Maxwell-dilaton theory of gravitation. There is an arbitrary coupling constant a for the dilaton. Unlike in classical mechanics, the relativistic 2-center problem seems to be chaotic for values of a between 0 and $\sqrt{3}$. For this latter value of a, the system is integrable, like in the classical case. We prove rigorously - using methods from potential scattering - that the motion of light in problems with 3 or more centers is chaotic for a discrete set of a's between 1 and $\sqrt{3}$. The value $a = 1$ is a turning point, where the qualitative features of the black holes change.

supervised by
Dr. M. Seri
Prof. Dr. D. Roest

Contents

1 Introduction

In this thesis, we will study versions of the N-center problem that are compatible with relativity. The classical two-center problem has been studied since the days of Jacobi, as it provides a good approximation of the three body problem. We can view it as a limit where one light body moves in the gravitational (or electromagnetic) field of the heavier ones. In this approximation, we have that the heaviest two bodies are stationary. This can be a realistic assumption on the time scale of one orbit of the light particle.

Having a relativistic version of this problem is useful for studying black hole binaries. The study of these objects has become important in recent years, as it is related to the measurement of gravitational waves. The LIGO / VIRGO collaborations have measured gravitational waves emitted from a pair of inspiralling black holes [1]. For these measurements, it is very important to have accurate templates to match the data to. Hence, getting a better understanding of these binaries is important in current research.

Modelling a pair of inspiraling black holes is mathematically very complicated. It's difficult to imagine what the full time-dependent metric would even look like. Hence, most things we know about the black hole binary system are through perturbation theory. In particular, there are the post-Newtonian and post-Minskowskian approximation, where we solve for the Hamiltonian order by order in v/c or G respectively.

We will consider a different approach. In order to make the two-center problem relativistic, we need to get rid of the *action at a distance*. Because of special relativity - if the sun would suddenly disappear - the earth would stay in its orbit for roughly 8 minutes before flying off to infinity, since this is how long light takes to travel from sun to earth. This is a fundamental problem with Newton's $1/r^2$ force theory, in which the force between two massive objects only depends on the distance between them.

To remedy this problem, we introduce a mediator particle, which travels at finite speed. Such a particle should have integer spin, like all force carriers in theoretical physics. The most common choices are spin 0 (dilaton), spin 1 (photon) and spin 2 (graviton). We write down an action in terms of the corresponding fields ϕ, A_μ and $g_{\mu\nu}$, and then look for solutions of the corresponding equations of motion. For example, the familiar action

$$S = \int \mathrm{d}^4x \sqrt{-g} \left(\frac{c^3}{16\pi G} R - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right)$$

leads to the Einstein-Maxwell equations [2]. The general action, also including the dilaton, is given by

$$S = \int \mathrm{d}^4x \sqrt{-g} \left(R - \partial_\mu \phi \partial^\mu \phi - e^{-2a\phi} F_{\mu\nu} F^{\mu\nu} \right),$$

where a is a coupling constant.

There is a one-parameter family of *static* solutions to the field equations generated by the above action containing an arbitrary number of centers. These correspond to extremal black holes, which are held fixed by balancing the attractive spin 0 and spin 2 forces with the repulsive spin 1 force. The family is labelled by a parameter a [3].

Let us consider the motion of test particles in such a geometry with two centers. Without loss of generality, we can place them on the z-axis, with separation $2s$, i.e. at $(0, 0, \pm s)$. We will mainly be interested in the motion of light in the (x, z)-plane. Light is the simplest type of particle, as it only couples to the background metric, and not to any of the other forces. This will greatly simplify the analysis of the equations of motion.

There have been a number of studies on the motion of light in the planar case. All of these conclude - based on very different methods - that this is a chaotic system for $0 \leq a < \sqrt{3}$ [4, 3, 5] However, practically all of the known methods eventually fall back on numerics. We would like to use analytical tools to prove the existence of chaos in these relativistic 2-center problems. The simplest way to prove the existence of chaos would involve constructing a proper symbolic code for the system. Such codes have been established in the past [3, 5], but none of these are proven to be *conjugate* to a shift map. This step is crucial for mathematically establishing the chaotic dynamics as we will see in Section 3

Both the classical and relativistic 2-center problems can be cast into the Hamiltonian framework. For these systems, there is the notion of *integrability* [6]. If a Hamiltonian system is integrable, then there are enough constants of motion to constrain what can happen to a particle moving around the phase space. In particular, we can write down the solution to Hamilton's equations using a finite number of integrations, and there can be no chaos. It has been known for a long time that the classical 2-center problem is integrable. It's somewhat surprising that the relativistic version is not. However, it has been proven that for $a = \sqrt{3}$ the integrability returns [3]. It is thus interesting to understand more deeply what goes wrong for $0 \leq a < \sqrt{3}$.

Another interesting aspect to study is the *physics* of extremal black holes with dilaton coupling. General relativity - while being very simple in terms of the action - is a very difficult theory to do scattering computations in. There are alternative theories of gravity, such as $N = 8$ supergravity, which have many more symmetries. Exploiting these, we can guess formulas for the scattering amplitudes based almost exclusively on the symmetries. The model for the 2-center problem above is related to many limits of the full $N = 8$ theory [3]. For example, we will look at the Kaluza-Klein limit in Section 5. We can hence consider our relativistic problem with dilaton coupling as a toy model for $N = 8$. Looking at the properties of black holes in this

toy theory, such as their event horizons and temperature is interesting from a physical point of view.

In order to study chaos in our relativistic 2-center problem, we will use tools from potential scattering [7, 8, 9]. Our system is Hamiltonian, and it can be rewritten in the standard "kinetic + potential" form by using a time rescaling. Here, it is important that the flow generated by Hamilton's equations is complete, i.e. that we can follow any trajectory for an arbitrarily long time. We hence need to regularize the motion by blowing up the singularities. We show that this is at least possible for a discrete set of values of a between 1 and $\sqrt{3}$, which accumulate at 1.

What happens at the other values of a is an interesting question. However, many of the standard methods from the dynamical systems literature cannot easily be applied. Hence, we will not treat any of these cases in detail. From the numerics, it's clear enough that there is chaos for all values of a between 0 and $\sqrt{3}$. Understanding where the chaos comes from in these systems and quantifying the *amount* of chaos depending on a is a good subject for further research.

This thesis is structured as follows: In section 2, we will look at how to make the N-center problem consistent with relativity. For two or more centers, this problem will usually be chaotic. In section 3, we will define what chaos is formally. We will also look at the notion of integrability, which is the opposite of chaos in some sense.

In section 4, we will look at the mathematical theory of scattering. This will lead to the notion of scattering degree, which we will apply later to relativistic N-center problems. As an example, we will calculate the scattering degree for centrally symmetric short range potentials

Section 5 provides some physical background on several aspects of the project. We will look at Nordström's theory to get some intuition about scalar theories of gravitation. We will also look at thermodynamics for black holes, and Penrose diagrams. These tools will be applied to understand what kind of objects our relativistic centers are.

Finally, in section 6, everything will come together as we analyse motion of light near extremal black holes in Einstein-Maxwell-dilaton theories. We will look at the effect of the parameter a, and at chaotic motion in the system. We will also present some numerical work.

This work is submitted as a combined Master's thesis in mathematics and physics. Trying to separate one from the other is in general a bad idea, but for the sake of administration, I have made an attempt. Sections 3, 4 and 6.2-6.5 constitute the mathematical part, while sections 2, 5 and 6.1 concern the physics. I have not tried to make these parts independent from each other, because that would introduce far too many redundancies.

2 N-Center solutions in general relativity

The goal of this section is to review the minimum about of general relativity required to discuss how to generalize the N-center problem to this setting. Classically, it is possible to create a static configuration of charged massive particles by tuning the charge and mass in such a way that the gravitational attraction and the Coulomb repulsion cancel out. Solutions with similar properties exist in General Relativity (GR). To discuss how such solutions are constructed and interpreted, we need to recall how to couple electromagnetism and gravity. Moreover, we will recall how (uncharged) particles move under the influence of gravity. These two problems are similar, in the sense that both can be tackled by using a suitable action.

2.1 The Einstein-Hilbert action

One of the core principles in GR is that flat Minkowski spacetime is replaced by some four-dimensional manifold, which is equipped with a metric $g_{\mu\nu}$. Unlike in classical physics, this metric is not positive definite. Instead, temporal and spatial components have a different sign. In this work, I will stick to the convention $(-,+,+,+)$, so that the metric of flat spacetime is

$$\eta = \begin{pmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{pmatrix}.$$

Generally, this is also written as

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\,,$$

where $\mathrm{d}s$ refers to an element of arclength along a parametrized curve $\gamma : \lambda \mapsto (t(\lambda), x(\lambda), y(\lambda), z(\lambda))$. The length of such a curve in flat space is given by

$$L(\gamma) := \int \mathrm{d}\lambda \sqrt{g_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu} = \int \mathrm{d}\lambda \sqrt{-\left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)^2 + \left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\lambda}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}\lambda}\right)^2}. \tag{2.1}$$

In GR, the metric becomes a dynamical variable. We would like to calculate how it evolves. Like in classical physics, we can do this by defining an action $S = S[g_{\mu\nu}]$ which depends on the metric and its derivatives. We note that spacetime is a manifold with a metric, and thus we have access to a natural volume form $\sqrt{-g}\,\mathrm{d}^4x$. Here, the minus sign is due to the signature. Because of this, it makes intuitive sense to look for an action of the form [2]

$$S = \int_M \mathrm{d}^4x\,\sqrt{-g}f,$$

for some scalar function f defined on spacetime. Since the only object we have available is the metric, we should try to construct a scalar from it. The most logical choice is the scalar curvature R associated to $g_{\mu\nu}$. Indeed,

$$S = \int_M \mathrm{d}^4x\,\sqrt{-g}R \tag{2.2}$$

is the famous Einstein-Hilbert action. Varying S with respect to the metric gives the Einstein field equations in vacuum:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0. \tag{2.3}$$

Here, $G_{\mu\nu}$ is the Einstein tensor. As a side note, contracting equation (2.3) with the inverse metric $g^{\mu\nu}$ gives

$$g^{\mu\nu} R_{\mu\nu} - \frac{1}{2} R g^{\mu\nu} g_{\mu\nu} = 0 \implies R = 0.$$

This means that we can rewrite the Einstein field equations in the simpler form

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

To close off this subsection, let me make a remark on the dimensionality. Generally, the action should have the dimension of angular momentum. Looking back at (2.2) shows that it has the wrong dimension. Luckily, we can just put a constant in front to fix it. The conventional factor (in SI units) is

$$S = \frac{c^3}{16\pi G} \int_M \mathrm{d}^4 x \, \sqrt{-g} R.$$

As we will see later on, this factor controls how strongly matter couples to gravity. The pre-factor is also often written as M_P^2, i.e. the square of the Planck mass.

Of course, R is not the only Lorentz scalar that can be formed with the metric. There are also terms like $R^2, R_{\mu\nu}R^{\mu\nu}$ and $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$, which contain four derivatives instead of two. We can - in principle - incorporate these into the Einstein-Hilbert action. For dimensional reasons, they are associated with some characteristic length scale l [10, p. 709]:

$$S = M_P^2 \int \mathrm{d}^4 x \, \sqrt{-g} \left[R + l^2 \left(\alpha R^2 + \beta R_{\mu\nu} R^{\mu\nu} + \gamma R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} \right) + \dots \right]$$

The only length scale we have available is the Planck length l_P. When deriving the equations of motion for the gravitational field, the l_P^2 turns into $1/l_P^2$, which is a very small number. Hence, these terms are strongly suppressed, and we will ignore them in the rest of this work.

2.2 The electromagnetic field

One goal of this section is to understand how we can couple electromagnetism and general relativity. To this end, it makes sense to briefly review how to do classical electromagnetism in the context of fields. The fundamental object in Maxwell's theory is the four-potential A_μ, which has as components the scalar potential ϕ and the magnetic vector potential $\mathbf{A}$. Based on their transformation properties, we can figure out that A_μ are the components of a 1-form [2]

$$A = A_\mu \, \mathrm{d}x^\mu$$

defined on spacetime. Its exterior derivative

$$F = \mathrm{d}A = (\partial_\mu A_\nu - \partial_\nu A_\mu)\,\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$$

has components $F_{\mu\nu}$. This anti-symmetric tensor is called the field strength tensor. We know from the standard courses in electricity and magnetism that $F_{\mu\nu}$ contains information about the electric and magnetic fields. Indeed,

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}.$$

The description in terms of A_μ has some redundancies. Indeed, if we add the exterior derivative of some scalar function α to A, then the field strength tensor remains unchanged:

$$A \to A + \mathrm{d}\alpha \implies F = \mathrm{d}A \to \mathrm{d}(A + \mathrm{d}\alpha) = \mathrm{d}A + \cancel{\mathrm{d}^2\alpha} = F.$$

We would like the action of this theory to be invariant under such "gauge" transformations. The simplest choice would be to take

$$S = \int_M \mathrm{d}^4x (F \wedge F),$$

however, that would not be a good idea. The 4-form $F \wedge F = \mathrm{d}A \wedge \mathrm{d}A = \mathrm{d}(A \wedge \mathrm{d}A)$ is exact, hence its integral over spacetime vanishes by Stokes' theorem. Luckily, since we have a volume form, we can construct a different 2-form to wedge F with. The Hodge star $\star : \Omega^k(M) \to \Omega^{n-k}(M)$ gives a different action

$$S_{\text{Maxwell}} = -\frac{1}{2}\int_M \mathrm{d}^4x F \wedge (\star F), \tag{2.4}$$

which becomes

$$S_{\text{Maxwell}} = -\frac{1}{4}\int_M \mathrm{d}^4x\,\sqrt{-g}\,(g^{\mu\nu}g^{\rho\sigma}F_{\mu\rho}F_{\nu\sigma}) = -\frac{1}{4}\int_M \mathrm{d}^4x\,\sqrt{-g}\,(F^{\mu\nu}F_{\mu\nu}) \tag{2.5}$$

in local coordinates. Recalling what $F_{\mu\nu}$ looks like in terms of $\mathbf{E}$ and $\mathbf{B}$, this reduces to the more familiar looking

$$F^{\mu\nu}F_{\mu\nu} = 2\left(\mathbf{B}^2 - \mathbf{E}^2\right).$$

Varying (2.5) with respect to A_μ gives the Maxwell equations in their covariant form:

$$\partial_\mu F^{\mu\nu} = 0. \tag{2.6}$$

The generalization of this Maxwell theory to curved spacetime is almost immediate: we simply replace the partial derivatives in the definition of $F_{\mu\nu}$ by covariant derivatives, which transform in the correct way under Lorentz transformations:

$$F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu.$$

With this redefinition of $F_{\mu\nu}$, the Maxwell action (2.5) is unchanged. In this case, varying with respect to A_μ will give

$$\nabla_\mu F^{\mu\nu} = 0. \tag{2.7}$$

It is also possible to add a source term. This is typically represented by a 1-form J. The action becomes

$$S = \int_M \mathrm{d}^4x \left(F \wedge (\star F) + A \wedge (\star J)\right), \tag{2.8}$$

which is an integral over a 4-form. If we require that this action is gauge invariant, then we need to impose [2]

$$\mathrm{d} \star J = 0 \iff \nabla_\mu J^\mu = 0. \tag{2.9}$$

Usually, we interpret this equation as: "The electromagnetic field only couples to conserved currents." The source term gives a right-hand side to Maxwell's equations (2.7)

$$d \star F = \star J \iff \nabla_\mu F^{\mu\nu} = J^\nu. \tag{2.10}$$

2.3 The stress-energy tensor

To couple matter[1] and gravity, one simply adds the Einstein-Hilbert action to their favourite matter action. For example, in the case of electromagnetism, this would look like

$$S_{\text{tot}} = \int \mathrm{d}^4x \sqrt{-g} \left(\frac{c^3}{16\pi G} R - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}\right), \tag{2.11}$$

with the matter action being

$$S_M = -\frac{1}{4} \int \mathrm{d}^4x \sqrt{-g} F_{\mu\nu} F^{\mu\nu}.$$

This is the relativistic version of the normal Maxwell action. We can now vary the action with respect to both $g_{\mu\nu}$ and A_μ, and we demand that it is stationary with respect to both variations. This gives two (coupled) equations of motion. In case of electromagnetism, we get [2]

$$\begin{cases} G_{\mu\nu} - 8\pi G T_{\mu\nu} & = 0 \\ \nabla_\mu F^{\mu\nu} & = 0 \end{cases} \tag{2.12}$$

Here, $T_{\mu\nu}$ is the *stress-energy tensor* associated to our matter action S_M, which is defined via the formula

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}}. \tag{2.13}$$

In case of electromagnetism, this becomes

$$T_{\mu\nu} = g^{\rho\sigma} F_{\mu\rho} F_{\nu\sigma} - \frac{1}{4} g_{\mu\nu} F^{\rho\sigma} F_{\rho\sigma}.$$

[1] Some authors consider only half-integer spin particles to be matter. The Bosons with integer spin are usually interpreted as carrying a force. We will not make this distinction here, and call every field other than $g_{\mu\nu}$ a matter field.

The stress-energy tensor contains information about the energy and momentum density. For example, in flat space, $T_{00} = \frac{1}{2}\mathbf{B^2} + \frac{1}{2}\mathbf{E}^2$, which should be a familiar expression for the energy density of the electromagnetic field. We also note that the constant G determines how strongly matter couples to the metric, as can be seen from the equations of motion (2.12).

2.4 The Majumdar-Papapetrou solution

In this subsection, we will discuss a specific class of exact solutions to (2.12), which were discovered independently by Majumdar and Papapetrou in the 1940s [11, 12]. In this solution, the metric $\mathrm{d}s^2$ takes the form

$$\mathrm{d}s^2 = -\frac{1}{U^2}\,\mathrm{d}t^2 + U^2\,\mathrm{d}\mathbf{x}\cdot\mathrm{d}\mathbf{x}\,. \tag{2.14}$$

If we take the electromagnetic potential to be

$$A = \frac{1}{U}\,\mathrm{d}t\,,$$

then we get a solution to the Einstein-Maxwell equations, provided that $U = U(\mathbf{x})$ is harmonic, i.e.

$$\nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0.$$

Hartle and Hawking [13] realised that this solution could be used to model a spacetime with multiple black holes. Indeed, if we take

$$U(x,y,z) = 1 + \sum_{i=1}^{N} \frac{M_i}{r_i}, \quad r_i = \|\mathbf{x} - \mathbf{r}_i\| \tag{2.15}$$

then clearly U is harmonic on $\mathbb{R}^3 \setminus \{\mathbf{r}_i\}$. We will see in a moment that this metric describes a static spacetime with N extremally charged Reissner-Nordström black holes.

2.4.1 Mass

Let us investigate the properties of these solutions by taking $n = 1$ and $\mathbf{r}_1 = 0$. The metric (2.14) becomes

$$\mathrm{d}s^2 = -\frac{1}{\left(1+\frac{M}{r}\right)^2}\,\mathrm{d}t^2 + \left(1+\frac{M}{r}\right)^2\,\mathrm{d}\mathbf{x}\cdot\mathrm{d}\mathbf{x}\,. \tag{2.16}$$

As $\|\mathbf{x}\| \to \infty$, also $r \to \infty$, and the metric becomes flat. This is the reason for including the seemingly unnecessary "+1" in U. Moreover, we can expand in powers of $1/r$ to obtain $\left(1+\frac{M}{r}\right)^2 = 1 + \frac{2M}{r} + O\left(\frac{1}{r^2}\right)$ and similarly $\left(1+\frac{M}{r}\right)^{-2} = 1 - \frac{2M}{r} + O\left(\frac{1}{r^2}\right)$. For large r, the metric hence becomes

$$\begin{aligned}
\mathrm{d}s^2 &= -\left(1-\frac{2M}{r}\right)\mathrm{d}t^2 + \left(1+\frac{2M}{r}\right)\mathrm{d}\mathbf{x}\cdot\mathrm{d}\mathbf{x} \\
&= -\left(1-\frac{2M}{r}\right)\mathrm{d}t^2 + \left(1+\frac{2M}{r}\right)\left(\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right) \\
&= -\left(1-\frac{2M}{r}\right)\mathrm{d}t^2 + \left(1+\frac{2M}{r}\right)\mathrm{d}r^2 + (r^2 + O(r))(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2),
\end{aligned}$$

where we can drop the $O(r)$ terms in the spherical part. It can be seen from this *weak field limit* that the parameter M should be interpreted as a mass. Indeed, consider the standard Schwarzschild metric for a spherically symmetric black hole in spherical coordinates [2]:

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2M}{r}\right)^{-1}\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2). \tag{2.17}$$

Recall the standard Taylor expansion $\left(1 - \frac{2M}{r}\right)^{-1} = 1 + \frac{2M}{r} + O\left(\frac{1}{r^2}\right)$. Plugging this in gives

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

We see that this agrees with the large r limit of the MP-solution, and hence M can be interpreted as a mass. Based on this argument, we will assume from now on that $M_i > 0$ for all i.

2.4.2 Charge

We can also calculate the black hole's charge by using Maxwell's equations. Here, we need the material from subsection 2.2. The total electric charge contained in some spatial submanifold Σ is given by

$$Q = \int_\Sigma \star J \tag{2.18}$$

By using Maxwell's equations (2.10) this integral can be rewritten in terms of the field strength tensor [2]:

$$Q = \int_\Sigma d \star F = \int_{\partial\Sigma} \star F \tag{2.19}$$

This leads to the following computation [14], of which the steps are justified below:[2]

$$\begin{aligned} 4\pi Q &= \int_V \mathrm{d}x^\alpha\, \nabla_\alpha \left(F^{\alpha 0}\sqrt{-g}\right) \\ &= \int_{\partial V} \mathrm{d}S_\alpha\, F^{\alpha 0}\sqrt{-g} \\ &= \int_{\partial V} \mathrm{d}S_\alpha\, U^2 \cdot \partial_\alpha(U^{-1}) \\ &= -\int_{\partial V} \mathrm{d}S_\alpha \frac{\partial}{\partial x^\alpha}\left(\frac{M}{r}\right) \\ &= 4\pi M \end{aligned}$$

Hence, in our system of units, the charge of the black hole is equal to its mass. Let us justify some steps in the above calculation. We integrate over a small spherical volume V with boundary ∂V. The first step is Gauss' theorem from electromagnetism,

[2]Here, the index $\alpha = 1, 2, 3$ only has the spatial components. The additional factor of 4π depends on convention

which is equation (2.19). We then apply Stokes' to make it into an integral over the boundary. Next, we calculate the relevant quantities from (2.14). We have

$$\sqrt{-g} = \sqrt{\frac{1}{U^2} \cdot U^6} = U^2$$
$$F_{\alpha 0} = \nabla_\alpha A_0 - \nabla_0 A_\alpha = \partial_\alpha A_0 - \partial_0 A_\alpha = \partial_\alpha \left(U^{-1}\right).$$

The remaining steps are simple. Computing the derivatives gives

$$\begin{aligned}\int_{\partial V} \mathrm{d}S_\alpha \, U^2 \cdot (-1) \cdot U^{-2} \cdot \partial_\alpha U &= -\int_{\partial V} \mathrm{d}S_\alpha \, \partial_\alpha \left(\frac{M}{r}\right) \\ &= \int_0^\pi \int_0^{2\pi} \frac{M\mathbf{r}}{r^3} \cdot \hat{\mathbf{r}} (r^2 \sin\theta \, \mathrm{d}\phi \, \mathrm{d}\theta) = 4\pi M.\end{aligned}$$

The metric (2.14) thus represents an extremally charged black hole.

2.4.3 Surface area

It is possible [14] to convert our metric to the standard form of a *Reissner-Nordström* black hole using the change of coordinates $r \mapsto r - M$. This solution represents a charged non-rotating black hole. It is well-known that these have two event horizons. The metric in its standard form is given by [2, p. 258]

$$\mathrm{d}s^2 = -f(r)^2 \, \mathrm{d}t^2 + f(r)^{-2} \, \mathrm{d}r^2 + r^2 \, \mathrm{d}{\Omega_2}^2 \tag{2.20}$$

and

$$f(r)^2 = 1 - \frac{2GM}{r} + \frac{Q^2}{r^2},$$

where Ω_2 is the metric on a 2-sphere. The two event horizons are located at the roots of $f(r)^2$, namely

$$r_\pm = GM \pm \sqrt{(GM)^2 - Q^2}. \tag{2.21}$$

For an extremal black hole, the inner and outer event horizon coincide. In this case, we have $|Q| = GM$. We will study (charged) black holes and their properties in much more detail in Section 5. For now, it suffices to understand that the black hole is extremal, because adding more charge while keeping the mass fixed is not allowed.

The above change of coordinates looks strange at first glance. However, we can understand it from the following perspective: the point at $r = 0$ does not represent a true point, but a surface with nonzero area. We can calculate this area directly from the metric by choosing as parametrization $t = 0$, $r = r_0$ and θ, ϕ trace out the sphere. We have

$$\left(1 + \frac{M}{r}\right)^2 r^2 (\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2) \to M^2 (\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2)$$

as $r \to 0$. Hence, the surface area tends to $4\pi M^2$, which equals the area of the event horizon.

2.5 Coupling a scalar field

Several groups understood that it was possible to extend the solution to the Einstein-Maxwell equations to a one-parameter family of solutions by adding a scalar field ϕ to the action [15, 16, 17]. The Einstein-Maxwell action becomes[3]

$$S = \int \mathrm{d}^4x \sqrt{-g} \left(R - \partial_\mu \phi \partial^\mu \phi - e^{-2a\phi} F_{\mu\nu} F^{\mu\nu} \right). \tag{2.22}$$

The parameter a is a coupling constant, that measures how strongly the dilaton field ϕ couples to electromagnetism. We recognise the original $\sqrt{-g}R$ term from general relativity, and the $F_{\mu\nu}F^{\mu\nu}$ term from Maxwell. This term is now multiplied by $e^{-2a\phi}$ to couple the dilaton to it. The new term $(\partial\phi)^2$ is just a kinetic term for the scalar field.

The solution is modified to

$$\begin{cases} g & = -U^{-\frac{2}{1+a^2}} \, \mathrm{d}t^2 + U^{\frac{2}{1+a^2}} \, \mathrm{d}\mathbf{x} \cdot \mathrm{d}\mathbf{x} \\ A & = \frac{1}{\sqrt{1+a^2}U} \, \mathrm{d}t \\ e^{-\phi} & = U^{\frac{a}{1+a^2}} \end{cases} \tag{2.23}$$

with

$$U(x, y, z) = 1 + \sum_{i=1}^{n} \frac{(1+a^2)M_i}{r_i}. \tag{2.24}$$

It's easy to see that this reduces to the original solution in case $a = 0$. The full solution again describes an extremal black hole. A computation similar to the one above shows that the various charges must satisfy $Q^2 = M^2 + \Sigma^2$, where Σ is the charge due to the dilaton field.

Depending on the value of a this solution is related to several different reductions from $N = 8$ supergravity [3]. We have already noted that $a = 0$ is just Einstein-Maxwell in 4 dimensions. The case $a = \sqrt{3}$ is related to a reduction of Einstein gravity from 5 to 4 dimensions, which is sometimes called Kaluza-Klein theory. The case $a = 1$ is a low-energy limit of superstring theory. I will say more about these reductions in a later section.

The $N = 8$ supergravity is a supersymmetric theory, in which a lot of observables are completely determined by symmetries. This means it's quite easy to do scattering calculations (e.g. in terms of Feynman diagrams) in this theory. In contrast, general relativity is a very difficult theory to do calculations in. For this reason, it becomes interesting to study $N = 8$ supergravity as an alternative theory of gravity.

The theory at hand is much easier to grasp than the full $N = 8$ supersymmetric theory, and varying the parameter a allows us to access some interesting limits from more complicated theories. This makes it a nice playground to ask physical questions.

[3] This is one of the simplest possible choices. In principle, one could add a potential $V(\phi)$ for the scalar field too. An alternative way to "couple" ϕ and $F_{\mu\nu}$ is to add a term like $\phi F_{\mu\nu}F^{\mu\nu}$. We use the exponential here because it is always positive. This has to do with the "conformal" nature of the dilaton.

2.6 Lagrangian and Hamiltonian formalism

In this subsection, we will obtain the equations of motion satisfied by a photon moving in the field of two extremal Reissner-Nordström black holes. We treat the case $a = 0$ (no dilaton coupling) in detail. When there is dilaton coupling, nothing significantly changes in the derivation. We only need to be careful about replacing U^2 by suitable powers of U in the equations.

When there are no external forces, light rays follow the geodesics of $g_{\mu\nu}$, which is given by equation (2.14), with U determined by (2.15). As we saw in equation (2.1), the geodesic equations can be found using the Lagrangian[4]

$$L = \frac{1}{2} g_{ab}(q(\lambda))\dot{q}^a(\lambda)\dot{q}^b(\lambda). \tag{2.25}$$

We recall that - for photons - we should always have $L = 0$, and that the action determined by L is invariant under reparametrisation. We shall make use of this gauge freedom later.

I want to now move to Hamiltonian formalism. As usual, the canonical momenta can be determined from the Lagrangian:

$$p_\mu = \frac{\partial L}{\partial \dot{q}^\mu} = g_{\mu\nu}\dot{q}^\nu.$$

Hence, the corresponding Hamiltonian is given by

$$H = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu, \tag{2.26}$$

where $g^{\mu\nu}$ is the inverse of the metric tensor. We remark again that motion of photons takes place on the $H = 0$ level set. Substituting the particular form of the metric yields

$$H = \frac{1}{2}\left(-U^2 p_t^2 + U^{-2}(p_x^2 + p_y^2 + p_z^2)\right). \tag{2.27}$$

The function U depends only on the spatial variables x, y, z. Hence, t is a cyclic coordinate and p_t is conserved. Since changing p_t is equivalent to reparametrising the trajectory, we can use our gauge freedom to set $p_t = -1$. We will do this from now on. We will choose to place the two centers in the (x, z)-plane and on the z-axis. With that choice, the Hamiltonian is axially symmetric. This yields conservation of angular momentum in the z-direction:

$$L_z = xp_z - zp_x \implies \dot{L}_z = 0.$$

Here, the dot denotes differentiation w.r.t. the parameter λ. Because of the axial symmetry, a particle in the (x, z)-plane with $p_y = 0$ will stay in that plane. We will later use this symmetry to reduce the system to a 2-dimensional configuration space.

[4]Since light only couples to the metric, the Lagrangian is quite simple. For particles which are charged under either the electromagnetic field or the dilaton, the equations of motion would be more complex. See [3, (4.2)].

Before writing down Hamilton's equations, I will make two observations. First, we note that for light rays, $p_\mu p^\mu = 0$. Writing this out in components yields

$$-U^2 p_t^2 + U^{-2}(p_x^2 + p_y^2 + p_z^2) = 0 \implies p_x^2 + p_y^2 + p_z^2 = U^4 p_t^2 = U^4.$$

Second, observe that

$$\frac{\partial(U^{-2})}{\partial x} = \frac{\partial(U^{-2})}{\partial(U^2)}\frac{\partial(U^2)}{\partial x} = -\frac{1}{U^4}\frac{\partial(U^2)}{\partial x}.$$

Hamilton's equations for our system are then given as follows:

$$\begin{cases} \dot{p_t} &= 0 \\ \dot{t} &= \frac{\partial H}{\partial p_t} = -U^2 p_t = U^2 \\ \dot{p_x} &= -\frac{\partial H}{\partial x} = \frac{1}{2}\frac{\partial(U^2)}{\partial x}p_t^2 - \frac{1}{2}\frac{\partial(U^{-2})}{\partial x}[p_x^2 + p_y^2 + p_z^2] = \frac{\partial(U^2)}{\partial x} \\ \dot{x} &= \frac{\partial H}{\partial p_x} = U^{-2} p_x \end{cases} \tag{2.28}$$

Similar equations hold for y and z. In the above, we have used the two observations in order to simplify $\partial_x(U^{-2})$ and $p_x^2 + p_y^2 + p_z^2$.

When there is dilaton coupling ($a \neq 0$), we know precisely how the metric changes. Redoing the above derivation tells us that we simply need to replace all the U^2-factors by $U^{\frac{2}{1+a^2}}$. Furthermore, we should remember that U itself changes slightly. There is a factor of $1 + a^2$ multiplying the mass of each center. However, this factor should not change anything physically, as it's simply a rescaling of the mass.

2.7 Chaos in (relativistic) N-center problems

Having derived the equations of motion for a photon in the extremal black hole geometry - with or without dilation coupling - we want to study the qualitative features of the dynamics. In particular, we want to understand whether the motion is regular or chaotic. In classical mechanics, the 2-center problem is *integrable*, hence the motion is regular [18, Section 4.3]. However, there is a lot of evidence that indicates the relativistic version of the problem is actually chaotic.

This was first investigated by Contopoulos [4] in 1990. In his work, he studied the motion of light rays in the MP-geometry with two centers. He computed numerically the fundamental periodic orbits and their stability. It turns out (in the case $a = 0$) that these fundamental orbits are highly unstable. There are trajectories that enter from infinity and get close to a periodic orbit. The trajectory can stay close to the periodic orbit for a couple of cycles before "bouncing" to a different periodic orbit, or escaping to infinity. This bouncing between different fundamental orbits is what causes the chaotic behaviour.

After Contopoulos, the problem was studied by Cornish and Gibbons [3] and by Shipley and Dolan [5]. Of these papers, only [3] includes the dilaton coupling in the analysis. Both these papers introduce a symbolic code for the system, which means that they construct a map from the physical orbits to sequences of "symbols". Here,

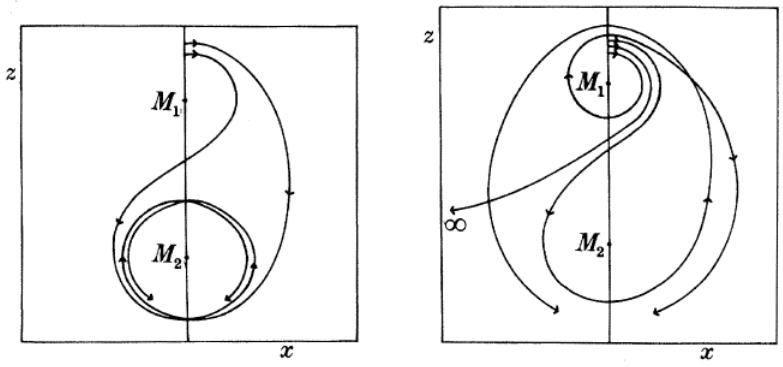


Figure 1: Taken from [4, Figures 5 and 6]. Some orbits of the relativistic 2-center problem ($a = 0$). There are four fundamental orbits: going around M_1, going around M_2, going around both and the figure-8. Trajectories can stay close to one of the fundamental orbits for a while, before scattering to infinity of going to another fundamental orbit.

it is important that orbits which are topologically distinct are mapped to different sequences. We will have a lot more to say about symbolic dynamics in Section 3.

Cornish and Gibbons found a symbolic code for the relativistic 2-center problem which has positive entropy for all values of a between 0 and $\sqrt{3}$, which implies that the dynamics are chaotic for all these values. When $a = \sqrt{3}$, something very surprising happens. The authors prove that the system with $a = \sqrt{3}$ is *integrable*, hence regular. Actually, in some sense, the system reduces to the classical 2-center problem. All the relativistic corrections cancel out in all orders of perturbation theory!

The problem with their argument is that they do not prove that their symbolic code "works". We will also make this more precise in Section 3. Essentially, the issue is that it is not clear whether each sequence actually corresponds to an allowed orbit in the system. Mathematically, this is key to proving that the dynamical system is chaotic.

In Section 4, we will introduce some material which will allow us - in certain cases - to construct a proper symbolic code for the system. For this, we will use tools from scattering theory - specifically the notion of topological degree. To apply these tools, it is crucial that the motion is *complete*, which means that we can follow any trajectory for an arbitrarily large amount of time. This leads to problems in systems where there are collisions.

We will show that the system is *regularizable* for a discrete set of values between 1 and $\sqrt{3}$. At least in these cases, we will be able to prove the existence of chaos in relativistic N-center problems. Our method does not quite work for the 2-center case. We will be able to construct a symbolic code, but its entropy will be zero. For this case, I believe that - with minor changes - the method can be adapted to still prove chaos. However, the fact that the problem is not easily regularizable for most values of a is a high barrier. These cases are beyond the scope of this work.

3 Symbolic dynamics and chaos

In this section, we will introduce some material from dynamical systems theory. Specifically, we shall make the notions of integrability and regular and chaotic motion more precise. All of this material is standard, and included for completeness. Some good references are [19] and [20].

3.1 What is chaos?

The dynamical systems we study in mathematics are typically *deterministic*, which means that the initial conditions fully determine the state of the system at all future times. This assumption is also baked into classical mechanics. Indeed, consider Newton's second law $F = m\ddot{x}$. The forces only depend on $x, \dot{x}$ and possibly t. This gives a second order differential equation for x. If we know the initial *state* of the mechanical system, i.e. the position and velocity at $t = t_0$, then Newton's law fully determines the state for all $t > t_0$.

However, we usually don't know the precise initial state. This could be e.g. due to measurement errors or imprecisions in setting up the system. We could imagine this as a small open ball U in the state space Σ. Mathematically speaking, a dynamical system is an action of some Abelian group (usually $\mathbb{Z}$ or $\mathbb{R}$) on Σ. This corresponds to a map[5] $\Phi : G \times \Sigma \to \Sigma$. Being a group action, this map should satisfy

- $\Phi_0(x) = x$ for all $x \in \Sigma$;
- $\Phi_t \circ \Phi_s(x) = \Phi_{t+s}(x)$ for all $t, s \in G$ and all $x \in \Sigma$;
- $\Phi_{-t}(x) = (\Phi_t)^{-1}(x)$ for all $t \in G$ and all $x \in \Sigma$.

Here, $\Phi_t(x) := \Phi(t, x)$. The third item implies in particular that each Φ_t is invertible.[6] We typically also require that Φ_t preserves some structure of Σ, e.g. that the map is continuous or smooth. Usually, the flow Φ is induced by some differential equation (like Newton's second law).

Given $U \subseteq \Sigma$, we can hence consider the set $\Phi_t(U)$ for all $t \in G$. The properties of this set determine whether or not we call the system chaotic. There is no universally accepted definition of a chaotic system, but all of the definitions should capture the following idea: "Chaos is when the present determines the future, but the approximate present does not approximately determine the future." This quote is due to Lorenz, who studied chaotic behaviour in weather prediction.

The following is the most common definition of a chaotic system, which is due to Devaney [19, p. 338].

Definition 3.1. Let (Σ, Φ) be a dynamical system, and assume that $d : \Sigma \times \Sigma \to \mathbb{R}^+$ is a metric on Σ. Assume that all functions involved are continuous w.r.t. d. We say that the system is chaotic if the following three properties are satisfied:

[5] Here, G is $\mathbb{Z}$ or $\mathbb{R}$. We interpret this as the time variable.

[6] Strictly speaking, it is not necessary to impose the third condition. In this case, we have the action of a monoid $\mathbb{N}_0$ or $\mathbb{R}^+$ on Σ.

1. Periodic orbits are dense in Σ;

2. (Topological transitivity) Given $U, V \subseteq \Sigma$ open there exist $x_0 \in U$ and $t \in G$ such that $\Phi_t(x_0) \in V$;

3. (Sensitive dependence on initial conditions) There exists a *sensitivity constant* $\beta > 0$ such that for any $x_0 \in \Sigma$ and any open neighbourhood $U \ni x_0$ there exist $y_0 \in U$ and $t \in G$ such that

$$d(\Phi_t(x_0), \Phi_t(y_0)) > \beta. \tag{3.1}$$

The third condition is certainly the most well-known, but it is not the only characteristic of a chaotic system. For example, consider the simple system defined by $\dot{x} = ax$, whose solutions are given by $x(t) = e^{at}x_0$. This system has sensitive dependence upon initial conditions, but it clearly does not satisfy either of the first two criteria.

One of the simplest examples of a chaotic system is the *baker map* [19]. We consider the function

$$f(x) = 2x \pmod 1 \tag{3.2}$$

which we view as a map from $D = [0, 1)$ to itself. By iterating this map we get a dynamical system

$$\Phi(n, x) = f^n(x) = f \circ f \circ \cdots \circ f(x) \tag{3.3}$$

on $\Sigma = D$. Since f is not invertible, this map is only defined on the monoid $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. It can be shown that

$$f^n(x) = 2^n x \pmod 1. \tag{3.4}$$

The graph of the first few iterates of f is shown in Figure 2.

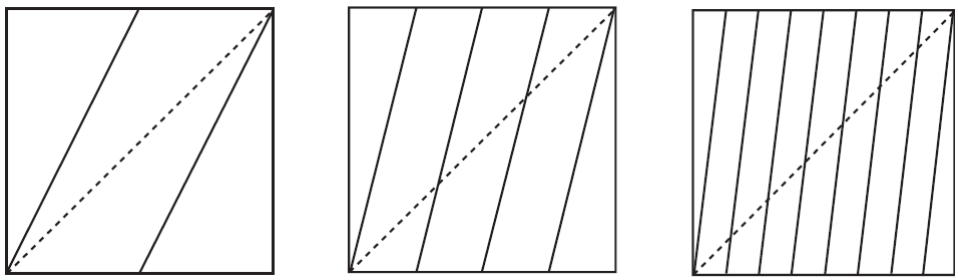


Figure 2: Taken from [19, Figure 15.9]. Plots of f, f^2 and f^3. Dotted line represents the diagonal $y = x$.

Theorem 3.2. *The baker map is chaotic.*

Proof. We can simply check the three conditions in the definition. Note that a point $x \in D$ is periodic if and only if $f^n(x) = x$ for some $n \in \mathbb{N}_0$. As can be seen from Figure 2, the graph of f^n crosses the diagonal once in every interval of the form $I_k^n = [\frac{k}{2^n}, \frac{k+1}{2^n})$. Hence, there is a periodic point in every such interval. For each n,

these intervals cover D and their length tends to zero as $n \to \infty$. Hence, it follows that periodic orbits are dense in D.

Given $U \subseteq D$ open, we may choose k, n such that $I_k^n \subseteq U$. Since $f^n(I_k^n) = D$, given any $V \subseteq D$ we can find $x_0 \in I_k^n \subseteq U$ such that $f^n(x_0) \in V$. This proves topological transitivity.

Choose $\beta = \frac{1}{2}$. Given $x_0 \in D$ and $U \ni x_0$ open, we again choose $I_k^n \subseteq U$. Since f^n maps I_k^n onto D, it's obvious that we can find y_0 such that the distance between $f^n(x_0)$ and $f^n(y_0)$ is at least $\frac{1}{2}$. For example, if $f^n(x_0) = 0.3$, we choose y_0 such that $f^n(y_0) = 0.8 + \epsilon$. We can do a similar construction if $f^n(x_0)$ is in the right-half of D. Thus, there is sensitive dependence on initial conditions, and the system is chaotic. □

In most situations, checking these conditions is not so easy. We could only do it in the above example because the map was so simple. Still, the example is not completely useless. We can prove that chaotic behaviour is preserved by conjugacies. Hence, if we can prove that a dynamical system is conjugate to a known chaotic system, then the original system must already be chaotic. Let's make this more precise.

Definition 3.3. Two functions $f : \Sigma \to \Sigma$ and $g : \Sigma' \to \Sigma'$ between topological spaces Σ, Σ' are said to be (topologically) conjugate if there exists a homeomorphism $h : \Sigma' \to \Sigma$ such that $f \circ h = h \circ g$. If h is only a continuous surjection, then f and g are said to be (topologically) semi-conjugate.

Let us now use f and g to define dynamical systems on Σ and Σ' respectively. This is done similarly to the baker map example. Namely, time is discrete, and $\Phi(n, x)$ is obtained by iterating f and g.

Proposition 3.4. *Suppose that $h : \Sigma' \to \Sigma$ is a semi-conjugacy, and that the system (Σ', g) is chaotic. Then, the system (Σ, f) is also chaotic.*

Proof. We will only prove transitivity, as the other two properties are similar. Suppose that $U, V \subseteq \Sigma$ are open. Then, by continuity of h, also $h^{-1}(U)$ and $h^{-1}(V)$ are open as subsets of Σ'. Since the system on Σ' is chaotic, there must exist points $x_0 \in h^{-1}(U)$ and $n \in \mathbb{N}_0$ such that $g^n(x_0) \in h^{-1}(V)$. Finally - by definition - $h(x_0) \in U$ and $h(g^n(x_0)) = f^n(h(x_0)) \in V$. The point $h(x_0)$ is thus the point we were looking for. □

3.2 Symbolic dynamics

With the previous proposition in mind, it makes sense to look for other examples of chaotic systems. In this subsection, I will introduce a simple example of a chaotic dynamical system, namely the shift map. This innocent looking system turns out to be a very powerful tool in proving the existence of chaos in a more complex systems. Indeed, we can sometimes prove that a system "contains a shift map", and hence it is chaotic.

Define $\mathcal{A} = \{1, 2, \ldots, m\}$. This set is referred to as the *alphabet*. On $\mathcal{A}$, we put the discrete metric

$$d(i, j) = \delta_{ij}, \quad i, j \in \mathcal{A}. \tag{3.5}$$

We now look at the space of bi-infinite sequences with entries in $\mathcal{A}$. We define the space Σ as a Cartesian product

$$\Sigma = \cdots \times \mathcal{A} \times \mathcal{A} \times \ldots$$

An example of an element $s \in \Sigma$ would be $\ldots 121212 \ldots$ Alternatively, an element of Σ is a function from $\mathbb{Z}$ to $\mathcal{A}$. In this way, we are allowed to write $s = \ldots s_{-2}s_{-1}.s_0s_1s_2 \ldots$, where $s_i := s(i)$ and $s : \mathbb{Z} \to \mathcal{A}$. A "decimal" point is inserted arbitrarily before the zeroth entry, which allows us to keep track of the indexing.

We can also make Σ into a metric space, for example by using

$$d(s, \bar{s}) = \sum_{i \in \mathbb{Z}} \frac{1}{2^{|i|}} \frac{d(s_i, \bar{s}_i)}{1 + d(s_i, \bar{s}_i)}. \tag{3.6}$$

Note that the fraction $d/(1 + d)$ is either 0 or $\frac{1}{2}$ depending on whether s_i and $\bar{s}_i$ are the same or different. The above sum is hence convergent, being dominated by a geometric series. For more properties of the metric space (Σ, d), see [20, Section 24.1].

We now define a dynamical system on Σ which effectively moves the decimal point one place to the right. More formally, we introduce the map $\sigma : \Sigma \to \Sigma$ which acts on a sequence s by

$$\sigma(s)(i) = s(i + 1), \quad i \in \mathbb{Z}. \tag{3.7}$$

We call σ the shift map. Note that the map σ depends on the choice of alphabet $\mathcal{A}$, but this is not reflected in the notation. Now, let us investigate the properties of σ in more detail.

Proposition 3.5. *The map $\sigma : \Sigma \to \Sigma$ is continuous.*

Proof. See [20, Proposition 24.2.1] □

Theorem 3.6. *[20, Modification of Proposition 24.2.2] The dynamical system (Σ, σ) defined by iterating the shift map on Σ is chaotic.*

Proof. For simplicity, we will assume that $\mathcal{A} = \{0, 1\}$ and that the sequences in Σ are only infinite in *one* direction, so that $\Sigma = \{0, 1\}^{\mathbb{N}}$ The general case works similarly.

Fix $s \in \Sigma$. We will first prove that there is a periodic orbit arbitrarily close to it. To this end, define s_N by repeating the first N entries of s periodically. By choosing N large enough, we find an orbit of period N close to s.[7]

To prove transitivity, we consider the element $s^* = (0\ 1|0\ 0\ 0\ 1\ 1\ 0\ 1\ 1| \ldots)$, which first has all possible sequences of length 1, then length 2, etc. Consider an arbitrary

[7] This works because of our choice of metric on Σ. By fixing the first N entries of the sequence, everything else is the tail of a geometric series, which can be as small as we like.

$s \in \Sigma$. By applying σ to s^* repeatedly, eventually $\sigma^K(s^*)$ will have the same initial N digits as s. The orbit or s^* thus comes arbitrarily close to any point in Σ, and we have transitivity.

Set $\beta = 2$ and fix $s \in \Sigma$. Consider $t = (s_1, s_2, \ldots, s_N, \hat{s}_{N+1}, \hat{s}_{N+2})$, which agrees with s in the first N components and then completely disagrees. By tuning N, we can make $d(s,t)$ as small as we like. However, $d(\sigma^N(s), \sigma^N(t)) = 2$, since they are maximally different. This proves sensitive dependence. □

By Proposition 3.4, if we can find a semi-conjugacy h from the space of sequences onto some other space Σ', then the dynamical system defined on Σ' is also chaotic. One can use this technique to prove that the famous logistic map

$$f_a(x) = ax(1-x)$$

is chaotic for $a > 4$ [19, Section 15.6]. In this work, we are looking for a more elaborate application. The involved dynamics are a lot more complicated than what we have looked at so far. Also, they are generated by an ODE, hence we have continuous time. To remedy this, we will first have to transform our scattering problem to a discrete time system.

The standard way to do this is to introduce a *Poincaré section*. Let M be the phase space - also called state space - of the n-dimensional dynamical system, and let $\Phi_t : M \to M$ be the flow. We are looking for an $(n-1)$-dimensional subset $S \subseteq M$ which is *transversal* to the flow. Intuitively, this means that a trajectory which hits S must immediately exit S.[8] Hence, a trajectory cannot get stuck in S.

We define the *Poincaré map* $P : \Lambda \subseteq S \to S$ as the first-return map. Specifically, pick a point $x \in S$ and look at its orbit $\Phi_t(x)$. We define $P(x)$ as the first point in S which the orbit through x meets. More mathematically, set

$$\tau = \min\{t > 0 : \Phi_t(x) \in S\}, \quad P(x) = \Phi_\tau(x).$$

If the orbit never returns to S, we leave P undefined. This defines a discrete dynamical system P on some subset Λ of S. Note that we have also reduced the dimensionality of the dynamical system by 1.

The next step is to introduce a "natural" partition of S, i.e. a function $F : S \to \mathcal{A}$. For example, if S is disconnected, we can label the connected components of S by $\mathcal{A}$ and send $x \in S$ to the number of its component. When chosen in this way, the map F is even continuous, which will be important in what follows.

An orbit of the dynamical system can now be described by a sequence of elements in $\mathcal{A}$ according to the order in which the orbit meets different areas of the partition. Each orbit is hence associated with a sequence of elements of $\mathcal{A}$. This sequence is finite if the orbit escapes after a finite amount of applications of P. By running the

[8] Mathematically, the vector field X associated to the flow Φ_t may not lie completely in T_xS at any $x \in S$. It must have a nonzero component in the "orthogonal" direction.

flow backwards in time, we can also define the negative entries of the sequence.

If Λ is chosen to be invariant under P (i.e. $P(\Lambda) \subseteq \Lambda$), then we get a genuine map $h : \Lambda \to \Sigma$, where Σ is the space of sequences with elements in $\mathcal{A}$. Note that this map is automatically a conjugacy. This should be intuitively clear. To prove it mathematically, we will have to be a bit more precise about the definition of h.

For $x \in \Lambda$, the element $h(x)$ is a sequence with elements in $\mathcal{A}$. Its i'th entry is - by definition -

$$h(x)(i) = F(P^i(x)),$$

where $F : \Lambda \to \mathcal{A}$ is the partition introduced earlier. To prove that h is a conjugacy, we need to prove that $h \circ P = \sigma \circ h$. For $x \in \Lambda$ and $i \in \mathbb{Z}$ arbitrary, we have

$$\begin{aligned}(\sigma(h(x)))(i) &= h(x)(i+1) \\ &= F(P^{i+1}(x)) \\ &= F(P^i(P(x))) \\ &= h(P(x))(i),\end{aligned}$$

by unravelling all the definitions. If the map $h : \Lambda \to \Sigma$ is invertible, then we obtain $P \circ h^{-1} = h^{-1} \circ \sigma$, i.e. a conjugacy from the shift map to the Poincaré map. This is the crucial step! *If* h is invertible, then Proposition 3.4 proves that the dynamical system (Λ, P) is chaotic. Since this is obtained by a reduction of the original continuous-time system, it must also be chaotic.

Invertibility is the subtle point. We need that every sequence of symbols we *can* write down actually *does* come from some physical trajectory of the system. In general, this is a very difficult problem. We need to prove existence of special types of solutions to the original (complicated) differential equation. In Section 4, we will use sophisticated methods - such as degree theory - to overcome these difficulties.

3.3 The notion of integrability

Let us take a moment to consider the opposite end of the spectrum. For mechanical systems, we consider the notion of an *integrable* system. The phase space of such a system has so much structure that chaotic motion becomes impossible. In general, a mechanical system can be defined on a symplectic manifold (M, ω). Typically, we choose an n-dimensional *configuration space* Q and consider its cotangent bundle $M = T^*Q$ with the canonical symplectic form [18]. The reader who has no idea what that means may safely consider $Q = \mathbb{R}^d$ and $M = \mathbb{R}^{2d}$ with coordinates (q, p) denoting positions and momenta. In these coordinates, the symplectic form is either

$$\omega = \pm \, \mathrm{d}q \wedge \mathrm{d}p$$

depending on convention.

Consider a pair of smooth function $f, g : M \to \mathbb{R}$ defined on the phase space. Recall that the *Poisson bracket* between f and g is defined in local coordinates by

the equation

$$\{f, g\} = \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p}. \tag{3.8}$$

A mechanical system on (M, ω) is defined by the choice of a Hamiltonian function $H : M \to \mathbb{R}$. The symplectic form ω sets up a bijection between vector fields and 1-forms on M. The vector field associated to $\mathrm{d}H$ generates the time evolution of our system. It can be shown that any function $f : M \to \mathbb{R}$ thus has time evolution

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, H\}. \tag{3.9}$$

In particular, we have

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \{q, H\} = -\frac{\partial H}{\partial p}, \quad \frac{\mathrm{d}p}{\mathrm{d}t} = \{p, H\} = \frac{\partial H}{\partial q}, \tag{3.10}$$

which are Hamilton's equations of motion.

We are now ready to give the definition of an integrable system.

Definition 3.7. A mechanical system defined by (M, ω, H) where $\dim(M) = 2d$ is said to be integrable if there exist d independent[9] functions $f_i : M \to \mathbb{R},\ i = 1, 2, \ldots, d$ such that $\{f_i, f_j\} = 0$ for all i, j.

In particular, one can always choose $f_1 = H$. With this choice, integrability is equivalent to the existence of d independent conserved quantities.[10] Note that any system with a one-dimensional configuration space is automatically integrable by conservation of energy. Also, any 2-dimensional system with a rotational symmetry is integrable. The two conserved quantities are energy and angular momentum.

We now look at the phase space structure of an integrable system, which I eluded to earlier. This theory is due to Liouville and Arnold [21, p. 272]. Consider a level set of the functions f_i:

$$M_{\mathbf{h}} = \{x \in M : f_i(x) = h_i\}, \quad \mathbf{h} \in \mathbb{R}^d. \tag{3.11}$$

Since the f_i are conserved, motion can't hop from one $M_{\mathbf{h}}$ onto another $M_{\mathbf{h}'}$. More can be said.

Theorem 3.8. *If $M_{\mathbf{h}}$ is compact and connected, then it is diffeomorphic to a d-dimensional torus. The phase flow on $M_{\mathbf{h}}$ is quasi-periodic, i.e. on each $\mathbb{S}^1$ the motion is periodic with frequency $\omega_i(\mathbf{h})$. Hamilton's equations of motion can be integrated in quadratures.*

The last point of the Theorem is where the word "integrable" comes from. It effectively means that we can solve the differential equations by using a finite number of integrals. We will see later how this works for a two-dimensional system with a

[9]Here, we say that the functions f_i are independent if their differentials df_i are linearly independent almost everywhere.

[10]Indeed, each f_i is conserved along solutions of the system because $\frac{\mathrm{d}f_i}{\mathrm{d}t} = \{f_i, H\} = 0$

rotational symmetry.

To see where the notion of integrability fits into our story, we consider the N-center problem of classical mechanics. This system has Hamiltonian

$$H = \frac{\|p\|^2}{2} - \sum_{i=1}^{N} \frac{M_i}{\|q - s_i\|}, \quad q, p \in \mathbb{R}^3. \tag{3.12}$$

It describes a test particle moving in the field of N stationary masses M_i located at $s_i \in \mathbb{R}^3$. If $n = 1$, this is the Kepler problem, which is special in a number of ways. First of all, the 1-center problem is actually equivalent to the 2-body problem. The latter separates into two parts in the center-of-mass frame. Namely, the center of mass undergoes free motion, while reduced system ($\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_2$) satisfies the equation for a 1-center problem. This is not the case for the 3-body and 2-center problems.

Solutions to the 1-center problem are completely determined by Kepler's three laws of planetary motion. In particular, the trajectories are conic sections. Hence, for bound motion, they are ellipses. This is rather special, because it implies that the orbits close precisely. This doesn't happen for arbitrary central force problems. In fact, there are only two central potentials for which all bound orbits close: Kepler's problem and the harmonic oscillator.[11]

It turns out that the closing of all bound orbits is related to *maximal superintegrability*. The Kepler problem is not just integrable, which would require three independent constants of motion. It turns out that there are seven constants of motion, of which five are independent [18]. The conserved quantities are energy (E), three components of angular momentum ($\mathbf{L}$) and three components of the Laplace-Runge-Lenz vector ($\mathbf{A}$). This last vector always points in the direction of the semi-major axis for bounded motion, and its conservation means that the orbits close. We call the Kepler problem *maximally superintegrable*, because it has $2d-1 = 5$ independent constants of motion, which is the most that any d-dimensional system can have.

For the classical 2-center problem, the rotational symmetry is partially broken. There are still two obvious constants of motion (energy and one component of angular momentum). It is even possible to find a third constant of motion [18], which makes the problem integrable. On the other hand, the 3-center problem is chaotic[12], hence it cannot be integrable.

What happens for relativistic problems? General relativity predicts a precession for the perihelion of Mercury, which has been confirmed experimentally. This was one of the great successes of Einstein's theory. To first order in v^2/c^2, the theory predicts

$$\delta\theta = 6\pi \frac{G^2 M^2}{l^2 c^2}.$$

From this, we see that the LRL-vector is no longer conserved. The rotational symmetry is however preserved, so the system is still superintegrable (but not maximally).

[11] This result is known as Bertrand's theorem [22].

[12] At least if the three centers are in general position.

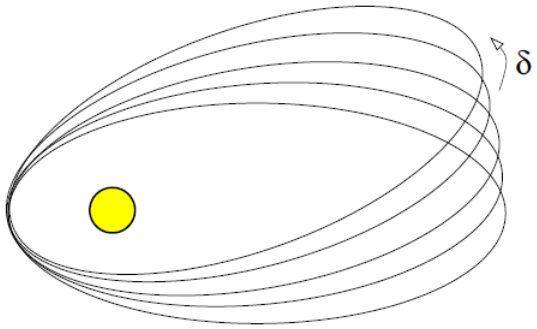


Figure 3: Taken from [2, Figure 7]. Precession of the orbit of Mercury due to relativistic effects.

As I have laid out before, there is evidence that the relativistic 2-center problem is chaotic, hence not integrable.

We can also introduce the photon (spin 1) and the dilaton into the theory. One can ask what happens to the precession of the perihelion and integrability in these cases. It turns out that the answer depends on the coupling strength a of the dilation. In case $a = \sqrt{3}$, the integrability of the 2-center problem returns. For all values of a between 0 and $\sqrt{3}$ there is good indication that the system is chaotic. We will get back to this in Section 6.

Regarding the perihelion precession: In a pure spin 0 theory, the perihelion of Mercury also precesses, but in the opposite direction w.r.t. general relativity. One can imagine that it might be possible to tune the coupling constant a in such a way that the precession due to spin 0 and spin 2 exactly cancel each other out - at least to first order in v^2/c^2. Such things are still under investigation.

4 Classical scattering theory

This section introduces the main ideas behind scattering theory. We will mainly be interested in the long term behaviour of particles coming in from infinity towards some potential $V(x)$. Under rather mild conditions on the potential, the trajectory of the particle approaches a straight line. We introduce the scattering transformation, which encodes the data of our scattering experiment. Moreover, we introduce the scattering degree $\deg(E)$. This generalization of the winding number - which originated in algebraic topology - can be used to prove existence of trajectories with a certain scattering angle. In some cases, we will be able to use the scattering degree to construct symbolic dynamics.

4.1 Scattering experiments and the differential cross-section

Let us consider a situation where the potential $V(x)$ is centred at the origin. Particles are incident from the left, at a distance b from the z-axis. We will assume for the moment that the system is rotationally symmetric, so we only need to consider motion in the plane. We call b the *impact parameter*, and θ the *scattering angle*. See Figure 4. Our goal is to determine θ as a function of b.

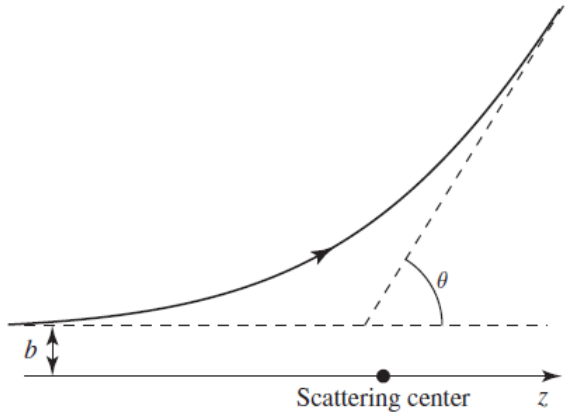


Figure 4: Taken from [23, Figure 10.1]. Definition of the impact parameter and scattering angle.

As an example, let's consider *hard sphere scattering*. The target is a heavy stationary sphere of radius R, and the incoming particle bounces off the sphere elastically. See Figure 5. It's easy to see from the figure that $\sin\alpha = \frac{b}{R}$ and that $\theta = \pi - 2\alpha$. Thus,

$$\sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right) = \frac{b}{R} \implies \cos\left(\frac{\theta}{2}\right) = \frac{b}{R}. \tag{4.1}$$

Of course, this relation only holds if $b \leq R$. In case $b \geq R$, we see that the particle is not deflected, and $\theta = 0$. The relation is thus

$$\theta = \begin{cases} 2\cos^{-1}\left(\frac{b}{R}\right), & b \leq R \\ 0, & b \geq R \end{cases} \tag{4.2}$$

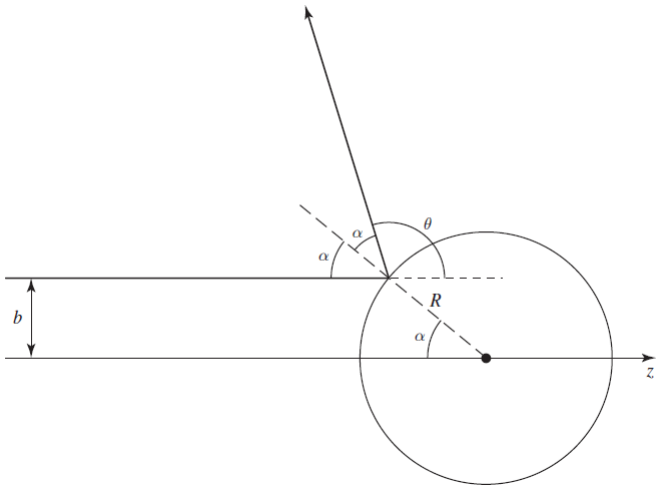


Figure 5: Taken from [23, Figure 10.2]. Geometry of hard sphere scattering.

I now introduce the concept of *differential cross-section*, which is a quantity that can be measured in experiment. Consider particles scattering from an infinitesimal area element $\mathrm{d}\sigma$ into the solid angle $\mathrm{d}\Omega$. In the limit, we expect these to be proportional, and the proportionality "constant" is denoted by $D(\theta)$. Hence, by definition, $\mathrm{d}\sigma = D(\theta)\,\mathrm{d}\Omega$. The function $D(\theta)$ is called the differential cross-section. Because of symmetry, it is independent of ϕ.

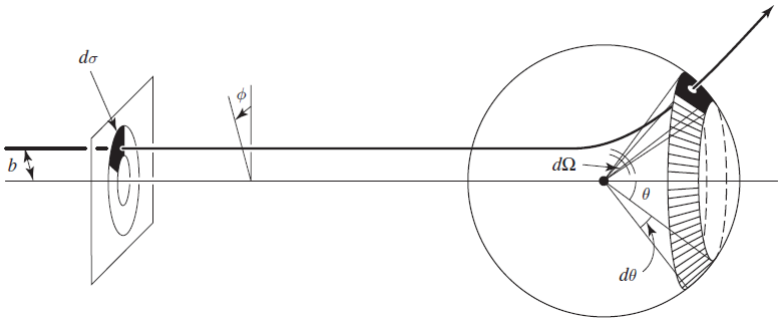


Figure 6: Taken from [23, Figure 10.3]. Infinitesimal areas involved in the definition of the differential cross section.

By using the standard formulae $\mathrm{d}\sigma = b\,\mathrm{d}b\,\mathrm{d}\phi$ and $\mathrm{d}\Omega = \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\phi$, we get

$$D(\theta) = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{b}{\sin(\theta)}\left|\frac{\mathrm{d}b}{\mathrm{d}\theta}\right|. \tag{4.3}$$

Here, we note that $D(\theta)$ should obviously be positive, since both $\mathrm{d}\sigma$ and $\mathrm{d}\Omega$ represent areas. Hence, we are supposed to take an absolute value, even though $b(\theta)$ is typically a decreasing function. We also define the *total cross-section* as

$$\sigma = \int \mathrm{d}\sigma = \int D(\theta)\,\mathrm{d}\Omega\,, \tag{4.4}$$

i.e. the integral of the differential cross-section over all solid angles.

In the example of hard sphere scattering, we find

$$\begin{aligned} D(\theta) &= \frac{b}{\sin(\theta)} \left| \frac{\mathrm{d}}{\mathrm{d}\theta} \left(R\cos\left(\frac{\theta}{2}\right)\right)\right| \\ &= \frac{1}{2}R \cdot \frac{R\cos\left(\frac{\theta}{2}\right)}{\sin(\theta)} \cdot \sin\left(\frac{\theta}{2}\right) \\ &= \frac{1}{4}R^2 \end{aligned}$$

and the total cross-section is

$$\sigma = \int \frac{1}{4} R \,\mathrm{d}\Omega = \pi R^2,$$

as expected for the area of a sphere exposed to a beam.

Finally, I will explain why the differential cross section can be measured in a scattering experiment. We consider a beam with luminosity L, which means $\mathrm{d}N = L\,\mathrm{d}\sigma$, where N is a number of events. Then, we have

$$\mathrm{d}N = L\,\mathrm{d}\sigma = LD(\theta)\,\mathrm{d}\Omega \implies D(\theta) = \frac{1}{L}\frac{\mathrm{d}N}{\mathrm{d}\Omega}.$$

The above is independent of the impact parameter, which usually cannot be controlled in an experiment. The differential cross-section can be determined by measuring the number of counts as a function of solid angle.

4.2 Classical potential scattering

Now that we understand which parameters are physically relevant, let us describe the mathematics of scattering theory. I want to develop the scattering transformation for a classical particle incident on a potential $V(q)$, and derive some of its properties. The main reference is [6]. One can define the scattering transformation in many different settings. In each case, it will be a slightly different type of object, but its core properties remain the same. See for example [24].

We will use the Hamiltonian formalism of classical mechanics. Consider a d-dimensional configuration space $\mathbb{R}^d_q$. The momenta are elements of the cotangent space $\mathbb{R}^d_p$, and hence the phase space is $\mathbb{R}^d_q \times \mathbb{R}^d_p$, which is isomorphic to $\mathbb{R}^{2d}$.

We will sometimes remove one or more points from the configuration space. In such cases, the phase space will be $\left(\mathbb{R}^d_q \setminus \{s_1, \ldots, s_n\}\right) \times \mathbb{R}^d_p$. More generally, it is possible to take T^*M as the phase space of a mechanical system, where M is the configuration space. In any case, the phase space locally looks like $\mathbb{R}^d_q \times \mathbb{R}^d_p$.

The Hamiltonian for our scattering process is written in the form "kinetic + potential". Indeed,

$$H(p,q) = \frac{1}{2}\|p\|^2 + V(q). \tag{4.5}$$

We will also define the *free Hamiltonian*

$$H_0(p,q) = \frac{1}{2}\|p\|^2. \tag{4.6}$$

The flows corresponding to these Hamiltonians will be denoted by Φ_t and Φ_t^0 respectively. At least locally, the flow of a differential equation is a group of diffeomorphisms acting on the phase space. We will need to put some restrictions on the potential V in order to get global results concerning existence of the flow and in order to compare Φ_t and Φ_t^0. We define three classes of potentials in increasing order of niceness: long range, short range and compactly supported.

Definition 4.1. [6, Definition 12.1] Let $V \in C^2(\mathbb{R}^d_q, \mathbb{R})$. We say that V is

- **long range** if
$$|\partial^i V(q)| \leq c\langle q\rangle^{-|i|-\epsilon} \quad \forall q \quad \forall |i| \leq 2$$
- **short range** if
$$|\partial^i V(q)| \leq c\langle q\rangle^{-|i|-1-\epsilon} \quad \forall q \quad \forall |i| \leq 2$$
- **compactly supported** if V is smooth and vanishes outside of some compact set $K \subseteq \mathbb{R}^d_q$.

We will always assume that V is at least C^2 in order to guarantee local existence and uniqueness of solutions to Hamilton's equations. By definition, the symbol $\langle q\rangle$ means

$$\langle q\rangle = \sqrt{\|q\|^2 + 1} \in [1, \infty),$$

which is of order $\|q\|$. Intuitively, a potential is short range if it falls of faster than $1/r$, and it is short range if it falls of faster than $1/r^2$. It is obvious that compactly supported $\subseteq$ short range $\subseteq$ long range.

Assume that V is (at least) a long range potential. Under this assumption, the limit $\lim_{\|q\|\to\infty} V(q)$ exists. By adding a suitable constant, we can and will **always** assume that this limit equals zero. We define the following flow-invariant subsets of the phase space P:

Definition 4.2. [6, Definition 12.3] Let V be a long range potential.

- Define the sets
$$b^{\pm} = \left\{ x \in P : \limsup_{t\to\pm\infty} \|q(t;x)\| < \infty \right\}.$$
We say that $x \in P$ is a *bound state* if $x \in b^+ \cap b^-$.
- Define the sets
$$s^{\pm} = \left\{ x \in P : \lim_{t\to\pm\infty} \|q(t;x)\| = \infty \right\}.$$
We say that $x \in P$ is a *scattering state* if $x \in s^+ \cap s^-$.
- Define the sets $t^{\pm} = b^{\pm} \cap s^{\pm}$. We say that $x \in P$ is a *trapped state* if $x \in t^+ \cup t^-$.

- For a real number E, we define Σ_E to be the pre-image of E under the Hamiltonian, i.e. all states with energy E. The sets $b_E^\pm, b_E, s_E^\pm, s_E, t_E^\pm$ are defined as the intersection of Σ_E with the respective set, e.g. $b_E = b \cap \Sigma_E$.

The following three Theorems give the asymptotic behaviour of a particle incident on $V(x)$. We split cases depending on whether V is long-range, short-range or compact support. I have combined information from Theorem 12.5, Corollary 12.8 and Theorem 12.11 from [6].

Theorem 4.3. *(Long range potentials) If V is a long range potential, then:*

- *The flow Φ_t is complete.*[13]
- *For any scattering state $x \in s_E^\pm$ the limit*

$$p^\pm(x_0) = \lim_{t \to \pm\infty} p(t; x)$$

exists and has the right energy:

$$\|p^\pm(x_0)\| = \sqrt{2E}.$$

*We call $p^\pm$ the **asymptotic momentum**. We call $\hat{p}^\pm := p^\pm / \|p^\pm\|$ the **asymptotic direction**.*

- *The energy shell Σ_E is a disjoint union of b_E, s_E and t_E. Moreover, b_E is compact, s_E is open and t_E has Liouville measure zero.*
- *If E is large enough, then $\Sigma_E = s_E$.*

Importantly, the Kepler potential $V(q) = \frac{1}{\|q\|}$ is *not* long range. For this potential, the asymptotic momentum does exist, but the speed as $t \to \infty$ will differ too much from $\sqrt{2E}$.

Theorem 4.4. *(Short range potentials) If V is a short range potential, then*

- *For any scattering state $x \in s_E^\pm$, the limit*

$$q_\perp^\pm(x) = \lim_{t \to \pm\infty} \Big(q(t; x) - \langle q(t; x), \hat{p}^\pm(x) \rangle \hat{p}^\pm(x) \Big)$$

*exists. We call $q_\perp^\pm$ the **asymptotic impact parameter**.*

- *The functions*

$$\Omega^\pm : \Big(\mathbb{R}_p^d \setminus \{0\} \Big) \times \mathbb{R}_q^d \to s^\pm, \quad \Omega^\pm(x) := \lim_{t \to \pm\infty} \Big(\Phi_{-t} \circ \Phi_t^0(x) \Big)$$

are well-defined homeomorphisms. Moreover, they conserve energy ($H \circ \Omega^\pm = H_0$), they respect time reversal and they conjugate the free and interacting flows ($\Omega^\pm \circ \Phi_t^0 = \Phi_t \circ \Omega^\pm$).

[13] This means that we can follow any solution for an arbitrary amount of time.

We call $\Omega^{\pm}$ the *Moeller transformations*. They interpolate between free states (which are acted on by Φ_t^0) and interacting states (acted on by Φ_t). By cleverly composing these homeomorphisms, we obtain the *scattering transformation*, which encodes the date of the scattering experiment.

Definition 4.5 (Scattering transform)**.** Let V be a short range potential. We define the **scattering transformation**

$$S(x) = (\Omega^-)^{-1} \circ \Omega^+(x) = \lim_{t\to\infty} \left(\Phi^0_{-t} \circ \Phi_{2t} \circ \Phi^0_{-t}(x)\right)$$

The asymptotic impact parameters and directions can be computed directly from the scattering transform. Moreover, it knows about the *time delay*, which effectively measures the time it takes for the particle to interact with the potential. These data together (d impact parameters, $d-1$ directions and 1 time) give complete information on the scattering experiment. For more details, see [6, Section 10.4].

Theorem 4.6. *(Compact support) Let V be compactly supported. Then, the Moeller transformations are smooth, and they preserve the symplectic form.*

4.3 Degree theory

We can use the material from the previous subsection to introduce a topological invariant which characterises the scattering in a classical potential at a fixed *nontrapping* energy $E > 0$. Recall that we have assumed that V is at least long range, C^2 and that $\lim_{\|q\|\to\infty} V(q) = 0$.

Recall the definition of trapped states from the previous subsection. We say that an energy E is *nontrapping* if $t_E = \emptyset$. The set of nontrapping energies is denoted by $\mathcal{NT}$. For such energies, the unbounded trajectories asymptotically look like straight lines in both forward and backward time.

We will define a cousin of the scattering transformation which effectively maps these straight lines to each other. The nontrapping condition is important here to ensure that our map will be defined everywhere. Recall the maps $\hat{p}^{\pm} = \frac{p^{\pm}}{\|p^{\pm}\|} : s^{\pm} \to \mathbb{S}^{d-1}$ and $q^{\pm}_{\perp} : s^{\pm} \to \mathbb{R}^d$. By construction, the asymptotic impact parameter is orthogonal to the asymptotic direction. Hence, the combination $(q^{\pm}_{\perp}, \hat{p}^{\pm})$ defines a point in the cotangent bundle $T^*\mathbb{S}^{d-1}$.

Finally, note that if x_0 and y_0 are scattering states and if $y_0 = \Phi_s(x_0)$ then they have the same asymptotics. This follows simply by shifting all the limits by s. Hence, if two points are on the same orbit, then $(q^{\pm}_{\perp}, \hat{p}^{\pm})$ has the same value on them. We will thus mod out by the action of $\Phi_{t\in\mathbb{R}}$ to obtain a bijection.

Theorem 4.7. *[7, p. 6] Let $E \in \mathcal{NT}$. The maps $A^{\pm}_E : s^{\pm}_E/\Phi_{\mathbb{R}} \to T^*\mathbb{S}^{d-1}$ are well-defined homeomorphisms.*

For a nontrapping energy, we define the map

$$S_E = (Q_E, \hat{P}_E) = A^+_E \circ (A^-_E)^{-1} : T^*\mathbb{S}^{d-1} \to T^*\mathbb{S}^{d-1}.$$

Intuitively speaking, this map is related to the scattering transformation S that we defined in the previous subsection by a similarity transformation. We can understand S' as $S' = \Omega^+ \circ (\Omega^-)^{-1}$, which is a map from in-states to out-states. Then, we can write $S' = \Omega^- \circ S \circ (\Omega^-)^{-1}$, and the maps are indeed similar.

The map $\hat{P}_E : T^*\mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$, which is the composition of S_E with the canonical projection, is continuous by the previous theorem. We consider its restriction to a single cotangent space $\hat{P}_{E,\theta} : T^*_\theta \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$. This map gives a relation between impact parameter $q_\perp$ as measured w.r.t. some reference angle θ and the corresponding asymptotic direction.

Lemma 4.8. *[7, Lemma 2.5] We have*

$$\lim_{\|q_\perp\| \to \infty} \hat{P}_{E,\theta} = \theta.$$

Thus, the map $\hat{P}_{E,\theta}$ extends uniquely to a continuous function $\hat{\mathbf{P}}_{E,\theta} : T^*_\theta \mathbb{S}^{d-1} \cup \{\infty\} \cong \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$.

For maps from the $(d-1)$-sphere to itself, we can define the *(topological) degree*. This is a concept from algebraic topology which generalises the winding number. We define $\deg(E) := \deg(\hat{\mathbf{P}}_{E,\theta})$. Since the map $\hat{P}_{E,\theta}$ is continuous in both its input $q_\perp$ and its parameter θ, the degree of E does not depend on the choice of reference direction θ. Moreover, it can be shown that the map

$$\deg : \mathcal{NT} \to \mathbb{Z}$$

is well-defined and locally constant on the set of nontrapping energies. [7, p. 7]

In general, the topological degree is defined for continuous maps from the sphere $\mathbb{S}^d$ to itself. Any continuous map $f : \mathbb{S}^d \to \mathbb{S}^d$ induces a group homomorphism f_* from d-th the reduced homology group $\tilde{H}_d(\mathbb{S}^d; A)$ to itself. Here, A is the Abelian group in which we take the coefficients for singular homology. There is a short introduction to algebraic topology in the appendix. Alternatively, one can look at the sources [25, 26].

Taking $A = \mathbb{Z}$, it can be shown that

$$\tilde{H}_d(\mathbb{S}^d; \mathbb{Z}) \simeq \mathbb{Z}.$$

Since every homomorphism from $\mathbb{Z}$ to itself is determined by where it sends the generator $1 \in \mathbb{Z}$, we can associate to each continuous map $f : \mathbb{S}^d \to \mathbb{S}^d$ a unique integer $\deg(f)$ which satisfies

$$f_*(1) = \deg(f)$$

and hence

$$f_*(a) = a \cdot \deg(f) \quad \forall a \in \tilde{H}_d(\mathbb{S}^d; \mathbb{Z}).$$

Lemma 4.9. *[25, Lemma 6.6] Let $f, f' : \mathbb{S}^d \to \mathbb{S}^d$ be continuous. Then,*

- *If f and f' are homotopic, then they have the same degree;*
- $\deg(f' \circ f) = \deg(f') \cdot \deg(f)$*;*

- *If f is a homotopy equivalence, then* $\deg(f)$ *is either* 1 *or* -1.

The above concept of degree is the most basic one. If we assume some additional structure on our topological spaces, then we can extend the concept quite a bit.

Theorem 4.10. *[27, Theorem 17.35] Suppose that M and N are compact connected oriented smooth manifolds of the same dimension n, and let $F : M \to N$ be a smooth map. Then, there is a unique integer* $\deg(F)$ *such that*

- *For every smooth n-form ω on N*

$$\int_M F^*\omega = \deg(F) \int_N \omega \tag{4.7}$$

- *If $q \in N$ is a regular value of F then*

$$\deg(F) = \sum_{x \in F^{-1}(q)} \operatorname{sgn}(x), \tag{4.8}$$

  where $\operatorname{sgn}(x)$ *is* 1 *if* $\mathrm{d}F_x$ *is orientation-preserving and* -1 *if it is orientation-reversing.*

Instead of singular homology, the above theorem uses knowledge from the de Rham cohomology. Specifically, if M is compact connected and orientable, then the highest degree cohomology group $H^n(M)$ is one-dimensional. Like the previous case, this means everything is determined by the action on a generator.

There are some cases in which we can explicitly work out the degree. We collect a number of results in the Theorem below.

Theorem 4.11. *Let $d = 2$.*

- *Let V be smooth and short range. Then, there is an energy threshold $E_{\mathcal{NT}}$ above which the motion is nontrapping. For $E \geq E_{\mathcal{NT}}$, we have* $\deg(E) = 0$. *[8, Proposition 2]*

- *For a centrally symmetric short range potential, if* $0 < E < V_{\max}$ *is nontrapping, then* $\deg(E) = 1$. *[8, Proposition 4]*

- *For $V(q) = \frac{-1}{|q|^\alpha}$ defined on $\hat{M} = \mathbb{R}^2 \setminus \{0\}$, the motion can be regularised if $\alpha = \frac{2n}{n+1}$ for $n \in \mathbb{N}$. In this case, all positive energies are nontrapping and* $\deg(E) = -n$. *[8, Proposition 3]*

We will look at parts of the proof in the next subsection.

Finally, we will look at an important application of degree theory to chaotic scattering. We consider the case of multiple compactly supported potentials. Suppose that V can be written in the form $V = \sum_{i=1}^N V_i$, where each V_i is smooth and supported in some ball

$$B_i = \{x \in \mathbb{R}^2 : \|x - s_i\| < R_i\}$$

We will assume that the supports are disjoint, and that they don't *shadow* each other, meaning that no straight line meets more than two of the B_i. We finally assume that

E is nontrapping for each of the individual potentials V_i and that $\deg(V_i) \neq 0$ for all $i = 1, 2, \ldots, N$. Under these assumptions, we can introduce symbolic dynamics based on the order in which a trajectory meets the balls.

The letters of the symbolic code are just $1, 2, \ldots, N$. We note that a sequence is only admissible if the same number never occurs twice in a row. We have the following:

Theorem 4.12. *[8, Theorem 2] Let $N \geq 2$. Under the above assumptions, let $J_l^r = \{j \in \mathbb{Z} : l \leq j \leq r\}$. For every admissible sequence (k_j) and every $\hat{p}^{\pm} \in \mathbb{S}^1$ there is a trajectory with energy E which meets the balls B_j precisely in the order prescribed by (k_j). Moreover,*

- *If $l \neq -\infty$ then this trajectory in s_E^- has initial direction $\hat{p}^-$. Otherwise it's in b_E^-.*
- *If $r \neq \infty$ then this trajectory in s_E^+ has final direction $\hat{p}^+$. Otherwise it's in b_E^+.*

In particular, E is a trapping energy for V.

In Section 6, we will apply this theorem to the relativistic N-center problems.

4.4 Central potentials

Let's now see some of these concepts in action. To minimise the amount of technical details, we consider a potential $V : \mathbb{R}^2 \to \mathbb{R}$ which is defined everywhere, C^2 and rotationally symmetric. We can thus write $V(x, y) = V(r)$. Moreover, I will assume that V is short-range. We will first review some basic classical mechanics, just for completeness. The main source is [21].

4.4.1 Solving the equations of motion

The Lagrangian is given by

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - V(r) = \frac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) - V(r).$$

Let us go to the Hamiltonian formalism. The conjugate momenta are

$$p_r = \frac{\partial L}{\partial \dot{r}} = \dot{r}, \quad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = r^2\dot{\phi}.$$

Notice that I have set the mass equal to 1 for convenience. The Hamiltonian is given by the Legendre transform of L. Indeed,

$$H = \frac{1}{2}p_r^2 + \left(\frac{p_\phi^2}{2r^2} + V(r)\right) = \frac{1}{2}p_r^2 + V_\ell(r). \tag{4.9}$$

We have two conserved quantities, namely H and p_ϕ, since the Hamiltonian is independent of t and ϕ. We will denote the values of these conserved quantities by E and ℓ respectively. Since there are as many conserved quantities as degrees of

freedom, the system is integrable. This essentially means we can write down the solution to the system by using a finite number of integrals. Let us try to do this explicitly.

First of all, notice that

$$p_r = \pm\sqrt{2(E - V_\ell(r))}, \tag{4.10}$$

where the particle either moves "inward" or "outward" depending on the sign of p_r. There are *turning points* for $E = V_\ell(r)$. We visualise this as a ball rolling in the graph of $V_\ell(r)$.

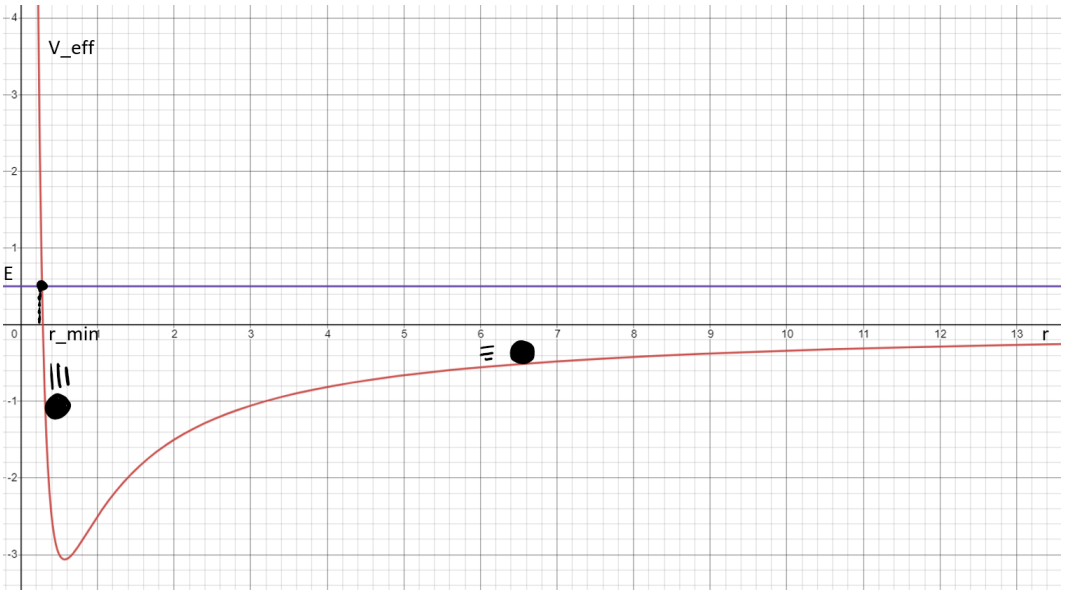


Figure 7: Effective potential for the Kepler problem. For $E < 0$, there are two turning points, and the motion is bounded. For $E \geq 0$ there is just one turning point, and the motion will be unbounded.

Consider a particle that comes in from infinity with some positive energy $E > 0$. Since the effective potential V_ℓ tends to zero at infinity, there will be one turning point $r_{\min}$, which is the largest positive solution to $V_\ell(r) = E$. The particle first moves inward to the turning point, and then outwards to infinity. To determine the time dependence, we can integrate:

$$\frac{\mathrm{d}r}{\mathrm{d}t} = p_r = \pm\sqrt{2(E - V_\ell(r))} \implies \int \mathrm{d}t = \int \frac{\pm\,\mathrm{d}r}{\sqrt{2(E - V_\ell(r))}}$$

For the inward moving part this gives

$$t = \int_r^\infty \frac{\mathrm{d}r'}{\sqrt{2(E - V_\ell(r'))}}, \tag{4.11}$$

until r reaches $r_{\min}$. This gives a function $t(r)$, which we can in principle invert to find r as a function of t. The outward moving part of the motion is solved similarly. We solve the angular part next. Since $p_\phi = \ell$ is constant, we can solve

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\ell}{r^2}.$$

We have already determined r as a function of t, but we would like to avoid writing an implicit formula. Instead, we can use the chain rule to first solve ϕ as a function of r:

$$\frac{\mathrm{d}\phi}{\mathrm{d}r} = \frac{\ell}{r^2}\frac{\mathrm{d}t}{\mathrm{d}r} = \frac{\frac{\ell}{r^2}}{\pm\sqrt{2(E - V_\ell(r))}}$$

For the inward moving part of the motion, we have

$$\phi - \phi_0 = \int_r^\infty \frac{\frac{\ell}{r'^2}\,\mathrm{d}r'}{\sqrt{2(E - V_\ell(r'))}} \tag{4.12}$$

Here, ϕ_0 is the angle at which the particle comes in from infinity, so as $r \to \infty$ we have $\phi(r) \to \phi_0$. This equation again is valid until r reaches $r_{\min}$. At that point, we will have

$$\phi - \phi_0 = \int_{r_{\min}}^\infty \frac{\frac{\ell}{r'^2}\,\mathrm{d}r'}{\sqrt{2(E - V_\ell(r'))}} + \int_{r_{\min}}^r \frac{\frac{\ell}{r'^2}\,\mathrm{d}r'}{\sqrt{2(E - V_\ell(r'))}}.$$

Like before, this implicitly determines ϕ as a function of r. If necessary, we would have to solve equation (4.11) for $r(t)$. Plugging this into equation (4.12) gives ϕ as a function of t. At this point, r, ϕ, p_r and p_ϕ are all known functions of t and the initial conditions.[14]

4.4.2 Scattering angle and degree

Recall the definition of $\hat{P}_{E,\theta}$ from the previous subsection. This function maps the asymptotic impact parameter as $t \to -\infty$ w.r.t. reference angle θ to the asymptotic direction as $t \to \infty$. Using the integrals from the previous subsection, it can be shown [8] that

$$\hat{P}_{E,\theta} = \theta - \Delta\phi(E, \sqrt{2E}q_\perp), \tag{4.13}$$

where

$$\Delta\phi(E, \ell) = 2\int_{r_{\min}}^\infty \frac{\frac{\ell}{r^2}}{\sqrt{2(E - V_\ell(r))}}\,\mathrm{d}r - \pi \tag{4.14}$$

To compute the degree of $\hat{P}_{E,\theta}$, we will use Theorem 4.10. We take $\omega = \mathrm{d}\phi$ to be the standard volume form on the circle, so that $\int_N \omega = 2\pi$. To compute the degree of $\hat{P}_{E,\theta}$, we need the pull-back:

$$\int_{\mathbb{S}^1} \hat{P}^*_{E,\theta}(\mathrm{d}\phi) = \int_{\mathbb{S}^1} \mathrm{d}\left(\phi \circ \hat{P}_{E,\theta}\right)$$

We use $q_\perp$ are a coordinate for $\mathbb{S}^1$, which hits all points except the one at infinity. Of course, this one point cannot change the value of the integral. Then, we obtain

$$\int_{\mathbb{S}^1} \hat{P}^*_{E,\theta}\omega = \int_{-\infty}^\infty \frac{\partial}{\partial q_\perp}\hat{P}_{E,\theta}\,\mathrm{d}q_\perp = -\int_{-\infty}^\infty \frac{\partial}{\partial q_\perp}\Delta\phi(E, \sqrt{2E}q_\perp)\,\mathrm{d}q_\perp\,.$$

Changing variables to $\ell = \sqrt{2E}q_\perp$ gives

$$\int_{\mathbb{S}^1} \hat{P}^*_{E,\theta}\omega = -\int_{-\infty}^\infty \frac{\partial}{\partial \ell}\Delta\phi(E, \ell)\,\mathrm{d}\ell = -2\int_0^\infty \frac{\partial}{\partial \ell}\Delta\phi(E, \ell)\,\mathrm{d}\ell\,,$$

[14] We note that E and ℓ are also uniquely determined by the initial conditions.

where we used that $\frac{\partial(\Delta\phi)}{\partial\ell}$ is parity-even by definition. Thus, we obtain for the degree

$$\deg(E) = \frac{\int_{\mathbb{S}^1} \hat{P}^*_{E,\theta}\omega}{\int_N \omega} = \frac{-2\int_0^\infty \frac{\partial}{\partial\ell}\Delta\phi(E,\ell)\,\mathrm{d}\ell}{2\pi} = -\frac{1}{\pi}\left[\Delta\phi(E,\ell)\right]_{\ell=0}^{\ell=\infty}.$$

For convenience, let me record this intermediate result as a Theorem:

Theorem 4.13. *For a planar centrally symmetric potential, the scattering degree is given by*

$$\deg(E) = -\frac{1}{\pi}\left[\Delta\phi(E,\ell)\right]_{\ell=0}^{\ell=\infty}. \tag{4.15}$$

Recall that the effective potential is given by $V_\ell(r) = \frac{\ell^2}{2r^2} + V(r)$, where V is short-range. We substitute $v = \frac{r_{\min}}{r}$ inside the integral defining $\Delta\phi$. This gives

$$\begin{aligned}
\Delta\phi &= 2\int_0^1 \frac{\ell\frac{v^2}{r_{\min}^2}\left(-\frac{1}{r_{\min}}\frac{r_{\min}^2}{v^2}\right)}{\sqrt{2\left(E-\frac{\ell^2}{2r^2}-V(r)\right)}}\,\mathrm{d}r - \pi \\
&= 2\int_0^1 \frac{\frac{\ell}{r_{\min}}}{\sqrt{2\left(E-\frac{\ell^2v^2}{r_{\min}^2}-2V\left(\frac{r_{\min}}{v}\right)\right)}}\,\mathrm{d}v - \pi \\
&= 2\int_0^1 \frac{\mathrm{d}v}{\sqrt{\frac{2r_{\min}^2}{\ell^2}\left(E-V\left(\frac{r_{\min}}{v}\right)\right)-v^2}} - \pi.
\end{aligned}$$

We need to understand the limiting behaviour of this integral as ℓ tends to zero and infinity. To see where we are going, consider the following result:

Theorem 4.14. *[28, Theorem 2] Let V be a planar centrally symmetric and short-range potential. Then,*

$$\deg(E) = \begin{cases} 0, & E > V_{\max} \\ 1, & 0 < E < V_{\max} \end{cases} \tag{4.16}$$

where $V_{\max}$ is the maximum value of V.

The authors of [28] use a homotopy argument, which is much more in line with algebraic topology. Instead, we shall prove the above result for $E < V_{\max}$ using the above integral equations. Let us consider only the second case, where $E < V_{\max}$.

Proof. We have to take the limit for ℓ going to zero and infinity respectively of the above integral in $\mathrm{d}v$. In order to switch limit with integral, we need to be sure that the integral converges for all values of ℓ. Recall that $r_{\min}$ is the smallest solution to the equation

$$V_\ell(r) = E \implies V(r) + \frac{\ell^2}{2r^2} = E. \tag{4.17}$$

For small values of v, i.e. large r, the integral looks like

$$\int_0^\epsilon \frac{\mathrm{d}v}{\sqrt{\frac{2r_{\min}^2}{\ell^2}\left(E-V\left(\frac{r_{\min}}{v}\right)\right)-v^2}} \sim \int_0^\epsilon \frac{\mathrm{d}v}{\sqrt{\frac{2r_{\min}^2}{\ell^2}E}} \sim \epsilon\frac{\ell}{r_{\min}(E,\ell)}\frac{1}{\sqrt{2E}}.$$

For small ℓ, $r_{\min}(E,\ell)$ converges simply to the smallest solution to $V(r) = E$, which is a finite number since $E < V_{\max}$. Hence, $\ell/r_{\min}$ cannot blow up. Conversely for large ℓ, notice that $r_{\min} \to \infty$, since

$$r_{\min}^2(E - V(r_{\min})) = \frac{\ell^2}{2}.$$

The right-hand side of this equation goes to infinity, while $E - V(r)$ is bounded. Moreover,

$$\frac{\ell^2}{2r_{\min}^2} = E - V(r_{\min}).$$

In order for the left-hand side to remain bounded, we must have that

$$r_{\min} \sim \frac{\ell}{\sqrt{2}}, \quad \ell \to \infty. \tag{4.18}$$

Thus, $\frac{\ell}{r_{\min}}$ is bounded for $\ell \to \infty$, and the integral converges.

Let us also check the behaviour close to $v = 1$, i.e. $r \sim r_{\min}$. Here, I prefer to use the form of the integral before substitution. Notice that

$$E - V_\ell(r_{\min} + \epsilon) = E - V_\ell(r_{\min}) - \epsilon V_\ell'(r_{\min}) + O(\epsilon^2) = -\epsilon V_\ell'(r_{\min}) + O(\epsilon^2),$$

and that $V_\ell'(r_{\min}) < 0$, since otherwise $r_{\min}$ would not be the largest solution. Hence,

$$\int_{r_{\min}}^{r_{\min}+\epsilon} \frac{\frac{\ell}{r^2}\,\mathrm{d}r}{\sqrt{2\,(E - V_\ell(r))}} \sim \int_{r_{\min}}^{r_{\min}+\epsilon} \frac{\frac{\ell}{r^2}\,\mathrm{d}r}{\sqrt{-2\epsilon V_\ell'(r_{\min})}} \sim \epsilon \frac{\ell}{r_{\min}^2} \frac{1}{\sqrt{\epsilon}} \frac{1}{\sqrt{-2V_\ell'(r_{\min})}}.$$

As before, this converges for $\epsilon \to 0$ independent of ℓ. Thus, we can safely take the limit for both large and small ℓ inside the integral.

First, we compute the limit for $\ell \to \infty$. Because of equation (4.18), we get

$$\lim_{\ell\to\infty} \frac{1}{\sqrt{\frac{2r_{\min}^2}{\ell^2}\left(E - V\left(\frac{r_{\min}}{v}\right)\right) - v^2}} = \frac{1}{\sqrt{1-v^2}},$$

and thus

$$\lim_{\ell\to\infty} \Delta\phi = 2\int_0^1 \frac{1}{\sqrt{1-v^2}} - \pi = 2 \cdot \frac{\pi}{2} - \pi = 0,$$

which is consistent with Lemma 4.8. As $\|q_\perp\|$ becomes large, $\Delta\phi$ tends to zero and thus $\hat{P}_{E,\theta}$ tends to θ. Finally

$$\lim_{\ell\to 0} \Delta\phi = \lim_{\ell\to 0} \left(2\ell \cdot \int_{r_{\min}}^{\infty} \frac{\frac{1}{r^2}\,\mathrm{d}r}{\sqrt{2(E - W_\ell(r))}} \right) - \pi = -\pi,$$

because the integral is bounded, while $\ell \to 0$. Thus, the final result for the scattering degree is

$$\deg(E) = -\frac{1}{\pi}\left[\Delta\phi\right]_{\ell=0}^{\ell=\infty} = -\frac{1}{\pi}\,(0 + \pi) = -1,$$

□

This is the end of the proof, but unfortunately the result is off by a minus sign. I will comment on this discrepancy, which is probably due to orientation, at the end of the next proof.

4.4.3 The regularizable case

We can also use Theorem 4.13 to prove another part of Theorem 4.11, namely the one pertaining to potentials of the form $V \propto -\frac{1}{r^\alpha}$. To this end, we need to compute

$$\lim_{\ell \to 0} \int_{r_{\min}}^{\infty} \frac{\frac{\ell}{r^2}\,\mathrm{d}r}{\sqrt{2\left(E + \frac{1}{r^\alpha} - \frac{\ell^2}{2r^2}\right)}}.$$

Following Knauf [8], we substitute

$$v = \frac{\left(\frac{\ell}{\sqrt{2}}\right)^{\frac{2}{2-\alpha}}}{r}.$$

Some tedious algebra then gives

$$E - V_\ell(r) = E - v^2 \left(\frac{\ell}{\sqrt{2}}\right)^{\frac{-2\alpha}{2-\alpha}} + v^\alpha \left(\frac{\ell}{\sqrt{2}}\right)^{\frac{-2\alpha}{2-\alpha}}$$

and

$$\frac{\ell}{r^2}\,\mathrm{d}r = -\ell \left(\frac{\ell}{\sqrt{2}}\right)^{\frac{-2}{2-\alpha}} \mathrm{d}v\,.$$

Filling this into the integral then gives

$$\Delta\phi(E,\ell) = 2\int_0^{v_{\max}} \frac{\mathrm{d}v}{\sqrt{E\ell^{\frac{2\alpha}{2-\alpha}} 2^{\frac{-\alpha}{2-\alpha}} - v^2 + v^\alpha}} - \pi,$$

where $v_{\max}$ is the smallest positive solution to

$$E\ell^{\frac{2\alpha}{2-\alpha}} 2^{\frac{-\alpha}{2-\alpha}} - v^2 + v^\alpha = 0,$$

similar to $r_{\min}$. As shown before, it is possible to exchange the limit $\ell \to 0$ with the integral. Upon doing this, the equation for $v_{\max}$ simply gives $v_{\max} = 1$, while the integrand becomes

$$\lim_{\ell \to 0} \Delta\phi = 2\int_0^1 \frac{\mathrm{d}v}{\sqrt{v^\alpha - v^2}} - \pi = \frac{2\pi}{2-\alpha} - \pi$$

Solving the above integral is a fun challenge to the reader. (I used a trigonometric substitution.) In the regularizable case with $\alpha = \frac{2n}{n+1}$, this becomes $n\pi$. Thus, from Theorem 4.13 we obtain

$$\deg(E) = -\frac{1}{\pi}\left[\Delta\phi\right]_{\ell=0}^{\ell=\infty} = -\frac{1}{\pi}\left(0 - (-n\pi)\right) = -n,$$

which is what I wanted to prove.

Here, the same formula for the scattering degree *does* give the correct result. I am confident that both results from [8] are correct, so the mistake must lie elsewhere. Think about it in the following way: look at Figure 8. In the first case, particles are reflected off the potential and ϕ goes from increasing to decreasing. In the second case particles swing around the center, and ϕ continuously increases. This difference does not seem to be reflected in the equation for $\Delta\phi$, and it comes with a change in orientation. We believe that this is where the mistake lies.

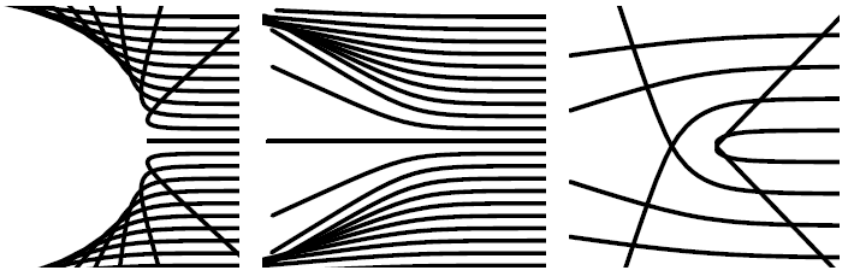


Figure 8: Taken from [9, Figure 1]. Left: $\deg(E) = 1$. Middle: $\deg(E) = 0$. Right: $\deg(E) = -1$.

5 Black holes and extremality

We have come a long way in developing mathematical tools to analyse scattering of light in relativistic N-center systems. At this point, let us take a step back and look at some physics. We will first have a closer look at the role of scalar fields in gravity. To this end, we will study Nordström's theory of gravity, which only involved a scalar ϕ and no metric tensor. These fields appear naturally in the low energy limit of string theories. Looking at a simple case first will hopefully make them seem less intimidating when we return to studying the dilaton.

After this introduction, we will consider two ways to look at physical properties of black holes, namely their thermodynamics and their Penrose diagrams. Both of these methods will be applied to our static solution to the Einstein-dilaton-Maxwell theory. Finally, we will have a closer look at the notion of extremality. Extremal solutions are highly symmetric, which often simplifies computations. We will see this in action later.

5.1 Scalar gravity

Scalar fields have been involved in developing a relativistically correct theories of gravity from the very beginning. The first gravitational theories with a (massless) scalar field were developed by Norwegian physicist Gunnar Nordström a couple years before Einstein wrote down general relativity. Nordström was also the first to use an extra dimension to model gravitational phenomena. In many modern theories - e.g. string theory - extra dimensions are needed for consistency. Hence, Nordström's theory was revolutionary in two ways. In this subsection, we will use his theory to build some intuition for the dilaton field, and how it arises from reductions of higher dimensional theories.

5.1.1 Nordström's theory

First, let us review how classical (i.e. without relativistic effects) gravity can be viewed as a field theory. We introduce a scalar field ϕ, which plays the role of the gravitational potential. The Lagrangian density is given by [29]

$$L = -\frac{1}{8\pi G}(\nabla\phi)^2 - \rho\phi, \tag{5.1}$$

where $\rho(\mathbf{x})$ is the mass density. We can easily determine that the corresponding equation of motion is

$$\nabla^2\phi = 4\pi G\rho, \tag{5.2}$$

which is Poisson's equation. If the right-hand side is chosen as a delta function

$$\rho(\mathbf{x}) = m\delta(\mathbf{x} - \mathbf{x}_0),$$

then it can be proven - for example by going to momentum space - that

$$\phi = -\frac{Gm}{|\mathbf{x} - \mathbf{x}_0|}, \tag{5.3}$$

so the theory recovers the gravitational potential due to a point mass.

Nordström's initial idea to make the theory relativistic was to replace the equation of motion by

$$\Box\phi = -4\pi G\rho, \tag{5.4}$$

where $\Box = \partial_t^2 - \nabla^2$. This is now a wave equation, and hence the finite speed of light is built in. To complete the theory, we need an equation of motion for particles moving in the field ϕ. For this, Nordström came up with

$$\dot{u}_\mu + \dot{\phi}u_\mu = \partial_\mu\phi, \tag{5.5}$$

where $u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$ is the four-velocity. This looks nice, but several problems with the theory were pointed out:

- Einstein said that the density should be proportional to the trace of an energy-momentum tensor.
- The field equation cannot be derived from an action principle.
- The field equation is linear. This is inconsistent with the fact that energy and mass are equivalent through $E = mc^2$. Hence, energy should be able to gravitate and we need a nonlinear equation.
- There is a problem with the strong equivalence principle. Consider the particle equation of motion. It satisfies the weak equivalence principle, since the motion does not depend on m. However, the gravitational mass and inertial mass will be different. Indeed, it can be proven [29, p. 4] that the inertial mass will depend exponentially on the field.

Based on these criticisms, Nordström went back to the drawing board. According to the first bullet point, Nordström replaced ρ by the trace T_m of $T_{\mu\nu}$, i.e. $\rho = g(\phi)T_m$. Here, $g(\phi)$ is some proportionality constant, which can be determined from the equivalence principle. This lead to $g(\phi) = 1/\phi$, and thus to the new and improved field equation

$$\phi\Box\phi = -4\pi GT_m, \tag{5.6}$$

which is nonlinear. It is worth noting that the electromagnetic energy-momentum tensor is traceless. Hence, this theory cannot explain the bending of light by black

holes. Luckily for Nordström, this experimental fact was not known at the time, so his theory survived for a bit longer.

Another improvement of the second theory is that the particle equation of motion can be determined from the variation of

$$S = \int \phi(x^\mu(\lambda))\sqrt{u^\mu(\lambda)u_\mu(\lambda)}\,\mathrm{d}\lambda\,.$$

This leads to the equation

$$\phi\dot{u}_\mu + \dot{\phi}u_\mu = \partial_\mu\phi. \tag{5.7}$$

Again, this satisfies the weak equivalence principle, since the motion does not depend on the mass m.

Nordström's theories are not the only possible gravitational theories based on scalar fields [29, p. 6]. One could consider an action of the form

$$S = -m\int A(\phi)\sqrt{u_\mu(\lambda)u^\mu(\lambda)}\,\mathrm{d}\lambda\,, \tag{5.8}$$

where $A(\phi)$ is some function to be determined. The only requirement is that $A(\phi)$ tends to $\phi + C$ as ϕ becomes small. This ensures that the theory has the right limit when gravity is weak. Note that Nordström's first theory has $A(\phi) = e^\phi$, while the second has $A(\phi) = \phi$.

Unfortunately, none of these theories can predict the deflection of light by mass. Moreover, they do not predict the correct perihelion shift for the orbit of mercury. In particular, if $A(\phi) = \phi$, the shift $\delta\theta$ of the perihelion has the opposite sign as in Einstein's general relativity. For these two reasons, you have probably never heard of Nordström's theories. They were overruled by experiment.

5.1.2 Gravity and extra dimensions

After developing his second scalar field theory of gravity, Nordström wanted to unify it with electromagnetism. This is where another revolutionary idea came to him: using the concept of extra spacetime dimensions. In modern theories, such as string theory, this has become commonplace. However, at the start of the 1900s, this was a completely new concept.

Consider the Maxwell Lagrangian coupled to a source term

$$L = -\frac{1}{4}F_{ab}F^{ab} - J_aA^a, \tag{5.9}$$

where $a = 0, 1, 2, 3, 5$. We have introduced a fifth dimension which is labelled by the number 5. This is a fossil from when we used to label time by 4, i.e. $x^\mu = (x, y, z, t)$. As usual, the field strength tensor is given by $F_{ab} = \partial_aA_b - \partial_bA_a$. The components of A^a are $(A^\mu, \phi/\sqrt{4\pi G})$ and $J^a = (J^\mu, \rho\sqrt{4\pi G})$. In this way, the theory incorporates the gravitational potential and mass density.

The Euler-Lagrange equations of this theory are given by [29, p. 8]

$$\partial_b \partial^b A_a - \partial_a(\partial^b A_b) = -J_a. \tag{5.10}$$

Let us impose current conservation $\partial_\mu J^\mu = 0$ for the ordinary 4-current. Acting with ∂^a on both sides of the equation of motion gives:

$$\begin{aligned} -\partial^5 J_5 &= -\partial^a J_a \\ &= \partial^a \partial_b \partial^b A_a - \partial^a \partial_a \partial^b A_b \\ &= 0 \end{aligned}$$

by interchanging $a \leftrightarrow b$ in one of the terms and by using equality of mixed partials. It follows that the mass density is not allowed to depend on the fifth dimension. If one imposes that $\partial_5 J^a = 0$ for all five components of the source, then the equation of motion separates into the standard Maxwell equations together with the original wave equation (5.4). This condition is called the *cylinder condition* in modern language.

Before, we have written down Maxwell's theory with extra dimensions to include gravitational effects. One might also wonder what happens if we write down Einstein gravity using extra dimensions. This was done by Kaluza and Klein at the start of the 1920s. Consider n extra dimensions labelled by y^α, $\alpha = 1, 2, \ldots, n$. The Einstein-Hilbert action for GR is written as

$$S = -\frac{1}{2} M_D^{2+n} \int \mathrm{d}^4 x \int \mathrm{d}^n y \, \sqrt{-\bar{g}} \bar{R}, \tag{5.11}$$

where M_D is the Planck mass in $D = 4+n$ dimensions. The power $2+n$ follows from dimensional analysis. At this point, Kaluza assumed that the full metric separates into

$$\mathrm{d}\bar{s}^2 = g_{\mu\nu} \, \mathrm{d}x^\mu \, \mathrm{d}x^\nu - b^2(x) h_{\alpha\beta} \, \mathrm{d}y^\alpha \, \mathrm{d}y^\beta \, . \tag{5.12}$$

Here, $b^2(x)$ is a conformal factor which roughly describes the size of the extra dimensions and $h_{\alpha\beta}$ describes their geometry. Let $V_n = \int \mathrm{d}^n y \sqrt{-h}$ and R_n the scalar curvature of the metric $h_{\alpha\beta}$, which we assume to be constant. On these assumptions, it can be proven that [29, p. 9]

$$S = -\frac{1}{2} M_D^{2+n} V_n \int \mathrm{d}^4 x \, \sqrt{-g} b^n \left[R - \frac{1}{b^2} R_n + n(n-1) g^{\mu\nu} \partial_\mu b \partial_\nu b \right] \tag{5.13}$$

Hence, we see that the "regular" Planck mass is given by

$$M_P^2 = V_n M_D^{2+n}. \tag{5.14}$$

In fact, this equation can also be derived from much more general considerations. The fact that we have a $1/r$-potential follows eventually from rotational symmetry. Recall that the potential should solve an equation like

$$\nabla^2 V = -\rho,$$

with ρ being the mass density. Going to momentum space, the Laplacian turns into k^2. Hence, we find [30, p. 41]

$$V(r) \propto \int \mathrm{d}^{3+n} k \left(e^{i \mathbf{k} \cdot \mathbf{x}} \frac{1}{k^2} \right) \propto \frac{1}{r^{1+n}},$$

where $\frac{1}{k^2}$ is the massless propagator in momentum space. It looks like this immediately contradicts experiment. However, we assume that the extra dimensions are compactified and have characteristic scale R. This new and improved potential $V(r)$ is only accurate for $r \ll R$, where all dimensions are felt.

We can make this more precise. In three spatial dimensions, Newton's law can be written as

$$V(r) = \frac{Gm_1m_2}{r} = \frac{m_1m_2}{M_P^2} \cdot \frac{1}{r}.$$

The Planck mass $M_P \sim 10^{19}$ GeV, which is a huge number. The fact that this is so large compared to the other three fundamental forces is sometimes called the *hierarchy problem* in theoretical physics. The introduction of extra dimension is a proposed solution to this problem. Indeed, if we have n extra dimensions, the potential becomes

$$V(r) = \frac{m_1m_2}{M_D^{2+n}} \cdot \frac{1}{r^{1+n}}$$

in the regime $r \ll R$. For everyday physics, we have $r \gg R$, and we obtain

$$V(r) = \frac{m_1m_2}{M_D^{2+n}} \cdot \frac{1}{R^n} \cdot \frac{1}{r}$$

Comparing coefficients gives

$$M_D^2 = \frac{M_P^2}{(M_DR)^n}, \tag{5.15}$$

which is equivalent to equation (5.14). In this way, if M_DR is a large number, the "true" scale of gravity M_D^2 can be much lower than M_P^2, hence solving the hierarchy problem. Based un current accelerator physics, we can already rule out $n = 1$. However, $n > 2$ is still possible based on experiment [30, p. 42].

Returning back to Kaluza's theory: we can set $\phi = b^n$ to obtain

$$S = -\frac{1}{2}M_P^2 \int \mathrm{d}^4x \, \sqrt{-g} \left[R\phi + (1 - 1/n)\frac{1}{\phi} g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - R_n\phi^{1-2/n} \right].$$

This action describes Einstein's gravity coupled to a scalar field ϕ. The last term can be interpreted as a potential $V(\phi)$. This scalar-tensor theory is related to Brans-Dicke theory [31].

One can also choose an alternative parametrization for the $(4 + 1)$-dimensional metric:

$$\mathrm{d}\bar{s}^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu + \phi^2\left(A_\mu\,\mathrm{d}x^\mu + \mathrm{d}x^5\right)^2. \tag{5.16}$$

If we assume the cylinder condition, i.e. that all components of the metric are independent of the fifth dimension, then the equations of motion separate into the 4-dimensional geodesic equation, the Lorentz force law and one equation for the scalar field ϕ. This way, we obtain the Einstein-Maxwell-dilaton theory as a reduction from 5-dimensional Einstein gravity.

5.2 Black holes and thermodynamics

Let us now turn to the properties of black hole solutions. First, we will consider the thermodynamic properties of black holes. In the 1970s, Bekenstein and Hawking argued that black holes should emit black body radiation with characteristic temperature [32]

$$kT_H = \frac{\hbar\kappa}{2\pi} \tag{5.17}$$

and entropy

$$S_H = \frac{A_{\text{hor}}}{4\hbar G}. \tag{5.18}$$

Here, κ is the surface gravity[15] of the black hole, while A_{hor} is the area of its event horizon. Since both the temperature and entropy contain $\hbar$ and G, it is generally assumed that Hawking radiation is in the domain of quantum gravity. The two equations above have been derived in many different settings, so most physicists believe that they are indeed fundamental to black holes [33].

There are some interesting physical puzzles associated with black hole thermodynamics. First of all, we typically define entropy via the equation

$$S = k_B \log \Omega, \tag{5.19}$$

where k_B is Boltzmann's constant and Ω is the number of microstates that represent the same macrostate. This suggests some quantum-gravitational way to count black hole microstates. We will get back to this puzzle later.

Secondly, the entropy of a thermodynamic system typically scales with its volume. However, for black holes, it scales with surface area. We can visualise that the information of a black hole is stored on the surface. This puzzle is sometimes referred to as the *holographic principle*. It could be that all of nature is fundamentally holographic; our universe being a shadow cast by a candle in some higher-dimensional space. Holography is related to very deep ideas in string theory, such as the *AdS-CFT correspondence*. We will not go into this, though a good starting point could be chapter IX.11 of [10].

Finally, let us discuss the *information loss paradox*. We have seen that black holes radiate thermally. There is no dependence on the *kind* of matter that fell into the black hole. This is a big problem, as it violates unitarity of quantum mechanics. In that theory, any process should, in principle, be reversible. Since all information about particles which fall into the black hole is seemingly destroyed, this principle is violated. Like the above puzzle, there are sophisticated proposals to resolve the paradox, which we will ignore in this work.

The starting point for black hole thermodynamics was Hawking's argument that the surface area of a black hole can never decrease. This is quite similar to the second law of thermodynamics, which states that the entropy of an isolated system always increases. This parallel eventually lead to the development of the four laws of black

[15] We will see later on what surface gravity means precisely for a black hole.

hole thermodynamics, which are valid for a static asymptotically flat black hole in four spacetime dimensions: [33, p. 2]

- The surface gravity κ is constant over the event horizon.
- For two stationary black holes differing only by small variations in M, J, Q,

$$\delta M = \frac{\kappa}{8\pi G}\delta A_{\text{hor}} + \Omega_H \delta J + \Phi_H \delta Q, \tag{5.20}$$

  where Ω_H is the angular velocity and Φ_H is the electric potential.
- The area of the event horizon of a black hole never decreases.
- It is impossible by any procedure to reduce the surface gravity of a black hole to zero in a finite number of steps.

We characterize the black hole by its mass M, charge Q and angular momentum J, consistent with the no-hair theorem. The parallels with the laws of classical thermodynamics are apparent. The surface gravity κ plays the role of temperature, and A_{hor} plays the role of entropy.

All these theoretical puzzles are nice, but they don't do us much good if we cannot even compute the Hawking temperature for a Schwarzschild black hole. Indeed, let us try to understand how to do this. Looking back at the defining equation (5.17), the only quantity we need to compute is the surface gravity κ. For a classical object, such as a planet, this is defined as the gravitational acceleration experienced by a person standing on the equator.

For the Earth, this can be easily computed from Newton's law of universal gravitation:

$$m\kappa = \frac{GmM}{R^2} \implies \kappa = \frac{GM}{R^2} \approx 9.8 \text{ m s}^{-2},$$

with M and R being the mass and radius of the Earth respectively. Fortunately, the Earth is not a black hole. Unfortunately, this concept breaks down for black holes, as black holes don't have a physical surface. We hence need to come up with a different way to define the surface gravity. I will follow the presentation of [34].

For a massive particle, we define the 4-acceleration to be the

$$a^\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}\lambda} + \Gamma^\mu_{\nu\rho} u^\nu u^\rho, \tag{5.21}$$

where $u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$ and $\Gamma^\rho_{\mu\nu}$ are the Christoffel symbols corresponding to the metric. The right-hand side of the above equation also occurs in the geodesic equation, and is zero for particles which follow a geodesic. In this way, a^μ is the correct relativistic generalisation of acceleration. We define the scalar acceleration through the norm:

$$a^2 = -a_\mu a^\mu.$$

Consider an observer hovering a distance r from the singularity of a black hole. For them, clearly $u_i = 0$ for the spatial components. The final component can be solved from the relation

$$u_\mu u^\mu = -1.$$

In case of the Schwarzschild metric, this becomes

$$u^\mu = \left(\left(1 - \frac{2GM}{r}\right)^{-\frac{1}{2}}, 0, 0, 0 \right).$$

It follows immediately that the only non-zero component of a^μ is given by

$$a^1 = \frac{\mathrm{d}u^1}{\mathrm{d}\lambda} + \Gamma^1_{00}(u^0)^2$$

Thus, for a Schwarzschild black hole, we find that

$$a = \frac{GM}{r^2}\left(1 - \frac{2GM}{r}\right)^{-\frac{1}{2}}, \tag{5.22}$$

which indeed goes to infinity at the event horizon.

To define the surface gravity of a black hole, we must hence cancel out this infinity by another one to produce a finite number. To this end, we imagine that the observer is attached to a long rope, which is held by a professor at infinity. Their goal is to make the observer - who conveniently has unit mass - hover at some distance r from the black hole. Imagine that the professor pulls slightly on the rope to raise the observer by some distance $\mathrm{d}\ell$. The observer will gain some potential energy $\mathrm{d}E_r = g_r\,\mathrm{d}\ell$. Here, g_r is the gravitational acceleration experienced at a distance r from the black hole. This is given by equation (5.22). In the process, the professor uses some energy $\mathrm{d}E_\infty = g_\infty\,\mathrm{d}\ell$. We want to know how much force g_∞ the professor needs to exert in order to keep the observer hovering at the horizon.

Because the professor is very smart, they immediately understand that [34, p. 44]

$$\frac{g_r}{g_\infty} = \frac{E_r}{E_\infty} = \left(1 - \frac{2M}{r}\right)^{-\frac{1}{2}}, \tag{5.23}$$

and hence that[16]

$$g_\infty = \frac{M}{r^2}. \tag{5.24}$$

For us, it will take a little bit longer. Let us work through it slowly. First of all, since the observer is supposed to be hovering, we obtain the first equality in equation (5.23). The second equality follows from conservation of energy. The energies E_r and E_∞ must be related by a redshift factor, since otherwise we could turn E_∞ into radiation, send it to r and back, and hence increase our energy. Because [34, p. 33]

$$\lambda_\infty = \lambda\left(1 - \frac{2M}{r}\right)^{-\frac{1}{2}},$$

[16]Here, we are working in units where $G = 1$, so that $r_s = 2M$ instead of the usual $2GM$.

we get (5.23). Equation (5.24) is then simply obtained by solving for g_∞ and substituting (5.22). We conclude that the force needed to keep the observer hovering at $r = 2M$

$$\kappa = \frac{M}{4M^2} = \frac{1}{4M}, \tag{5.25}$$

or $c^4/(4GM)$ in SI-units. This force is interpreted as the surface gravity of the black hole. Also note that (5.24) reproduces the Newtonian surface gravity.

This method to compute the surface gravity is nice, but it doesn't generalize nicely to rotating black holes. There is a more sophisticated way to define κ in that case, which involves a Killing vector field K^μ. In one sentence, this is a vector field such that the metric is constant along its integral curves. Thus, Killing vector fields generate isometries. For the static spacetimes we are considering in this work, the vector field $K = \frac{\partial}{\partial t}$ works, since none of the metric components depend on this coordinate.

Given a Killing vector field such that $K_\mu K^\mu$ is normalized to -1 at infinity, we define the surface gravity implicitly via the equation

$$K^\mu \nabla_\mu K^\nu = \kappa K^\nu. \tag{5.26}$$

This defines a differential equation, which can be solved [35, p. 109] to give $\kappa = 1/4M$ - the same result as using the other method.

Thus, the Hawking temperature of a Schwarzschild black hole is

$$T = \frac{1}{8\pi M} \tag{5.27}$$

after setting $\hbar$ and Boltzmann's constant to unity. This has a number of consequences. First of all, because of the Stefan-Boltzmann law, we have [33, p. 17]

$$\frac{\mathrm{d}M}{\mathrm{d}t} = -\epsilon\sigma T^4 A_{\mathrm{hor}}.$$

The black hole will lose mass slowly. Since $T \propto 1/M$ and $A_{\mathrm{hor}} \propto M^2$, the rate of change of M scales like M^3, which means the black hole will evaporate in finite time. For sufficiently massive black holes, the lifetime is estimated as

$$\tau = 2.16 \times 10^{66} \left(\frac{M}{M_{\mathrm{sun}}}\right)^3 \text{ years},$$

which is a very long time. For some context, the current age of the universe is about 1.3×10^{10} years. It suffices to say that black holes will outlive us all.

We can also compute the specific heat of a Schwarzschild black hole:

$$C = T\frac{\partial S}{\partial T} = -\frac{1}{8\pi T^2} \tag{5.28}$$

As the black hole evaporates, temperature will increase, giving rise to a negative specific heat. This is quite uncommon in thermodynamics, but it makes sense for

gravitational systems. For the details, see [33, Section 6.2].

Finally, let us look at what happens for (extremally) charged black holes. Using either one of the methods I described earlier, one can derive that

$$T = \frac{1}{2\pi} \frac{\sqrt{M^2 - Q^2}}{(M + \sqrt{M^2 - Q^2})^2}. \tag{5.29}$$

Indeed, for $Q = 0$ this reduced to the $1/8\pi M$ which we had for Schwarzschild. For $Q = M$ (so the extremal case), something interesting happens: the Hawking temperature reduces to zero! As we have just learned, this implies that extremal black holes do not radiate. As a charged black hole loses mass, the above expression smoothly goes to zero, without encountering a singularity. This means that - assuming that the black hole somehow retains its constant charge - an extremal situation is the stable endpoint of Hawking radiation.

5.3 Causal structure and Penrose diagrams

5.3.1 Conformal transformations

Now, let us change point of view. Consider a metric $g_{\mu\nu}(x)$ which solves Einstein's equations (so we are not considering A_μ and ϕ for the moment). A basic question is how many degrees of freedom this metric describes. Since $g_{\mu\nu}$ is a symmetric 4×4 matrix, the naive answer is $\frac{1}{2} \times 5 \times 4 = 10$ degrees of freedom. However, there is a redundancy in our description. If two metric are related by an invertible change of coordinates, then they describe the same physical system. It turns out that this redundancy brings the number of degrees of freedom down from 10 to 6. [2, p. 146]

The corresponding symmetry is the invariance of Einstein's theory under diffeomorphisms. This symmetry implies [2, p. 148] the Bianchi identity

$$\nabla_\mu G^{\mu\nu} = 0,$$

where $G_{\mu\nu}$ is the Einstein tensor. The Einstein equations in free space ($G_{\mu\nu} = 0$) form a system of 10 coupled differential equations. Invariance under diffeomorphisms means that there is a dependence amongst these equations. Indeed, the Bianchi identity gives us four relations, which leaves us with six degrees of freedom.

Instead of diffeomorphisms, we can look at the slightly weaker notion of *conformal transformations*. If two metrics $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ are related by

$$\tilde{g}_{\mu\nu} = \Omega^2(x) g_{\mu\nu}, \tag{5.30}$$

with $\Omega(x)$ a smooth nowhere zero function of spacetime, then we say that the metrics are conformally equivalent. As we shall see, even though they might not be diffeomorphic, that they share many features. We say that they have the same *causal structure*. It is clear that a vector is lightlike w.r.t. g if and only if it is lightlike w.r.t. $\tilde{g}$. Hence, a conformal transformation maps null geodesics to null geodesics [2, p. 168].[17] The two metrics also agree on which vectors are spacelike and which are timelike, though the lengths may be changed.

[17] The mapping might screw up the affine parametrization, but the shape is the same.

5.3.2 Penrose diagrams (basics)

Now that we understand that conformal transformations do not change the causal structure of spacetime, we would like to have a method for *visualising* what this structure looks like. This is done by means of a Penrose diagram. The idea is to use a conformal transformation in order to map the spacetime to a compact set, on which we can draw out the causal structure.

As an example, consider the Minkowski spacetime in one spatial dimension. Recall that the metric is given by

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2\,. \tag{5.31}$$

We first introduce light cone coordinates $u = t - x$ and $v = t + x$. [2, p. 169] The metric in these coordinates is given by

$$\begin{aligned} \mathrm{d}t &= \frac{\mathrm{d}u + \mathrm{d}v}{2} \\ \mathrm{d}x &= -\frac{\mathrm{d}u - \mathrm{d}v}{2} \\ \implies \mathrm{d}s^2 &= -\frac{1}{4}(\mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}v + \mathrm{d}v^2) + \frac{1}{4}(\mathrm{d}u^2 - 2\,\mathrm{d}u\,\mathrm{d}v + \mathrm{d}v^2) \\ &= -\,\mathrm{d}u\,\mathrm{d}v \end{aligned}$$

where both u and v can be any real numbers. We can compactify these ranges by setting $u = \tan\tilde{u}$ and $v = \tan\tilde{v}$. Similarly to the above computation, the metric becomes

$$\mathrm{d}s^2 = -\frac{1}{\cos^2\tilde{u}\cos^2\tilde{v}}\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v}\,,$$

where now $\tilde{u}$ and $\tilde{v}$ take values in $(-\pi/2, \pi/2)$. So far, we have just done changes of coordinates, so this metric is fully equivalent to the original Minkowski metric. The function $\cos\tilde{u}\cos\tilde{v}$ does not vanish over the interval we are considering. Hence, the metric

$$\mathrm{d}\tilde{s}^2 = -\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v}$$

is conformally equivalent to the Minkowski metric. Since this metric is regular when $\tilde{u}, \tilde{v}$ approach the boundary of the domain, it is conventional to add the boundary points to the domain. This is called *conformal compactification*. It is also conventional to draw the $\tilde{u}$ and $\tilde{v}$ axes as diagonal lines. This way, we obtain the *Penrose diagram* of 1-dimensional Minkowski spacetime, which has the shape of a diamond.

Recall that light rays move according to $\mathrm{d}\tilde{s}^2 = 0$, so in this case $\mathrm{d}\tilde{u} = 0$ or $\mathrm{d}\tilde{v} = 0$. Hence, light rays *always* move at an angle 45 degree w.r.t. the horizontal. Since the causal structure is preserved under conformal transformations, massive particles must move up inside the future light cone as time increases. As we can see from Figure 9, all massive particles eventually end up at i^+ as $t \to \infty$. Light rays will eventually end up on $\mathcal{I}^+$, while all spacelike curves end at i^0. Finally, Penrose diagrams make it easy to visualise whether or not two events can influence each other.

We can of course play the same game for Minkowski spacetime with three spatial dimensions. The metric is written as

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}\mathbf{x}^2 = -\,\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\,\mathrm{d}{\Omega_2}^2\,,$$

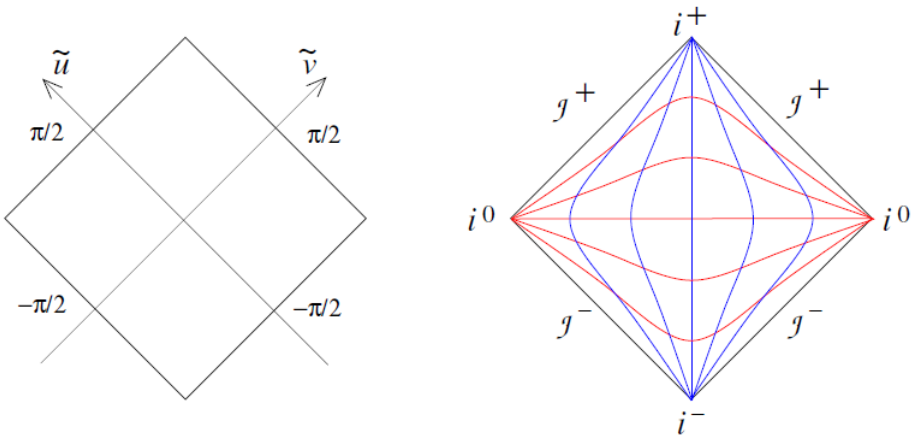


Figure 9: Taken from [2, Figure 33]. Left: Penrose diagram for (1+1)-dimensional Minkowski spacetime. Right: The same diagram with some matter trajectories.

where Ω_2 represents the metric on a 2-sphere:

$$\mathrm{d}{\Omega_2}^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\,.$$

Like before, we change coordinates to $u = t - r$ and $v = t + r$ and then compactify. The metric becomes

$$\mathrm{d}\tilde{s}^2 = -4\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v} + \sin^2(\tilde{u} - \tilde{v})\,\mathrm{d}{\Omega_2}^2\,.$$

There are two differences with the one-dimensional case. First of all, there is a term involving the metric of $\mathbb{S}^2$. We can imagine that there is a 2-sphere of radius $\sin(\tilde{u} - \tilde{v})$ at every point in the $(\tilde{u}, \tilde{v})$-plane. This sphere will be projected out to obtain a two-dimensional Penrose diagram. Secondly, note that r represents a radial coordinate, so $r \geq 0$. In particular, we must have $v \geq u$. The compactified coordinates hence take values in $-\pi/2 \leq \tilde{u} \leq \tilde{v} \leq \pi/2$. The resulting Penrose diagram is depicted in Figure 10.

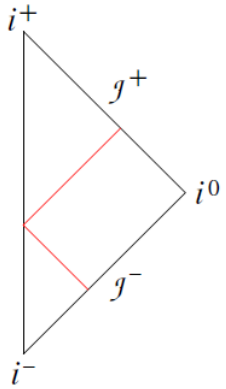


Figure 10: Taken from [2, Figure 35]. Penrose diagram for (3 + 1)-dimensional Minkowski spacetime. There is an implicit $\mathbb{S}^2$ at each point.

The vertical boundary on the left is not a boundary of spacetime. It represents the origin $r = 0$, and light rays bounce off it. The point where they "bounce" represents

when r changes from increasing to decreasing. The extra boundary is just an artefact of the coordinate singularity of spherical coordinates at $r = 0$. It is also possible to draw Penrose diagrams for (anti) de Sitter spacetime. These have fundamentally different behaviours at infinity. The curious reader is referred to [2, p. 174 - 180] for the details.

5.3.3 Penrose diagrams for black holes

We now consider a Schwarzschild black hole of mass M with $r_s = 2GM$. Recall that the metric is given by [10, p. 419]

$$\mathrm{d}s^2 = -\left(1 - \frac{r_s}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{r_s}{r}\right)^{-1} + r^2\,\mathrm{d}{\Omega_2}^2\,, \tag{5.32}$$

where $r_s = 2GM$. This metric has two singularities, namely at $r = r_s$ and at $r = 0$. By computing the scalar curvature, it can be shown that the singularity at $r = 0$ is a physical singularity, because the curvature blows up there. The singularity at $r = r_s$ is an artefact of "poorly" chosen coordinates.

At this point in a GR course, one would rewrite this metric in a bunch of different forms, each with its own advantages and disadvantages. The Wikipedia page for the Schwarzschild metric lists 7 different choices of coordinates. We will skip straight to the coordinates which are relevant for the Penrose diagram: Kruskal coordinates. We set $u = t - r_*$ and $v = t + r_*$, where [2, p. 239]

$$r_* = r + 2GM\log\left(\frac{r - 2GM}{2GM}\right).$$

The metric in these coordinates becomes

$$\mathrm{d}s^2 = -\left(1 - \frac{2GM}{r}\right)\mathrm{d}u\,\mathrm{d}v + r^2\,\mathrm{d}{\Omega_2}^2\,.$$

Unfortunately, the metric is still degenerate at $r = 2GM$. Hence, we apply another change of coordinates to $U = -e^{-\frac{u}{4GM}}$ and $V = e^{\frac{v}{4GM}}$. Notice that U is negative and V is positive outside of the event horizon. The metric becomes

$$\mathrm{d}s^2 = \frac{32(UV)^2}{r}e^{-\frac{r}{2GM}}\,\mathrm{d}U\,\mathrm{d}V + r^2\,\mathrm{d}{\Omega_2}^2\,.$$

Here, $r = r(U, V)$ is some function of U and V which can be determined from the definitions of the coordinates. The original coordinates require that $U < 0$ and $V > 0$, but there is nobody stopping you from extending the metric to the entire (U, V)-plane. The metric is only singular at the origin of this plane.

For a brief moment, let us admire the *Kruskal diagram* of the Schwarzschild black hole, which is shown in Figure 11. It can be shown that the horizon $r = 2GM$ is mapped to $U = 0$ or $V = 0$, while the singularity $r = 0$ goes to a hyperbola $UV = 1$. The exterior of the black hole corresponds to the right quadrant of this diagram. The top quadrant is the interior of the black hole. As time increases, particles must move up inside the light cone. They will eventually end up at the singularity $r = 0$. Finally, the lower quadrant is something called a white hole. For more details, see [2, 10].

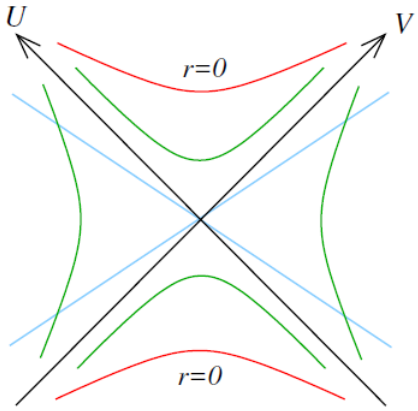


Figure 11: Taken from [2, Figure 48]. Kruskal diagram for a Schwarzschild black hole.

To obtain the Penrose diagram, we change coordinates one last time to $U = \tan \tilde{U}$ and $V = \tan \tilde{V}$. The resulting diagram is shown in Figure 12. The hyperbola $UV = 1$ now becomes a horizontal line, which is the zigzag at the top of the diagram. The causal structure is clearly different from Minkowski. A particle which somehow ends up in the top quadrant can never reach i^+ any more. It is forced towards the singularity.

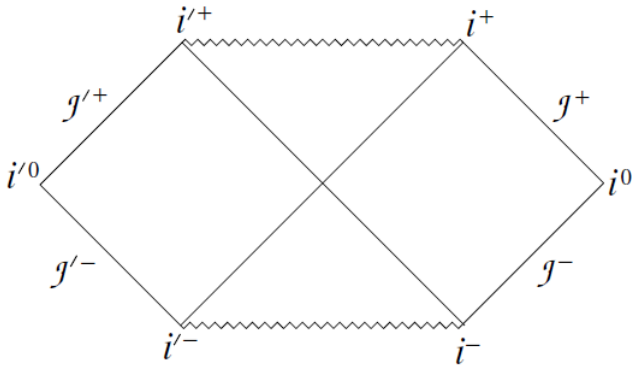


Figure 12: Taken from [2, Figure 51]. Penrose diagram for a Schwarzschild black hole. There is an implicit $\mathbb{S}^2$ at each point. (I will stop mentioning this from now on.)

This analysis can also be done for charged black holes. Consider the Reissner-Nordström solution with metric [2, p. 258]

$$\mathrm{d}s^2 = -f(r)^2\,\mathrm{d}t^2 + f(r)^{-2}\,\mathrm{d}r^2 + r^2\,\mathrm{d}{\Omega_2}^2 \tag{5.33}$$

and

$$f(r)^2 = 1 - \frac{2GM}{r} + \frac{Q^2}{r^2}.$$

We first investigate where the metric is singular. The function $f(r)^2$ can be factored as

$$f(r)^2 = \frac{1}{r^2}(r - r_+)(r - r_-), \quad r_\pm = GM \pm \sqrt{(GM)^2 - Q^2}.$$

We hence distinguish three cases, according to the amount of roots that $f(r)^2$ has. For $|Q| > GM$, the black hole is called *super-extremal*. In this case, the singularity at $r = 0$ is not covered by any event horizon. This contradicts the cosmic censorship conjecture, so these solutions are not physical.

If $|Q| < GM$, then the metric has three singularities, namely at $r = 0$ and at $r = r_\pm$. Since the curvature does not blow up at $r = r_\pm$, we again expect that these singularities are due to our choice of coordinates. They could again correspond to event horizons. Finally, if $|Q| = GM$, the black hole is called *extremal*. In this case, if any more charge is added to the black hole, a naked singularity will be revealed. The two horizons merge for an extremal hole.

Let us first analyse the causal structure in the sub-extremal case. After some fiddling with the coordinates, one can find that $r = r_+$ is again an event horizon. It is impossible to escape to infinity after reaching $r < r_+$. But that about the inner horizon at $r = r_-$? Like in the Schwarzschild case, we need to change to Kruskal-like coordinates. However, in the case of a charged black hole, these coordinates do not extend to the whole (U, V)-plane. Instead, they only extend to the horizon at $r = r_-$.

What does that mean physically? Once a particle reaches the region $r_- < r < r_+$ it is not forced to go towards the singularity at $r = 0$. The coordinates can be extended to a new region of spacetime, which actually looks the same as the old spacetime. Consider the completed Penrose diagram in Figure 13. The right side of the "diamond" is our universe. A brave observer crosses the horizon $r = r_+$ to end up in the upper part of the diamond. They will no longer be able to communicate with the rest of the universe, like in the Schwarzschild case.

What changes is that the observer can enter the lower region of another diamond. For Schwarzschild, this corresponded to a "white hole". The observer can exit the white hole and enter into a new universe. It has similar properties to their home universe, but the two are not in causal contact. This procedure can repeat ad infinitum. There is a whole tower of universes, both to the past and the future. Of course, this is all just speculation based on extending coordinates. It is not possible to confirm the existence of these additional universes in an experiment.

For an extremal Reissner-Nordström black hole, the inner and outer horizon coincide, and we get the Penrose diagram in Figure 14. Like in the sub-extremal case, there is an infinite tower of universes. However, since the inner and outer horizons are on top of each other, there is no white hole this time. The observer goes from the black hole exterior directly into the new universe.

5.4 The role of extremality

We have met the notion of an extremal black hole a few times by now. A charged black hole is extremal if it has the maximal amount of charge for the given mass. In some system of units, $|Q| = M$. As we have seen, adding even more charge would reveal a naked singularity, which is not allowed in physics. When we added in the

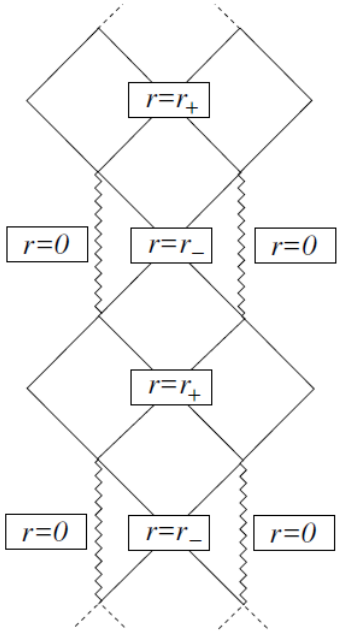


Figure 13: Taken from [2, Figure 56]. Penrose diagram for a non-extremal Reissner-Nordström black hole.

dilaton, the extremality equation became $M^2 = Q^2+\Sigma^2$, where Σ is the dilaton charge.

Extremality can be defined more generally for supersymmetric theories. I will not go into this at all, but the main highlight is that extremal states are states with mass equal to charge [36]. In this context, one often also uses the word *BPS-state*. As a general principle in physics, something with more symmetries is easier to handle. We have seen something similar with integrability in the context of classical mechanics. There, having enough conserved quantities made us able so integrate Hamilton's equations. In this subsection, I will tell two stories that hopefully illustrate the role of extremality.

The first story brings us back to black hole thermodynamics. I mentioned that it would be very nice to have an interpretation for the entropy of a black hole in terms of microstates. The first such calculation came from Strominger and Vafa in 1996 [37]. They derived the entropy formula for a very specific class of five-dimensional extremal black holes in string theory. To this end, they counted the number of BPS states leading to the same black hole from string theory. Their result has spawned similar calculations for other types of (near-)extremal black holes.

The second story is a bit more relevant for this work. We have considered the motion of light rays in the background of two (or more) extremal Reissner-Nordström black holes. An application of this is the imaging of black hole binary mergers, which is relevant nowadays because these mergers are an important source of gravitational waves. Studying the motion of light in the background can - for example - give us information on what the shadow of a binary system looks like.

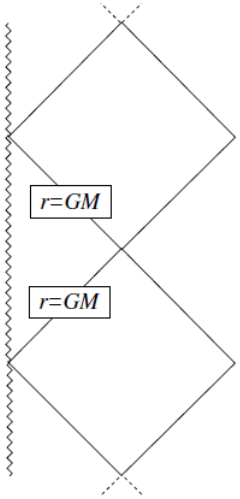


Figure 14: Taken from [2, Figure 58]. Penrose diagram for an extremal Reissner-Nordström black hole.

Studying the full system, where the two black holes orbit each other and eventually merge, is very complicated. One would need to write down a full time-dependent metric describing the inspiral. This includes the initial phase, where the black holes are orbiting each other, the inspiral phase with the emitted gravitational waves and the final state with just a single black hole. All we can really do is approximate the metric tensor in the regime before the merger.

Here, extremality comes to the rescue. Because the black hole has $Q = M$, the attraction due to gravity and the repulsion due to electromagnetism exactly cancel, and we get a *static* solution, for which we can explicitly write down the metric. Given this, it's not very hard to solve the corresponding geodesic equations on a computer. However, there is always a drawback. From a physical perspective, adding a bunch of charge to the black hole is very unrealistic. Indeed, if a black hole were to be charged, it would attract charge of the opposite type and become neutral. For this reason, all black holes we know of in nature are essentially neutral.

For this reason, I want to look at some recent simulations that have been done on more realistic black hole binaries, in order to compare the result to our "unrealistic" extremal case. For a review about numerically solving Einstein's equations, see for example [35]. In 2015, Bohn et al. [38] used a ray-tracing algorithm to generate images of the binary's shadow.

Of course, it's not actually possible to "see" a black hole. Therefore, the authors use an artificial background sphere located infinitely far from the black hole. This background can be seen in Figure 15. There is also a bright spot which the camera is pointed at. Images such as Figure 16 are obtained by evolving geodesics backwards in time until they either hit the black hole or until they get far enough away. Colours are assigned based on where on the background sphere the light rays end up.

With numerical solutions to Einstein's equations at hand, we can now look at images of a realistic black hole binary system. These images are shown in Figure 17. An interesting feature of the solution is the "eyebrow structure" that the black holes have. These eyebrows are due to geodesics that start close to one black hole but end up going around and falling into the other black hole. This can be compared to the bouncing between fundamental orbits that we encountered at the end of Section 2.

The authors [38] also plotted some close-starting trajectories and looked at how they evolved in time. The results can be seen in Figure 18. When zooming in on smaller and smaller regions of initial conditions, we see that there is some kind of self-similarity. Such a structure is very common for chaotic systems. This gives a good indication that chaos is indeed a feature of binary black hole systems.

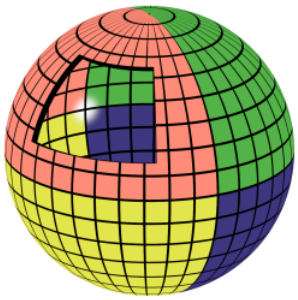


Figure 15: Taken from [38, Figure 3]. The background sphere.

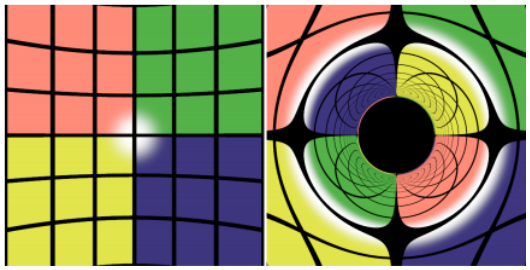
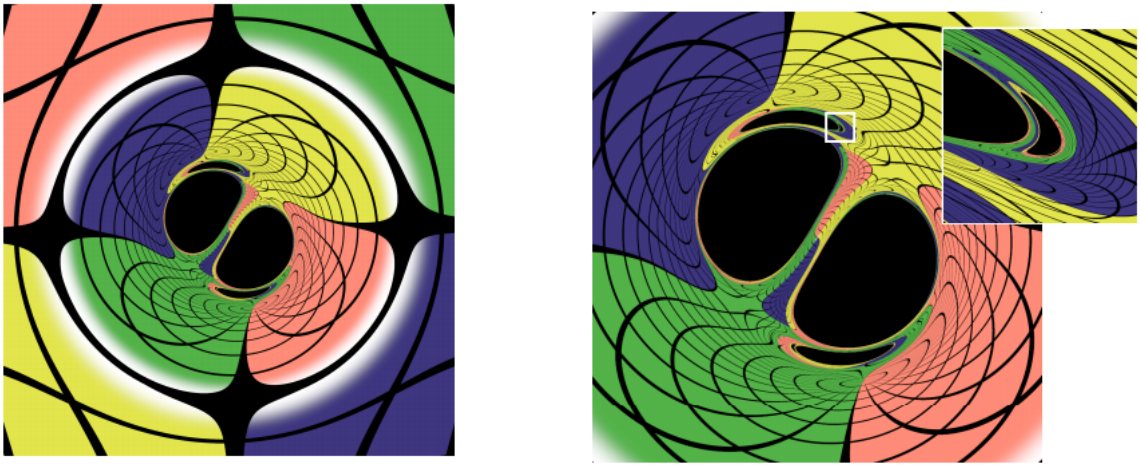


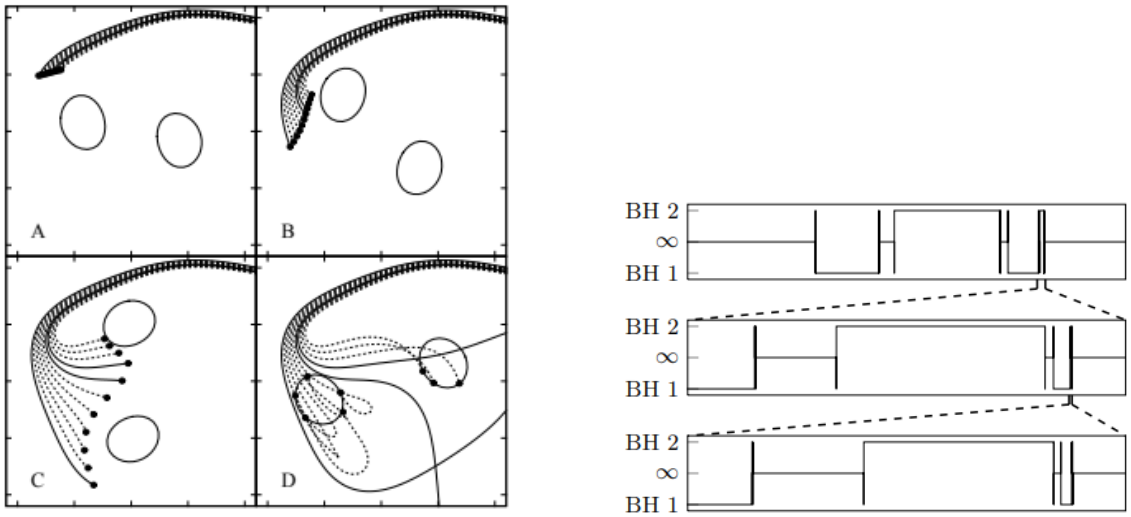
Figure 16: Taken from [38, Figure 4]. Image of a Schwarzschild black hole.

To close off this section, let me return to the comment that extremally charged black holes are very unrealistic. Shipley and Dolan [5] made images similar to Bohn et al. for the extremal case ($a = 0$). The shadow can be viewed in Figure 19. We again see the eyebrow structure that I commented on earlier. In fact, it looks like there is a fractal pattern of eyebrows that are getting smaller and smaller. This can be seen in the zoomed in image. In Figure 20, we can see that also in the unrealistic case there is self-similarity when the initial condition is varied. This somehow validates using the extremal black hole for analysing the qualitative structure. Either way, things are *much* easier to calculate in this case.



(a) Taken from [38, Figure 5]. (b) Taken from [38, Figure 6]

Figure 17: Left: Image of a realistic black hole binary. Right: Zoomed in version. Eyebrow structure is clearly visible.



(a) Taken from [38, Figure 8] (b) Taken from [38, Figure 9]

Figure 18: Left: Some trajectories in a realistic black hole binary. Right: Fate of the trajectory versus initial condition. Upon zooming in, self-similarity is revealed, giving an indication for chaotic motion.

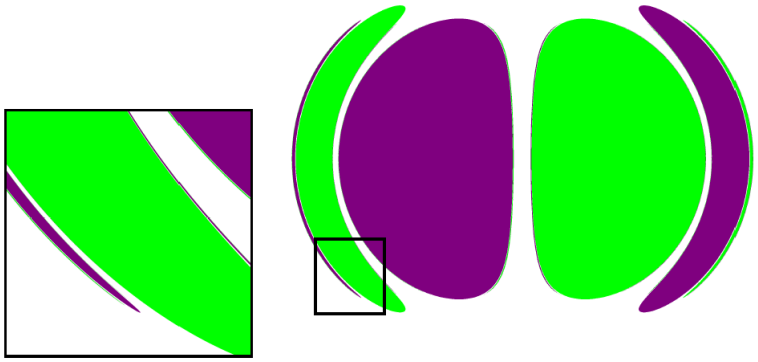


Figure 19: Taken from [5, Figure 1]. Shadow of an pair of extremal Reissner-Nordström black holes.

6 Applications to relativistic N-center problems

In this final section, I will attempt to bring together everything that we have learned so far. Recall that we looked at a class of static solutions to the Einstein field equations coupled to electromagnetism (spin 1) and a dilaton field (spin 0) with action

$$S = \int \mathrm{d}^4x \, \sqrt{-g} \left(R - \partial_\mu \phi \partial^\mu \phi - e^{-2a\phi} F_{\mu\nu} F^{\mu\nu} \right). \tag{6.1}$$

We will first look at the physical properties of the solution

$$\begin{cases} g & = -U^{-\frac{2}{1+a^2}} \, \mathrm{d}t^2 + U^{\frac{2}{1+a^2}} \, \mathrm{d}\mathbf{x} \cdot \mathrm{d}\mathbf{x} \\ A & = \frac{1}{\sqrt{1+a^2}U} \, \mathrm{d}t \\ e^{-\phi} & = U^{\frac{a}{1+a^2}} \end{cases} \tag{6.2}$$

in terms of Penrose diagrams and thermodynamics. After this, we will study the motion of light rays in the fixed geometry. Here, we will be able to apply methods from degree theory.

6.1 Physical properties of the solution

The above solution was derived independently by [16] and [17]. Actually, the authors of [16] derive for more general static solutions to the action

$$\begin{aligned} S = \int \mathrm{d}^d x \, \sqrt{-g} \Big[& \frac{1}{2k^2} R(g) - \frac{4}{(d-2)k^2} (\nabla \phi)^2 \\ & - \frac{1}{4} e^{-\frac{4}{d-2} a\phi} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2(d-2)!} e^{-\frac{4}{d-2} a'\phi} F^2_{(d-2)} - U(\phi) \Big], \end{aligned} \tag{6.3}$$

where $k^2 = 8\pi G$, d is the number of spacetime dimensions, and a, a' are coupling constants. We have previously understood the meaning of the first three terms in this action. Recall from our discussion on electromagnetism that $F_{\mu\nu}$ are the components of a 2-form. It leads to black hole solutions which are electrically charged.

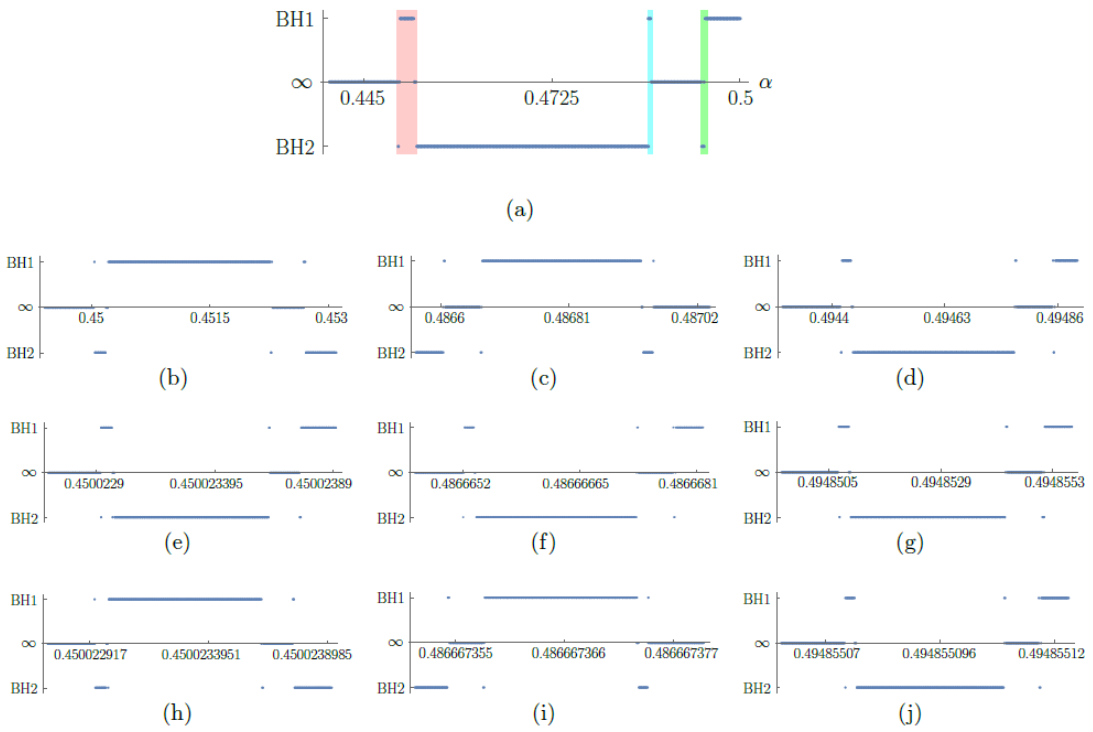


Figure 20: Taken from [5, Figure 7]. Fate of the trajectory plotted against the initial condition, which is labelled by an angle α. Again, we see self-similarity.

Of course, nobody is preventing us from adding magnetic charges too. This is done using a $(d-2)$-form $F_{(d-2)}$. It has its own coupling constant a'. We will not consider this term further. We will also always set the potential term $U(\phi)$ for the dilaton field to zero.

At this point, [16] points out a number of interesting special cases of the model. The case $a = 1$ is suggested by superstring theory - either the actual 10-dimensional theory, or some 4-dimensional effective variant. The case $a = \sqrt{\frac{d+n-2}{n}}$ corresponds to a Kaluza-Klein like theory with n extra compactified dimensions. In particular, for $d = 4$ and $n = 1$ we get $a = \sqrt{3}$. Finally, $a = 0$ is just the Einstein-Maxwell theory, and the dilaton drops out.

We will now turn to electrically charged but magnetically neutral black hole solutions to the theory. For simplicity, let us set $d = 4$. Reference [17] writes the metric for a single black hole in the following form:

$$\mathrm{d}s^2 = -\lambda^2\,\mathrm{d}t^2 + \lambda^{-2}\,\mathrm{d}r^2 + R^2\,\mathrm{d}{\Omega_2}^2 \tag{6.4}$$

with

$$\begin{cases} \lambda^2 &= \left(1 - \frac{r_+}{r}\right)\left(1 - \frac{r_-}{r}\right)^{\frac{1-a^2}{1+a^2}} \\ R &= r\left(1 - \frac{r_-}{r}\right)^{\frac{a^2}{1+a^2}}, \end{cases} \tag{6.5}$$

with $r_\pm$ being two free parameters. They are related to the mass and charge by

$$\begin{cases} M &= \frac{r_+}{2} + \left(\frac{1-a^2}{1+a^2}\right)\frac{r_-}{2} \\ Q &= \left(\frac{r_+ r_-}{1+a^2}\right)^{\frac{1}{2}}. \end{cases} \tag{6.6}$$

For an extremal solution, we set r_+ equal to r_-. Denoting this value by $r_\pm$, we obtain for the mass:

$$r_\pm = (1+a^2)M,$$

which is the "rescaled" mass we have seen before. We also recognise from the equation for λ^2:

$$\lambda^2 = \left(1 - \frac{r_\pm}{r}\right)^{1+\frac{1-a^2}{1+a^2}} = \left(1 - \frac{(1+a^2)M}{r}\right)^{\frac{2}{1+a^2}}.$$

With this observation, the metric looks very similar to the one from [3] which we have been considering. The only difference is a change to "isotropic coordinates", which only cover the region outside of the event horizons.

Finally, we look at the thermal and causal properties of the solution. All of these are given (both for $d = 4$ and $d \neq 4$) in [16]. The isotherms are plotted in Figure 21. Note that a clear change occurs at the critical value $a = 1$. We are looking at the extremal case, where the charge Q is as large as it can be for the given mass.[18] Here, the Hawking temperature is zero for $a < 1$ and infinite for $a > 1$. It is remarked by [16, p. 15] that the case $a < 1$ is similar to Reissner-Nordström, while $a > 1$ is similar to Kaluza-Klein. In the special case $a = 1$, the Hawking temperature is finite and nonzero for an extremal hole.

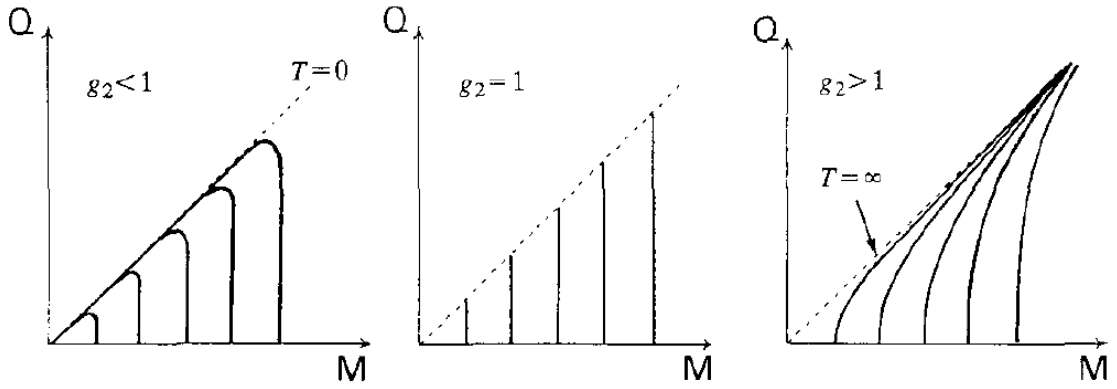


Figure 21: Taken from [16, Figure 1(a)]. Isotherms for Einstein-Maxwell-dilaton static black holes. The constant g_2 is denoted by a in this work.

Finally, let us discuss the Penrose diagrams, which are given in Figure 22. In (c) and (d), the dilaton coupling is set to zero, and we recover the Penrose diagrams for (extremal) Reissner-Nordström, which we have seen before. In (a), we see a nonextremal black hole with dilaton coupling. It has the same causal structure as

[18]The dilaton charge Σ of the black hole can be determined from Q and M by imposing a certain asymptotic value of the scalar field ϕ. This value ϕ_0 is set to zero in [17].

Schwarzschild, with a singularity at $r = 0$ covered by an event horizon.

The extremal case has a different causal structure. See Figure 22 (b). There is a naked singularity at $r = 0$ for all nonzero values of a. Hence, the term black hole might not be appropriate for this class of compact objects. Physicists do not like the idea of naked singularities in nature. Ideas like the *cosmic censorship conjecture* say that naked singularities should not exist, even though it is possible to construct solutions to the scalar-Einstein equations which have them [39]. Whether these solutions are physically realisable is a question for another time.

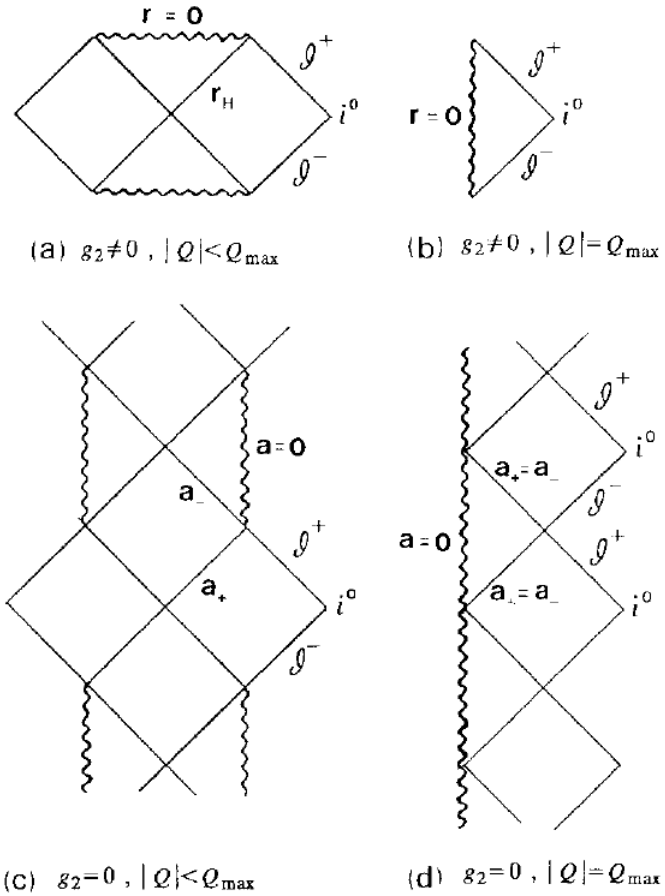


Figure 22: Taken from [16, Figure 3]. Penrose diagrams for Einstein-Maxwell-dilaton static black holes. The constant g_2 is denoted by a in this work. (a) Nonzero dilaton coupling, nonextremal. (b) Nonzero dilaton coupling, extremal, (c) No dilaton coupling, nonextremal, (d) No dilaton coupling, extremal.

6.2 Motion of light rays

To study the motion of light in this geometry, we consider the geodesic equation in its Hamiltonian form, with $H = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu$. We will assume throughout that the centers are located in the (x, z)-plane. In this case, the configuration space is two-dimensional, since we can eliminate the t and y-components.

Recall that the Hamiltonian is given by

$$H = \frac{1}{2}\left(-U^{\frac{2}{1+a^2}}p_t^2 + U^{-\frac{2}{1+a^2}}(p_x^2+p_y^2+p_z^2)\right). \tag{6.7}$$

We set p_t^2 equal to unity and p_y^2 to zero, in order to reduce the dimensionality of the system. Recall that the motion of light takes place on the $H=0$ level set. Hence, we can multiply through by the nonzero function $U^{\frac{2}{1+a^2}}$. This has the effect of doing a time rescaling. The actual trajectories in configuration space will be the same. We obtain the new Hamiltonian

$$H' = -\frac{1}{2}U^{\frac{4}{1+a^2}} + \frac{1}{2}(p_x^2+p_z^2) + \frac{1}{2}, \tag{6.8}$$

which is now in the form "kinetic + potential". The $\frac{1}{2}$ is added to make the potential go to zero at infinity.

From now on, we define $\alpha = \frac{4}{1+a^2}$ and $\vec{x} = (x,z)^T$. Similarly, $\vec{p} = (p_x, p_z)^T$. We will also absorb the factor of $1+a^2$ into the mass M_i. The term $\frac{1}{2}$ is added to make the potential go to zero at infinity. With these simplifications, the Hamiltonian becomes

$$H' = -\frac{1}{2}\left(1+\sum_{i=1}^{n}\frac{M_i}{\|\vec{x}-\vec{s_i}\|}\right)^{\alpha} + \frac{1}{2}\|\vec{p}\|^2 + \frac{1}{2}. \tag{6.9}$$

Observe that for light rays $\frac{1}{2}g^{\mu\nu}p_\mu p_\nu = 0$, so the Hamiltonian has constant value $\frac{1}{2}$ along trajectories.

Since $a \geq 0$, we are interested in $\alpha \leq 4$. We will see later on what the influence of the parameter α is. As we have already noted, $\alpha = 1$ or $a = \sqrt{3}$ corresponds directly to the classical 2-center problem, which is regular. Substituting $\alpha = 1$ in H' gives

$$H' = -\frac{1}{2}\left(1+\sum_{i=1}^{n}\frac{M_i}{\|\vec{x}-\vec{s_i}\|}\right) + \frac{1}{2}\|\vec{p}\|^2 + \frac{1}{2} = \sum_{i=1}^{n}\frac{M_i}{\|\vec{x}-\vec{s_i}\|} + \frac{1}{2}\|\vec{p}\|^2 = \frac{1}{2}.$$

Thus, we are effectively setting the Kepler Hamiltonian equal to $\frac{1}{2}$, so that we are considering scattering at a positive energy.

To summarise, we will study the planar motion of particles in a potential given by

$$V(x) = -\frac{1}{2}\left(1+\sum_{i=1}^{n}\frac{M_i}{\|\vec{x}-\vec{s_i}\|}\right)^{\alpha} + \frac{1}{2} \tag{6.10}$$

at $E = \frac{1}{2}$.

As a first step, let us study the form of $V(x)$ close to $x = s_i$.[19] We recall the standard Taylor expansion

$$(1+\epsilon)^\alpha = 1 + \alpha\epsilon + \frac{1}{2}\alpha(\alpha-1)\epsilon^2 + O(\epsilon^3).$$

[19]We drop the vector arrows from x and s_i when there is no possible confusion.

Suppose for the moment that there is only one center, located at $x = s$. Let's take the dominant term outside of the brackets:

$$\begin{aligned} V(x) &= \frac{1}{2} - \frac{1}{2}\frac{M^\alpha}{\|x-s\|^\alpha}\left(1 + \frac{\|x-s\|}{M}\right)^\alpha \\ &= \frac{1}{2} - \frac{1}{2}\frac{M^\alpha}{\|x-s\|^\alpha}\left(1 + \frac{\alpha\|x-s\|}{M} + \dots\right) \\ &= \frac{1}{2} - \frac{1}{2}\frac{M^\alpha}{\|x-s\|^\alpha} - \frac{1}{2}\frac{\alpha M^{\alpha-1}}{\|x-s\|^{\alpha-1}} + O\left(\frac{1}{\|x-s\|^{\alpha-2}}\right). \end{aligned}$$

Thus, the dominant term behaves like $V(r) \propto -\frac{1}{r^\alpha}$, which we studied in Section 4.

6.3 A crude cutoff

As a first step, we will investigate if the chaos we observe in the relativistic N-center problem is due to the behaviour close to the centers. To this end, we will introduce smooth bump functions $\psi_i : \mathbb{R}^2 \to \mathbb{R}$ supported in $B_i = \{x \in \mathbb{R}^2 : \|x - s_i\| < R_i\}$.

We replace the original potential by

$$V(x) = -\sum_{i=1}^{N} \frac{M^\alpha}{2\|x-s_i\|^\alpha}\psi_i(x) = \frac{1}{2}\sum_{i=1}^{N} V_i(x). \tag{6.11}$$

From Theorem 4.11, we know that for $\alpha = \frac{2n}{n+1}$ ($n \in \mathbb{N}$) the individual potentials without ψ_i define a regularisable flow. Moreover, each positive energy is nontrapping and the degree equals $-n$. These properties persist upon multiplying by the bump functions, as is shown in the following Lemma:

Lemma 6.1. *For potentials of the form*

$$V(x) = -\frac{M^\alpha}{2\|x-s_i\|^\alpha}\psi_i(x) \tag{6.12}$$

with $\alpha = \frac{2n}{n+1}$ each positive energy is nontrapping and $\deg(E) = -n$.

Proof. We first prove the result on the degree. Write

$$\begin{aligned} V(x) &= -\frac{M^\alpha}{2\|x-s_i\|^\alpha} + \frac{M^\alpha}{2\|x-s_i\|^\alpha} - \frac{M^\alpha}{2\|x-s_i\|^\alpha}\psi_i(x) \\ &= -\frac{M^\alpha}{2\|x-s_i\|^\alpha} + \frac{M^\alpha}{2\|x-s_i\|^\alpha}(1-\psi_i(x)) \\ &= -\frac{M^\alpha}{2\|x-s_i\|^\alpha} + W(x). \end{aligned}$$

Note that the function W is smooth at s_i, since $1-\psi_i$ is zero in some neighbourhood of s_i. Hence, it follows from [9, Proposition 4.1] that the degree is still $-n$. To prove that E is nontrapping, we use [9, Remark 3.3]. If the estimate $\langle x, \nabla V(x)\rangle \leq 0$ holds, then all positive E which are regular values of V are nontrapping. It's a straightforward computation to show that the estimate holds for $-\frac{1}{r^\alpha}$ potentials. By choosing a bump function whose radial derivative is nonpositive, we preserve the previous inequalities. Hence, all positive energies are nontrapping and the Lemma is proven. □

The above is quite flexible, as we still have complete control of the size of the balls on which the cutoff functions are supported. We can make the balls small enough so that they don't overlap or shadow each other. This is possible as long as the centers are in general position, i.e. no three centers lie on a single straight line. In this way, we can satisfy all the assumptions of Theorem 4.12.

We can thus conclude the existence of trajectories which hit the supports in any specified order. This gives us a well-defined symbolic code for the problem. The existence of this symbolic code is independent of the size of the supports. They just have to be small enough. Hence, we can shrink the supports until the Taylor approximation is good enough. In this sense, the coding should also work for the original problem, and we can say that the chaos is caused by the behaviour of particles close to a center. I will make this more precise in the next subsection.

If $n \geq 3$, this symbolic code immediately proves that the system is chaotic for α of the form $\frac{2n}{n+1}$, as the symbolic code has positive entropy. Indeed, the code is simply given by bi-infinite sequences of the symbols $1, 2, \ldots, n$ with the rule that a symbol cannot follow itself. We will give a precise definition of entropy later on. This method does **not** prove anything for the two-center problem as in this case the code always looks locally like $\ldots 12121212 \ldots$. This code has entropy zero.

Intuitively, it makes sense that this symbolic coding does not prove chaos in the case of two centers, as it cannot tell the difference between a loop around both centers and a figure 8. According to [4], both of these trajectories should exists and both are unstable. We thus need a new idea to prove chaos for relativistic 2-center problems.

6.4 The higher order terms

Let us look a little bit closer at the Taylor expansion of the potential. We studied earlier what happens when we only keep the dominant term, which behaves like $-\frac{1}{r^\alpha}$. Now, I also want to keep the second order term around. Recall that the 1-center potential is given by

$$V(x) = \frac{1}{2} - \frac{1}{2}\frac{M^\alpha}{\|x-s\|^\alpha} - \frac{1}{2}\frac{\alpha M^{\alpha-1}}{\|x-s\|^{\alpha-1}} + O\left(\frac{1}{\|x-s\|^{\alpha-2}}\right)$$

for x close to s. For $\alpha \in (1,2)$, we have that $\alpha - 1 \in (0,1)$ and $\alpha - 2 \in (-1,0)$. Hence, only the first two terms in the Taylor expansion are singular, while all higher order terms are regular. By using Taylor's theorem with remainder, we can write

$$V(x) = \frac{1}{2} - \frac{C}{\|x-s\|^\alpha} - \frac{D}{\|x-s\|^{\alpha-1}} + W(x)$$

with C, D positive constants and $W(x)$ a smooth function. When there are N centers, we can do this expansion near each center. This gives

$$\begin{aligned} V(x) &= -\frac{1}{2}\left(1 + \sum_{i=1}^N \frac{M_i(1+a^2)}{\|x-s_i\|}\right)^\alpha + \frac{1}{2} \\ &= \sum_{i=1}^N \left(-\frac{C_i}{\|x-s_i\|^\alpha} - \frac{D_i}{\|x-s_i\|^\alpha} + W_i(x)\right) + \text{Cst.} \end{aligned}$$

Like before, we can use a smooth bump function ψ to bring us to the compact support case. Indeed,

$$\begin{aligned} V(x) &= -\sum_{i=1}^{N} \psi(\|x-s_i\|)\left(\frac{C}{\|x-s_i\|^{\alpha}} + \frac{D}{\|x-s_i\|^{\alpha-1}} + W_i(x)\right) \\ &\quad -\sum_{i=1}^{N}(1-\psi(\|x-s_i\|))\left(\frac{C}{\|x-s_i\|^{\alpha}} + \frac{D}{\|x-s_i\|^{\alpha-1}} + W_i(x)\right) + \text{Cst.} \\ &= -\sum_{i=1}^{N} \psi(\|x-s_i\|)\left(\frac{C}{\|x-s_i\|^{\alpha}} + \frac{D}{\|x-s_i\|^{\alpha-1}}\right) + W(x) + \text{Cst.} \end{aligned}$$

The term involving $(1-\psi(\|x-s_i\|))$ is smooth, because the part in brackets is smooth except at $x = s_i$, where the bump function is zero. Hence, this part can be absorbed into a single smooth function $W(x)$.

We now consider a centrally symmetric potential of the form

$$V(r) = -\frac{C}{r^{\alpha}} - \frac{D}{r^{\alpha-1}} \tag{6.13}$$

with C, D positive constants and $\alpha \in (1,2)$. We have shown earlier that - when $D = 0$ and $\alpha = \frac{2n}{n+1}$ - the motion can be regularized and the scattering degree is given by $\deg(E) = -n$. Since the term involving $-\frac{1}{r^{\alpha-1}}$ is of lower order, one could expect that it doesn't affect the scattering degree. Let us try to understand why this is indeed the case.

Assume for a moment that the potential can be regularized. We will (unfortunately) consider again the integrals that define the scattering angle in case of a centrally symmetric potential. If we add the extra term involving D, we will obtain

$$\Delta\phi(E,\ell) = 2\int_0^{v_{\max}} \frac{\mathrm{d}v}{\sqrt{E\ell^{\frac{2\alpha}{2-\alpha}}2^{\frac{-\alpha}{2-\alpha}} - v^2 + Cv^{\alpha} + Dv^{\alpha-1}\left(\frac{\ell}{\sqrt{2}}\right)^{\frac{2}{2-\alpha}}}} - \pi,$$

where now $v_{\max}$ is the smallest positive solution to the equation

$$E\ell^{\frac{2\alpha}{2-\alpha}}2^{\frac{-\alpha}{2-\alpha}} - v^2 + Cv^{\alpha} + Dv^{\alpha-1}\left(\frac{\ell}{\sqrt{2}}\right)^{\frac{2}{2-\alpha}} = 0.$$

When $\ell \to 0$, notice that $v_{\max}$ converges to the solution of the simpler equation

$$Cv^{\alpha} = v^2 \implies v_{\max} = C^{\frac{1}{2-\alpha}}.$$

Hence, upon interchanging the limit with the integral, we obtain

$$\Delta\phi = 2\int_0^{C^{\frac{1}{2-\alpha}}} \frac{\mathrm{d}v}{\sqrt{Cv^{\alpha} - v^2}} = \frac{2\pi}{2-\alpha} - \pi,$$

precisely like before. Moreover, the estimate

$$\left\langle x, \nabla\left(-\frac{C}{\|x\|^{\alpha}} - \frac{D}{\|x\|^{\alpha-1}}\right)\right\rangle = -C\left\langle x, \nabla\left(\frac{1}{\|x\|^{\alpha}}\right)\right\rangle - D\left\langle x, \nabla\left(\frac{1}{\|x\|^{\alpha-1}}\right)\right\rangle \leq 0$$

shows that the potential is still nontrapping.

Thus, assuming we can regularize the collisions in this potential, the same arguments as before allow us to conclude the existence of chaotic scattering. Since the potential

$$V(x) = -\sum_{i=1}^{N} \psi(\|x - s_i\|) \left(\frac{C}{\|x - s_i\|^{\alpha}} + \frac{D}{\|x - s_i\|^{\alpha-1}} \right) + W(x) + \text{Cst}$$

is *identical* to the original potential - modulo smooth perturbations - this argument is exact.

Finally, let me comment on the regularization. In [7, Section 2], the Kepler problem ($\alpha = 1$) is regularized using phase space extension. Intuitively, the only natural way to regularize the problem is by making trajectories that are about to hit a center "reflect" back. In case of the 1-center problem with

$$V(r) = -\frac{1}{r^{\alpha}},$$

the collision orbits are precisely those with zero angular momentum. To this end, it is important that the limit $\lim_{\ell \to 0} \Delta\phi$ is a multiple of π. As we have seen, this happens precisely for the special values $\alpha = \frac{2n}{n+1}$. I have also argued above that this limit is unchanged by adding the second term in the Taylor series. Hence, I expect that the motion can be regularized without much more work. This is also backed up by some computations done by my supervisor Marcello Seri. Unfortunately, due to time constraints, I did not have time to go into this further.

6.5 What about the 2-center problem?

So far, we have been unable to treat the 2-center problem rigorously, because our symbolic code cannot distinguish between figure-8 orbits and those that just go around the two centers. In this subsection, I will discuss the two symbolic codes that have been proposed by [3] (from now on referred to as Cornish-Gibbons or CG) and [5] (Shipley-Dolan or SD). We will study the symbolic code proposed by SD in detail, as well as some other aspects of relativistic 2-center problems.

6.5.1 Fractal methods

In a chaotic system, one of the defining properties is the sensitive dependence upon initial conditions. Hence, a good way to start understanding whether or not your system is chaotic is to look at the long term behaviour of the solution as a function of initial condition. To this end, we put a photon between the two centers at $(x_0, z_0) = (-10, z_0)$ with $z_0 \in (-1, 1)$ and shoot it off at some angle $\theta \in (-\pi, \pi)$. Since photons must move at the speed of light, this completely determines the initial conditions $(x_0, z_0, p_{x0}, p_{z0})$. By convention, the centers are located on the z-axis with $z = \pm 1$ and $M = 1$ for both black holes.

By solving the equations of motion (2.28) (with the dilaton coupling constant a inserted in the proper places) numerically, we can investigate the chaotic behaviour as a function of a. Roughly speaking, there are three things that can happen to our photon:

- It falls into the upper black hole ($z = +1$);
- It falls into the lower black hole ($z = -1$);
- It escapes to infinity.

CG make another distinction in the third case. Namely, the photon can escape either to $z = +\infty$ or $z = -\infty$. I give each of these four outcomes a different colour and plot the outcome on the (θ_0, z_0)-plane for several values of a. Below, I have plotted the cases $a = 0, 1$ and $\sqrt{3}$ in Figures 23, 24 and 25 respectively. A point is given a red (resp. blue) asterisk if it is absorbed by the upper (resp. lower) black hole. It is given a cyan (resp. magenta) circle if the particle escapes to $z = +\infty$ (resp. $z = -\infty$).

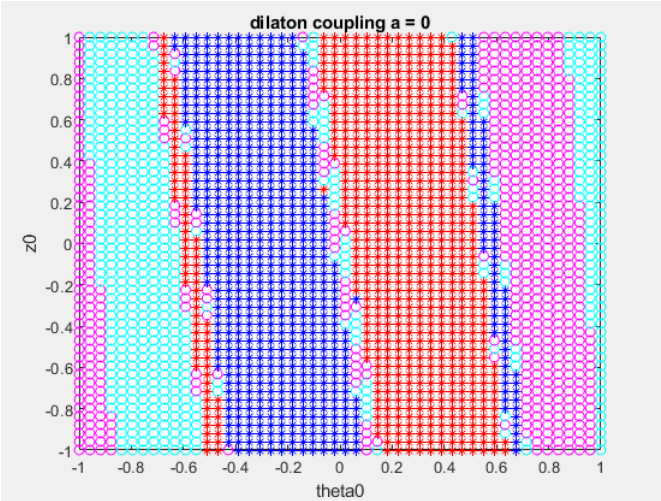


Figure 23: Outcome for the given initial condition in case $a = 0$.

If one uses more seed points than I did, the boundaries of the different regions look fractional for $0 \leq a < \sqrt{3}$ and regular for $a \geq \sqrt{3}$. Cornish an Gibbons apply box counting [40] to estimate the fractional dimension. Their result is as follows:

$$D(a = 0) = 1.36 \pm 0.02$$
$$D(a = 1) = 1.50 \pm 0.02$$
$$D(a = \sqrt{3}) = 1.01 \pm 0.02.$$

They claim that the dimension D peaks at $a = 1$, so that the chaos is most pronounced at this value. For my symbolic code, the entropy is independent of a, so I cannot back up this claim.

It is worth noting again the phase transition at $a = 1$. It can be seen in Figure 25 that all trajectories (except those aimed directly at one of the centers) escape to infinity. In fact, this is a feature of potentials of the form

$$V(r) \propto -\frac{1}{r^\alpha}. \tag{6.14}$$

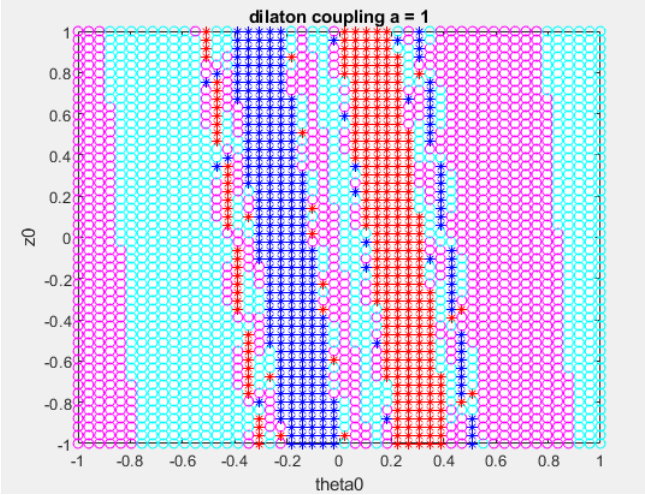


Figure 24: Outcome for the given initial condition in case $a = 0$.

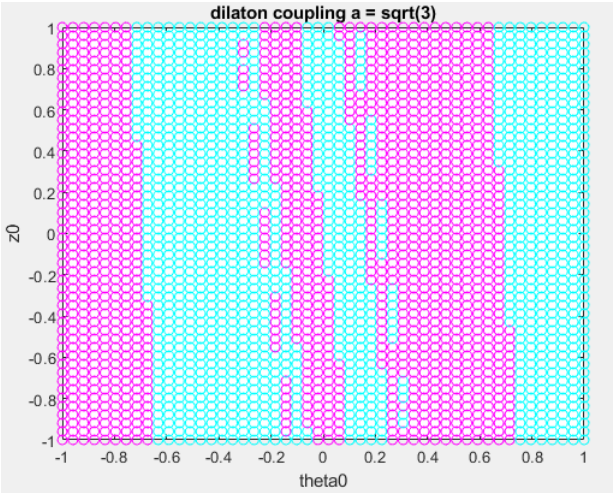


Figure 25: Outcome for the given initial condition in case $a = 0$.

For $\alpha \in (1, 2)$, the only collision trajectories are those with zero angular momentum. On the other hand, for $\alpha \geq 2$ there are trajectories with $\ell \neq 0$ that reach the center in finite time. This is sometimes called the "strong force regime". Note that $\alpha < 1$ corresponds to negative values of the dilaton coupling, so these values are not considered.

In using the degree to construct symbolic dynamics, continuity is an essential ingredient. This continuity is lost either when the collisions cannot be regularized, or when the nontrapping assumption is violated. Since for $a \leq 1$ we start losing a lot of trajectories which are captured by the black hole, I do not expect that the degree method is feasible in this case.

6.5.2 CG Symbolic dynamics

Now, let us consider the symbolic code introduced by CG in 1997. If the two centers are located at $s_{\pm} = (0, \pm 1)$ in the (x, z)-plane, we take the line segments defined by $z > 1$, $-1 < z < 1$ and $z < -1$ with $x = 0$. The Poincaré section consists of these

three parts crossed with $\mathbb{R}^2_p$. We record the symbol "+1" if the trajectory crosses the upper part of the Poincaré section, "0" for the middle part and "−1" for the lower part. Note that it is impossible to record the same symbol twice in a row.

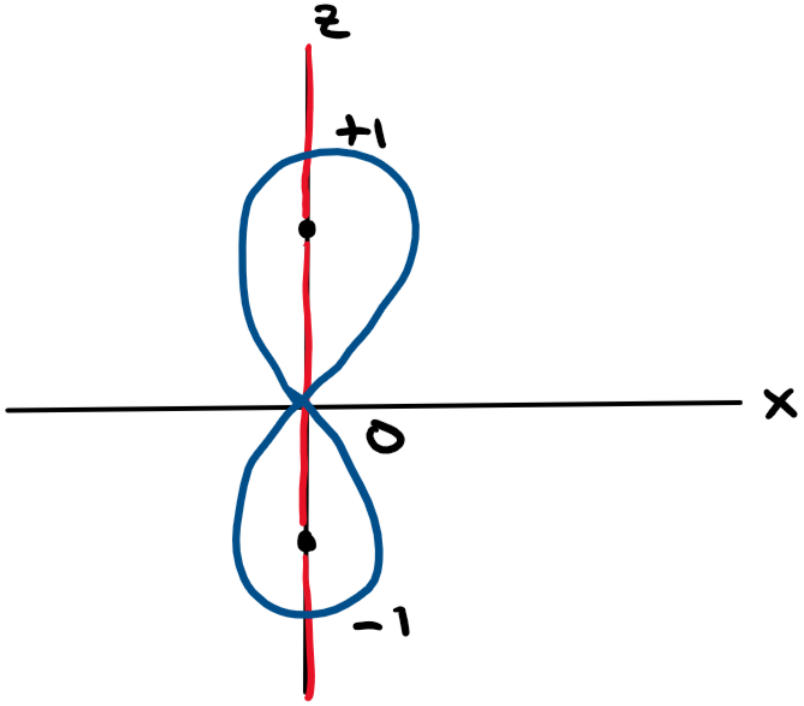


Figure 26: Setup for the symbolic code used by Cornish and Gibbons

The setup is depicted in Figure 26. As an example, the (periodic) solution drawn in blue is associated to the symbolic sequence $\overline{0,+1,0,-1}$, where the bar means that the block is repeated. For convenience, it is possible to change symbols to $A = \{(+1,-1),(-1,1)\}$, $B = \{(0,-1),(-1,0)\}$ and $C = \{(0,+1),(+1,0)\}$. The blue orbit then becomes $\overline{BC}$.

The length of a symbolic sequence is defined as the sum of the absolute values of the numbers that appear it. For example, one period of the blue orbit $0,+1,0,-1$ has length 2. Moreover, we define the *entropy* of the symbolic code as

$$S = \lim_{k\to\infty} \frac{\log(N(k))}{k}, \tag{6.15}$$

where $N(k)$ is the number of symbolic sequences of length k. The entropy is a measure of complexity, and hence can be used to quantify the amount of chaos in the system. A natural question to ask is thus: "What is the entropy of the symbolic code as a function of the coupling constant a?"

Let us work through an example of a code with just two symbols 1 and 2, which are both of length 1. Then, the number of sequences of length k is obviously given by

$$N(k) = 2^k.$$

Thus, the entropy of this simple symbolic code is

$$S = \lim_{k\to\infty} \frac{\log\left(2^k\right)}{k} = \lim_{k\to\infty} \frac{k\log(2)}{k} = \log(2).$$

In a similar code with N symbols $1, 2, \ldots, N$, one obviously finds $S = \log(N)$.

We now turn back the CG code in terms of A, B and C. Assume for a moment that all possible combinations of these letters are allowed and actually correspond to physical orbits. In this case, one might expect that the entropy is $\log(3)$. However, notice that an A is actually twice as long as a B or a C. Hence, the entropy is actually somewhere in between $\log(2)$ and $\log(3)$. Cornish and Gibbons show using simple combinatorics that

$$S = \log\left(1 + \sqrt{2}\right).$$

In the case $0 \leq a \leq 1$, CG argue that all possible strings of A, B, C *do* actually exist, while for $a > 1$ the code gets "pruned", i.e. not all combinations are allowed any more. Their argument is similar in flavour to the our degree argument, but it does not seem to be fully rigorous.

The argument goes as follows: The authors write down the equation for $\Delta\phi$ for a single center using a *different* potential than the one we used, namely

$$V(r) = \left(1 + \frac{(1+a^2)M}{r}\right)^{-\frac{2}{1+a^2}} r^{-1}.$$

Then, they argue that the *maximum* scattering angle occurs when a trajectory approaches close to one of the centers, where the effect of the other one can be safely ignored. For $0 \leq a \leq 1$, they argue that the maximum scattering angle is infinite. Hence, we can string an infinite amount of B's and C's together.

For $a > 1$, the maximum scattering angle $\Delta\phi_{\max}$ is no longer infinite. The longest string of the form $BB\ldots B$ that can occur is given by

$$n = \lfloor \frac{\Delta\phi_{\max}}{\pi} \rfloor,$$

where we use the floor function. Once $\Delta\phi_{\max}$ drops below 2π, it is no longer possible to, for example, have a loop around one of the centers between two figure eights. At this point, the symbolic code effectively has two symbols, namely A and BC, and the entropy drops to $\log(2)$.

Finally, when $\Delta\phi_{\max}$ goes below π, the figure eight orbit is no longer possible, and the symbolic code collapses to a single letter. The entropy drops to zero, and the system is no longer chaotic. This corresponds to the limit of $a = \sqrt{3}$ where the system becomes integrable, hence regular.

I conjecture that the scattering degree corresponds to $\Delta\phi_{\max}$ by just multiplying by $-\pi$. These results are summarised in Table 1. The correspondence between the maximum scattering angle and degree suggests, for example, that for $10/6 \leq a^2 \leq 2$ the scattering degree is effectively -2. Thus, if we could somehow get rid of the need to regularize the motion, this table suggests what the results should be.

a^2	α	degree	$\Delta\phi_{\max}$
3	1	-1	π
2	4/3	-2	2π
10/6	6/4	-3	3π
...	...	...	...
1	2	$-\infty$	∞

Table 1: Relation between a^2, α and the scattering degree, together with a conjectured correspondence to the maximum scattering angle.

7 Conclusion and outlook

The goal of this thesis was to rigorously prove the existence of chaos in relativistic N-center problems. To this end, we considered a family of static solutions to the Einstein-Maxwell-dilaton equations, generated by the action

$$S = \int \mathrm{d}^4x \sqrt{-g} \left(R - \partial_\mu \phi \partial^\mu \phi - e^{-2a\phi} F_{\mu\nu} F^{\mu\nu} \right).$$

The family is parametrized by the dilaton coupling constant a, and solution contains N extremal black holes.

We studied the motion of light in the field of these N centers. Since light only couples to the metric, and not to electromagnetism and dilaton, we only had to solve the geodesic equation. In its Hamiltonian form, we rewrote it in the form "kinetic + potential" by doing a time rescaling. In this form, we could apply methods from potential scattering.

Close to a center, we showed that the dominant term in the potential is

$$V(r) \propto -\frac{1}{r^\alpha},$$

where

$$\alpha = \frac{4}{1+a^2}.$$

More precisely, we used Taylor's theorem to prove that for one center located at $x = s$

$$V(x) = -\frac{C}{\|x-s\|^\alpha} - \frac{D}{\|x-s\|^{\alpha-1}} + W(x),$$

where $W(x)$ is a smooth perturbation. If $\alpha = \frac{2n}{n+1}$, potentials of this form can be regularized and their scattering degree is given by $\deg(E) = -n$. For these values of α, we could hence rigorously prove the existence of chaos when $N \geq 3$ by using Theorem 4.12.

What happens for the other values of α? Though at the moment we cannot say anything rigorously, I have argued at the end of Section 6 that there should be chaos for all values of $\alpha \in (1, 2)$, i.e. $a \in (1, \sqrt{3})$. We have seen that a qualitative change happens at $a = 1$. When $0 \leq a \leq 1$, many trajectories will be absorbed by the black

holes, even though they have nonzero angular momentum. This makes it impossible to apply the continuity arguments on which our degree method relies. We see numerically that chaos does persist for these values of a, though we currently cannot prove it rigorously.

The constraint $\alpha \in (1, 2)$ is often added when studying Kepler-like problems. For example, the recent paper [41] proves the existence of trajectories with a prescribed scattering angle in the classical N-center problem with potentials of the form $-\frac{1}{r^\alpha}$. Interestingly, they do not need to regularize the motion, as their argument is based on variational calculus. They do however impose that $\alpha \in [1, 2)$. Looking more deeply into papers like this could be a good extension of the research.

For $a = \sqrt{3}$, we have seen that the system becomes integrable. Hence, there is no chaos in this case. In fact, the potential in this case becomes exactly the Kepler potential and all the relativistic corrections cancel out to all orders in perturbation theory. From a physics point of view, it might be interesting to look into whether this cancellation can also happen in other theories of gravity. For example, one can look at the paper [42]. The authors study the perihelion shift in $N = 8$ supergravity. If it vanishes, that could be a hint that there is an additional symmetry.

Finally, our method was unable to prove the existence of chaos in the 2-center problem. This is because the symbolic code could not tell the difference between a figure-8 orbit and a loop around both centers. We have thought about how to remedy this problem. For example, one could add the line connecting the two centers to the Poincaré section. Unfortunately, we can no longer apply Theorem 4.12 in this case.

However, recall that the degree $\deg(E)$ tracks how often a trajectory can rotate about the center before scattering away. Perhaps, if we make this connection more precise, we can prove the existence of trajectories which get close to one center, rotate around a specified amount of times, and then either go to the other center or scatter away. At the moment, it is unclear to me whether or not this would actually work, but it is definitely worth looking into.

A Quick notes on homology and degree

The notion of degree for a continuous map $f : \mathbb{S}^d \to \mathbb{S}^d$ from a sphere to itself has been quite central in this work. In this appendix, I will introduce the bare minimum of algebraic topology required to understand this definition. Briefly put, we will see the definition of singular homology, some of its properties and the application to self-maps of $\mathbb{S}^d$.

Algebraic topology turns problems of algebra into topology and vice versa. The hope is to turn a difficult topology problem into a manageable algebra problem (or vice versa). To this end, we associate a number of algebraic objects to a given topological space in a nice way, called homology groups. This defines functors from the category of topological spaces to that of Abelian groups. I will not define these terms precisely, but I will indicate where they occur in the construction.

We define the *standard n-simplex* as the topological space [25, Definition 1.11]

$$\Delta^n = \{(t_0, t_1, \ldots, t_n) \in \mathbb{R}^{n+1} : t_i \geq 0,\ \sum_{i=0}^{n} t_i = 1\},$$

which carries the subspace topology induced from $\mathbb{R}^{n+1}$. Note that Δ^0 is a point, Δ^1 is an interval, and Δ^2 is a filled triangle. Also the boundaries $\partial\Delta^n$ of the simplices are of importance. Notice how the boundary of Δ^1 is made out of two copies of Δ^0. Similarly, the boundary of Δ^2 is a triangle, made of three copies of Δ^1.

Next, let X be an arbitrary topological space. Consider the space

$$S(X)_n = \{\sigma : \Delta^n \to X : \sigma \text{ is continuous}\}.$$

This space basically represents all the ways of mapping a point / line / triangle / etc. continuously into X. Furthermore, we let A be an arbitrary Abelian group.[20] Define the space [25, Definition 2.2]

$$C_n(X; A) = A[S(X)_n],$$

called the *A-linearization* of $S(X)_n$. It is the Abelian group consisting of linear combinations of elements of $S(X)_n$ with coefficients in A. Why would that be useful? Consider an element $\sigma \in S(X)_1$. For example, one could imagine $X = \mathbb{R}^3$ and $\sigma : \Delta^1 \to X$ is some curve in X. As discussed earlier, its boundary is a collection of two copies of Δ^0 inside X. Letting $x_0 = \sigma(0, 0)$ and $x_1 = \sigma(1, 1)$, we can define

$$\partial\sigma = x_1 - x_0 \in A[S(X)_0],$$

where we interpret x_i as the function which maps the point Δ^0 to x_i. The minus sign in the above equation takes care of the "orientation". Similarly, the boundary of a solid triangle in $S(X)_2$ is an alternating sum of curves.

[20] We will only need $A = \mathbb{Z}$, but we might as well keep it general.

We can thus define (by linearity) the boundary operator $\partial : C_n(X;A) \to C_{n-1}(X;A)$. This operator has the property that $\partial^2 = \partial \circ \partial = 0$. Geometrically, it should be obvious that the boundary of something which is already a boundary is empty. Mathematically, we obtain this property by a curious cancellation of all the alternating signs in the definition of ∂. For the full definition of ∂ and a proof of $\partial^2 = 0$, see [25] or the more available (and excellent textbook) [26].

The Abelian groups $C_n(X;A)$ together with the boundary operator form what is known as a *chain complex*. It can be visualized by the following chain of arrows:

$$\ldots \xrightarrow{\partial_{n+1}} C_n(X;A) \xrightarrow{\partial_n} C_{n-1}(X;A) \xrightarrow{\partial_{n-1}} \ldots \xrightarrow{\partial_2} C_1(X;A) \xrightarrow{\partial_1} C_0(X;A) \xrightarrow{\partial_0} 0$$

Since $\partial_n \circ \partial_{n+1} = 0$, we have trivially that $\mathrm{im}(\partial_{n+1}) \subseteq \ker(\partial_n)$. Hence, we can consider the quotient group

$$H_n(X;A) = \frac{\ker(\partial_n)}{\mathrm{im}(\partial_{n+1})},$$

which is well-defined, because all groups involved are Abelian. We call this the *n-th (singular) homology group* of the topological space X. Intuitively, these groups can detect holes in X. A nontrivial class can come from a triangle which is not the boundary of a solid triangle in X. This can happen - for example - if $X = \mathbb{R}^2 \setminus \{0\}$ and the triangle goes around the origin.

Now consider two topological spaces X, Y and a continuous map $f : X \to Y$ between them. If $\sigma : \Delta^n \to X$ is an element of $S(X)_n$, then we can define

$$f_*(\sigma) = f \circ \sigma : \Delta^n \to Y$$

Extending this by linearity, we get a map $f_* : C_n(X;A) \to C_n(Y;A)$, which is a group homomorphism by construction. It can be shown [25, p. 11] that this yields a well-defined homomorphism on the quotient group

$$f_* : H_n(X : A) \to H_n(Y;A).$$

Thus, for a fixed $n \geq 0$ and a fixed Abelian group A, we have assigned to every topological space X an Abelian group $H_n(X;A)$. Moreover, every continuous map $f : X \to Y$ induces a group homomorphism f_* (which is more correctly denoted by $H_n(f;A)$) between the corresponding homology groups $H_n(X;A)$ and $H_n(Y;A)$. This assignment satisfies the following two properties [25, p. 12]:

- $H_n(g \circ f;A) = H_n(g;A) \circ H_n(f;A)$;
- $H_n(\mathrm{id}_X;A) = \mathrm{id}_{H_n(X;A)}$;

which makes H_n into a (covariant) functor. Moreover, we have the following important Theorem [25, Lecture 4]:

Theorem A.1. *(Homology of the spheres) Let $n, k \geq 0$. We have:*

$$\begin{cases} H_0(\mathbb{S}^0;\mathbb{Z}) \cong \mathbb{Z} \oplus \mathbb{Z} \\ H_k(\mathbb{S}^0;\mathbb{Z}) \cong \{0\}, \quad k > 0 \\ H_0(\mathbb{S}^n;\mathbb{Z}) \cong \mathbb{Z}, \quad n > 0 \\ H_k(\mathbb{S}^n;\mathbb{Z}) \cong \{0\}, \quad k > 0,\ n > 0, k \neq n \\ H_n(\mathbb{S}^n;\mathbb{Z}) \cong \mathbb{Z}, \quad n > 0. \end{cases} \tag{A.1}$$

As you can see, there are a number of annoying special cases when either n or k equals zero. We can mitigate this in two ways. One way is to simply ignore that $n = 0$ exists and only define homology for $n > 0$. A more mathematical way is to define homology differently for $n = 0$. Consider the unique map c from X into a 1-point space $\{*\}$. It induces a group homomorphism $c_* : H_0(X; A) \to H_0(\{*\}; A)$. We set

$$\tilde{H}_n(X; A) = \begin{cases} \ker(c_*), & n = 0, \\ H_n(X; A), & n > 0. \end{cases} \tag{A.2}$$

which is the *reduced homology group*. One can prove that this still gives a functor from topological spaces to Abelian groups [25, p. 34]. With this new and improved definition, the homology of the spheres becomes

$$\tilde{H}_k(\mathbb{S}^n; \mathbb{Z}) \cong \begin{cases} \{0\}, & k \neq n, \\ \mathbb{Z}, & k = n, \end{cases} \tag{A.3}$$

a result which now holds for all $n \geq 0$.

We can finally define the degree of a continuous map $f : \mathbb{S}^n \to \mathbb{S}^n$. There is an induced group homomorphism

$$f_* : \tilde{H}_n(\mathbb{S}^n; \mathbb{Z}) \cong \mathbb{Z} \to \mathbb{Z} \cong \tilde{H}_n(\mathbb{S}^n; \mathbb{Z}).$$

Recall that $\mathbb{Z}$ is generated by $1 \in \mathbb{Z}$, so the map f_* is completely determined by $f_*(1)$. We have that

$$f_*(n) = n f_*(1) =: n \deg(f).$$

The mapping degree has the following properties:

Lemma A.2. *[25, Lemma 6.6] Let $f, f' : \mathbb{S}^d \to \mathbb{S}^d$ be continuous. Then,*

- *If f and f' are homotopic, then they have the same degree;*
- $\deg(f' \circ f) = \deg(f') \cdot \deg(f)$*;*
- *If f is a homotopy equivalence, then* $\deg(f)$ *is either* 1 *or* -1*.*

We will use the mapping degree for its applications in scattering theory. It can also be used to prove topological statements such as the hairy ball theorem [25, Theorem 6.13]. It says that there is no continuous nowhere vanishing tangent vector field to $\mathbb{S}^{2k}$.

References

[1] B. P. Abbott et al. Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102, 2016.

[2] D. Tong. General relativity. `http://www.damtp.cam.ac.uk/user/tong/gr.html`, 2019. [Online; accessed 11-February-2021].

[3] N. J. Cornish and G. W. Gibbons. A tale of two centres. *Classical and Quantum Gravity*, 14(7):1865, 1997.

[4] G. Contopoulos. Periodic orbits and chaos around two black holes. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 431(1881):183–202, 1990.

[5] J. O. Shipley and S. R. Dolan. Binary black hole shadows, chaotic scattering and the cantor set. *Classical and Quantum Gravity*, 33(17):175001, 2016.

[6] A. Knauf. *Mathematical physics: classical mechanics*, volume 109. Springer, 2018.

[7] M. Klein and A. Knauf. *Classical planar scattering by coulombic potentials*, volume 13. Springer Science & Business Media, 2008.

[8] A. Knauf. Qualitative aspects of classical potential scattering. 1999.

[9] A. Knauf and M. Krapf. The non-trapping degree of scattering. *Nonlinearity*, 21(9):2023, 2008.

[10] A. Zee. *Einstein gravity in a nutshell*, volume 14. Princeton University Press, 2013.

[11] S. D. Majumdar. A class of exact solutions of einstein's field equations. *Physical Review*, 72(5):390, 1947.

[12] A. Papapetrou. A static solution of the equations of the gravitational field for an arbitary charge-distribution. In *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, volume 51, pages 191–204. JSTOR, 1945.

[13] J. B. Hartle and S. W. Hawking. Solutions of the einstein-maxwell equations with many black holes. *Communications in Mathematical Physics*, 26(2):87–101, 1972.

[14] S. Chandrasekhar. *The mathematical theory of black holes*, volume 69. Oxford University Press, 1998.

[15] G.W. Gibbons. Antigravitating black hole solitons with scalar hair in n = 4 supergravity. *Nuclear Physics B*, 207(2):337–349, 1982.

[16] G.W. Gibbons and K. Maeda. Black holes and membranes in higher-dimensional theories with dilaton fields. *Nuclear Physics B*, 298(4):741–775, 1988.

[17] D. Garfinkle, G.T. Horowitz, and A. Strominger. Charged black holes in string theory. *Physical Review D*, 43(10):3140, 1991.

[18] W. Thirring. *Classical mathematical physics: dynamical systems and field theories*. Springer Science & Business Media, 2013.

[19] M. W. Hirsch, S. Smale, and R. L. Devaney. *Differential equations, dynamical systems, and an introduction to chaos*. Academic press, 2012.

[20] S. Wiggins and M. Golubitsky. *Introduction to applied nonlinear dynamical systems and chaos*, volume 2. Springer, 1990.

[21] V. I. Arnol'd. *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media, 2013.

[22] F. C. Santos, V. Soares, and A. C. Tort. An english translation o bertrand's theorem. *arXiv preprint arXiv:0704.2396*, 2007.

[23] D. J. Griffiths and D. F. Schroeter. *Introduction to quantum mechanics*. Cambridge University Press, 2018.

[24] M. Reed and B. Simon. *III: Scattering Theory*, volume 3. Elsevier, 1979.

[25] S. Sagave. Lecture notes for the mastermath course algebraic topology, 2017.

[26] J. J. Rotman. *An introduction to algebraic topology*, volume 119. Springer Science & Business Media, 2013.

[27] J. M. Lee. *Introduction to Smooth Manifolds*. Springer, 2013.

[28] N. Martynchuk and H. Waalkens. Knaufâ€™s degree and monodromy in planar potential scattering. *Regular and Chaotic Dynamics*, 21(6):697–706, 2016.

[29] F. Ravndal. Scalar gravitation and extra dimensions. *arXiv preprint gr-qc/0405030*, 2004.

[30] A. Zee. *Quantum field theory in a nutshell*, volume 7. Princeton university press, 2010.

[31] C. Brans and R. H. Dicke. Mach's principle and a relativistic theory of gravitation. *Physical review*, 124(3):925, 1961.

[32] J. D. Bekenstein. Black holes and the second law. In *JACOB BEKENSTEIN: The Conservative Revolutionary*, pages 303–306. World Scientific, 2020.

[33] S. Carlip. Black hole thermodynamics. *International Journal of Modern Physics D*, 23(11):1430023, 2014.

[34] D. J. Raine and E. G. Thomas. *Black holes: an introduction*. Imperial College Press, 2010.

[35] J. Centrella et al. Black-hole binaries, gravitational waves, and numerical relativity. *Reviews of Modern Physics*, 82(4):3069, 2010.

[36] G.W. Moore. Pitp lectures on bps states and wall-crossing in d= 4, n= 2 theories. *Lecture Notes. http://www. physics. rutgers. edu/˜ gmoore*, 2010.

[37] A. Strominger and C. Vafa. Microscopic origin of the bekenstein-hawking entropy. *Physics Letters B*, 379(1-4):99–104, 1996.

[38] A. Bohn et al. What does a binary black hole merger look like? *Classical and Quantum Gravity*, 32(6):065002, 2015.

[39] M. D. Roberts. Scalar field counterexamples to the cosmic censorship hypothesis. *General relativity and gravitation*, 21(9):907–939, 1989.

[40] C. P. Dettmann, N. E. Frankel, and N. J. Cornish. Fractal basins and chaotic trajectories in multi-black-hole spacetimes. *Physical Review D*, 50(2):R618, 1994.

[41] A. Boscaggin, A. Bottois, and W. Dambrosio. The spatial n-centre problem: scattering at positive energies. *Calculus of Variations and Partial Differential Equations*, 57(5):1–23, 2018.

[42] S. Caron-Huot and Z. Zahraee. Integrability of black hole orbits in maximal supergravity. *Journal of High Energy Physics*, 2019(7):1–43, 2019.