Physics in a Gravitational Field and General Covariance

1 From the Einstein Equivalence Principle to Geodesics

. Motivation: The Einstein Equivalence Principle

The highly successful Newtonian theory of gravity can be succinctly summarised by two sets of differential equations, one describing the dynamics (motion) of particles in a given gravitational field (described by a potential ϕ), and the other describing the dynamics of the gravitational field itself, namely how ϕ is to be determined from a given mass configuration. The former takes the standard Newtonian form

$$m\ddot{\vec{x}} = \vec{F}_g = -m\vec{\nabla}\phi \tag{1.1}$$

(but we will come back in some detail below to the question if/why the same mass parameter m appears on both sides of this equation, so as to incorporate the observation, going back to Galileo, that "all bodies fall at the same rate in a a gravitational field"). The latter is the Poisson equation

$$\Delta\phi = 4\pi G_N \mu = (4\pi G_N/c^2)\rho \ , \tag{1.2}$$

with G_N denoting, here and throughout, Newton's constant, i.e. the gravitational coupling constant, and where μ is the mass density, and $\rho = \mu c^2$ the associated rest mass energy density - I will set $c = 1$ in the following and use ρ.

Let us start with the field equation. It is immediately evident that this cannot be the final story. Not only is this equation not Lorentz invariant. Because of the absence of time-derivatives in (1.2), it actually describes an "action at a distance" and an instantaneous propagation of the gravitational field to every point in space (if you wiggle your mass distribution here now, this will immediately effect the gravitational potential arbitrarily far away). This is something that Einstein had just successfully exorcised from other aspects of physics, and clearly Newtonian gravity had to be revised as well.

It is then also immediately clear that what would have to replace Newton's theory is something rather more complicated. The reason for this is that, according to Special Relativity, mass is just another form of energy. Then, since gravity couples to masses, in a relativistically invariant theory gravity will also have to couple to energy. In particular, therefore, gravity would have to couple to gravitational energy, i.e. to itself. As a consequence, the new gravitational field equations will, unlike Newton's, have to be non-linear: the field of the sum of two masses cannot equal the sum of the gravitational fields of the two masses because it should also take into account the gravitational energy of the two-body system.

Now, having realised that Newton's theory cannot be the final word on the issue, how does one go about finding a better theory?

I will first very briefly discuss (and then dismiss) what at first sight may appear to be the most natural and naive approach to formulating a relativistic theory of gravity, namely the simple replacement of Newton's field equation (1.2) by its relativistically covariant version

$$\Delta\phi = 4\pi G_N \rho \quad \longrightarrow \quad \Box\phi = 4\pi G_N \rho \ , \tag{1.3}$$

where $\Box$ is the Lorentz invariant d'Alembert or wave operator. While this looks promising, something can't be quite right about this equation. We already know (from Special Relativity) that ρ is not a scalar but rather the 00-component of a tensor, the energy-momentum tensor, so if actually ρ appears on the right-hand side, ϕ cannot be a scalar, while if ϕ is a scalar something needs to be done to fix the right-hand side.

Turning first to the latter possiblility, one option that suggests itself is to replace ρ by the trace $T = T^\alpha_\alpha$ of the energy-momentum tensor. This is by definition / construction a scalar, and it will agree with ρ in the non-relativistic limit (where rest mass dominates over other contributions). Thus a first attempt at fixing the above equation might look like

$$\Box\phi = 4\pi G_N T \ . \tag{1.4}$$

This is certainly an attractive equation, but it definitely has the drawback that it is too linear. Recall from the discussion above that the universality of gravity (coupling to all forms of matter) and the equivalence of mass and energy lead to the conclusion that gravity should couple to gravitational energy, invariably predicting non-linear (self-interacting) equations for the gravitational field. However, the left hand side could be such that it only reduces to $\Box$ or Δ of the Newtonian potential in the Newtonian limit of weak time-independent fields. Thus a second attempt at fixing the above equation might look like

$$\Box\Phi(\phi) = 4\pi G_N T \ , \tag{1.5}$$

where $\Phi(\phi) \approx \phi$ for weak fields.

Such scalar relativistic theories of gravity (or rather some minor variants thereof) were proposed and studied among others by Abraham, Mie, and Nordstrøm. As it stands, the field equation appears to be perfectly consistent (and it may be interesting to discuss if/how the Einstein equivalence principle, which will put us on our route towards metrics and space-time curvature is realised in such a theory). However, regardless of this, this theory is incorrect simply because it is ruled out experimentally. The easiest way to see this (with hindsight) is to note that the energy-momentum tensor of Maxwell theory (6.47) is traceless (6.121), and thus the above equation would predict no coupling of gravity to the electro-magnetic field, in particular to light, hence in such a theory there would be no deflection of light by the sun etc.[1]

The other possibility to render (1.3) consistent is the, a priori perhaps much less compelling, option to think of ϕ and $\Delta\phi$ or $\Box\phi$ not as scalars but as (00)-components of

[1]For more on the history and properties of scalar theories of gravity see the review by D. Giulini, *What is (not) wrong with scalar gravity?*

some tensor, in which case one could try to salvage (1.3) by promoting it to a tensorial equation

$$\{\text{Some tensor generalising } \Delta\phi\}_{\alpha\beta} \sim 4\pi G_N T_{\alpha\beta} \quad . \tag{1.6}$$

This is indeed the form of the field equations for gravity (the Einstein equations) we will ultimately be led to (see section 18.4), but Einstein arrived at this in a completely different, and much more insightful, way.

Let us now, very briefly and in a streamlined way, try to retrace (one aspect of) Einstein's thoughts, namely on the relation between inertial and gravitational mass, which, as we will see, will lead us rather quickly to the geometric picture of gravity sketched in the Introduction.

To that end we return to the Newtonian equation of motion (1.1). Recall that in this Newtonian theory, there are two *a priori* completely independent concepts of mass:

- inertial mass m_i (or acceleration mass), which accounts for the resistance of a body or particle against acceleration and appears universally on the left-hand-side of the Newtonian equation of motion

  $$m_i\vec{a} = \vec{F} \tag{1.7}$$

  in conjunction with the acceleration $\vec{a}$;
- gravitational mass m_g which is the mass the gravitational field couples to, i.e. it is the gravitational charge of a particle,

  $$\vec{F}_g = -m_g\vec{\nabla}\phi \quad . \tag{1.8}$$

Now it is an important empirical fact that the inertial mass of a body is equal to its gravitational mass. This realisation, at least with this clarity, is usually attributed to Newton, although it goes back to experiments and observations by Galileo usually paraphrased as "all bodies fall at the same rate in a gravitational field". (It is not true, though, that Galileo dropped objects from the leaning tower of Pisa to test this - he used an inclined plane, a water clock and a pendulum).

These experiments were later on improved, in various forms, by Huygens, Newton, Bessel and others and reached unprecedented accuracy with the work of Baron von Eötvös (1889-. . .), who was able to show that inertial and gravitational mass of different materials (like wood and platinum) agree to one part in 10^9. In the 1950/60's, this was still further improved by R. Dicke to something like one part in 10^{11}. More recently, rumours of a 'fifth force', based on a reanalysis of Eötvös' data (but buried in the meantime) motivated experiments with even higher accuracy and no difference between m_i and m_g was found.

Newton's theory would in principle be perfectly consistent with $m_i \neq m_g$, just as the formally analogous equation for an electrically charged particle with charge q_e in an electrostatic field $\vec{E} = -\vec{\nabla}\phi$,

$$m_i\ddot{\vec{x}} = -q_e\vec{\nabla}\phi \quad , \tag{1.9}$$

is perfectly acceptable for any ratio q_e/m_i, and Einstein was very impressed with the observed equality of m_i and m_g. This should, he reasoned, not be a mere coincidence but is probably trying to tell us something rather deep about the nature of gravity.

To see what this could be, let us recall that there is a very common class of "forces" for which the equality between the inertial mass and the coupling constant is actually built in and automatic. These are the "pseudo-forces" or "fictitious forces" $\vec{P}$ (like centrifugal forces) which arise when one transforms the Newtonian equations of motion via a non-linear coordinate transformation to accelerated (or other non-Cartesian) coordinates,

$$x^i \to z^m = z^m(x^i) \tag{1.10}$$

(like spherical coordinates). These "forces" arise from the non-trivial transformation behaviour of the acceleration $\ddot{\vec{x}}$ under such non-linear coordinate transformations, and are therefore inevitably and automatically proportional to m_i,

$$m_i\ddot{\vec{x}} = \vec{F} \quad \Rightarrow \quad m_i\ddot{\vec{z}} = \vec{F} + \vec{P} \quad \text{with} \quad \vec{P} \sim m_i \tag{1.11}$$

(cf. section 1.4 for the explicit expression for $\vec{P}$ when one performs an analogous calculation in relativistic mechanics in Minkowski space). Conversely, such pseudo-forces can be eliminated by transforming the equations of motion to a suitable (inertial) coordinate system or reterence system.

With his unequalled talent for discovering profound truths in such simple observations, he concluded (calling this "der glücklichste Gedanke meines Lebens" (the happiest thought of my life)) that the equality of inertial and gravitational mass suggests a close relation between inertia and gravity itself, suggests, in fact, that locally effects of gravity and acceleration (or non-linear transformations of the reference system) are indistinguishable,

gravitational mass = inertial mass <u>because</u> (locally) GRAVITY = ACCELERATION

He substantiated this with some classical thought experiments, *Gedankenexperimente*, as he called them, which in the meantime have morphed into and have come to be known as the *elevator thought experiments*, which we will now discuss.

1. Consider somebody in a small sealed box (elevator) somewhere in outer space. In the absence of any forces, this person will float. Likewise, two stones he has just dropped (see Figure 1) will float with him.

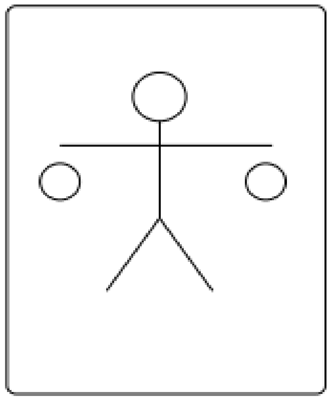

Figure 1: Experimenter and his two stones freely floating somewhere in outer space, i.e. in the absence of forces.

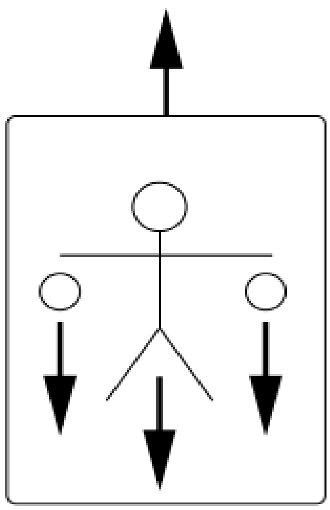

Figure 2: Constant acceleration upwards mimics the effect of a gravitational field: experimenter and stones drop to the bottom of the box.

2. Now assume (Figure 2) that somebody on the outside suddenly pulls the box up with a constant acceleration. Then of course, our friend will be pressed to the bottom of the elevator with a constant force and he will also see his stones drop to the floor.

3. Now consider (Figure 3) this same box brought into a constant gravitational field. Then again, he will be pressed to the bottom of the elevator with a constant force and he will see his stones drop to the floor. With no experiment inside the elevator can he decide if this is actually due to a gravitational field or due to the fact that somebody is pulling the elevator upwards.

   Thus our first lesson is that, indeed, locally the effects of acceleration and gravity are indistinguishable.

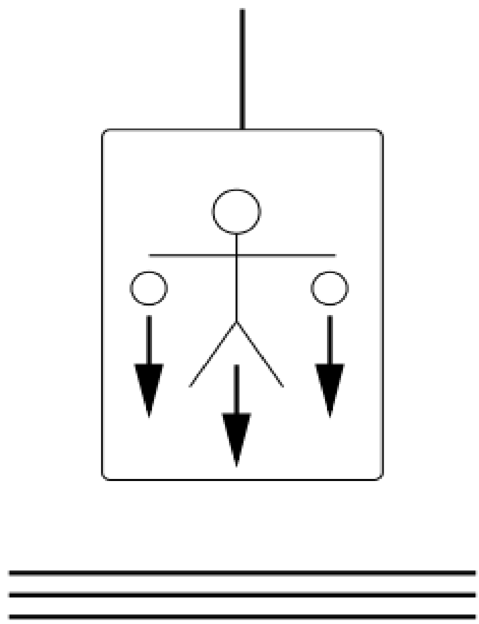

Figure 3: Effect of a constant gravitational field: indistinguishable for our experimenter from that of a constant acceleration in Figure 2.

4. Now consider somebody cutting the cable of the elevator (Figure 4). Then the elevator will fall freely downwards but, as in Figure 1, our experimenter and his stones will float as in the absence of gravity.

   Thus lesson number two is that, locally the effect of gravity can be eliminated by going to a freely falling reference frame (or coordinate system). This should not come as a surprise. In the Newtonian theory, if the free fall in a constant gravitational field is described by the equation

$$\ddot{x} = g \;\; (+ \text{ other forces}) \;\; , \tag{1.12}$$

   then in the accelerated coordinate system

$$\xi(x,t) = x - gt^2/2 \tag{1.13}$$

   the same physics is described by the equation

$$\ddot{\xi} = 0 \;\; (+ \text{ other forces}) \;\; , \tag{1.14}$$

   and the effect of gravity has been eliminated by going to the freely falling coordinate system ξ. The crucial point here is that in such a reference frame not only our observer will float freely, but because of the equality of inertial and gravitational mass he will also observe all other objects obeying the usual laws of motion in the absence of gravity.

5. In the above discussion, I have put the emphasis on constant accelerations and on 'locally'. To see the significance of this, consider our experimenter with his

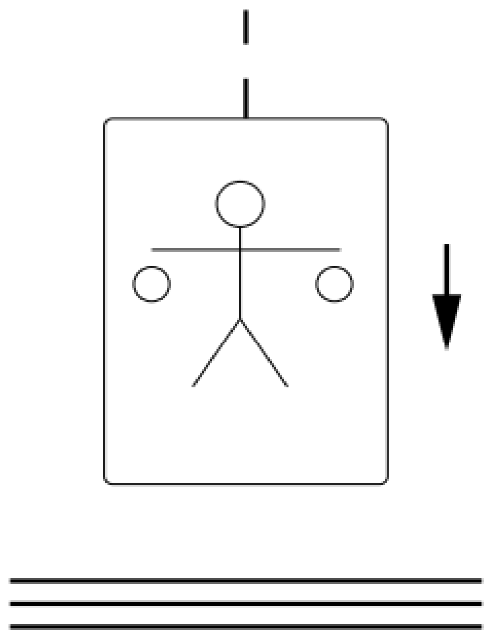

Figure 4: Free fall in a gravitational field has the same effect as no gravitational field (Figure 1): experimenter and stones float.

elevator in the gravitational field of the earth (Figure 5). This gravitational field is not constant but spherically symmetric, pointing towards the center of the earth. Therefore the stones will slightly approach each other as they fall towards the bottom of the elevator, in the direction of the center of the gravitational field.

6. Thus, if somebody cuts the cable now and the elevator is again in free fall (Figure 6), our experimenter will float again, so will the stones, but our experimenter will also notice that the stones move closer together for some reason. He will have to conclude that there is some force responsible for this.

   This is lesson number three: in a non-uniform gravitational field the effects of gravity cannot be eliminated by going to a freely falling coordinate system. This is only possible locally, on such scales on which the gravitational field is essentially constant.

Einstein formalised the outcome of these thought experiments in what is now known as the *Einstein Equivalence Principle* which roughly states that physics in a freely falling frame in a gravitational field is the same as physics in an inertial frame in Minkowski space in the absence of gravitation. Two formulations are

> At every space-time point in an arbitrary gravitational field it is possible to choose a *locally inertial* (or *freely falling*) *coordinate system* such that, within a sufficiently small region of this point, the laws of nature take the

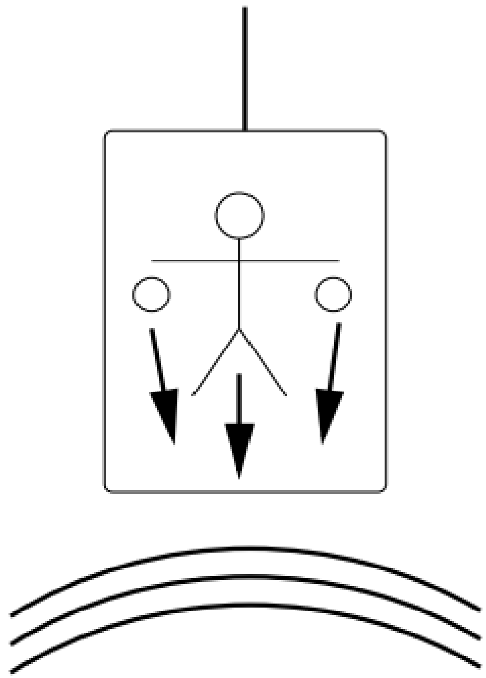

Figure 5: Experimenter and his stones in a non-uniform gravitational field: the stones will approach each other slightly as they fall to the bottom of the elevator.

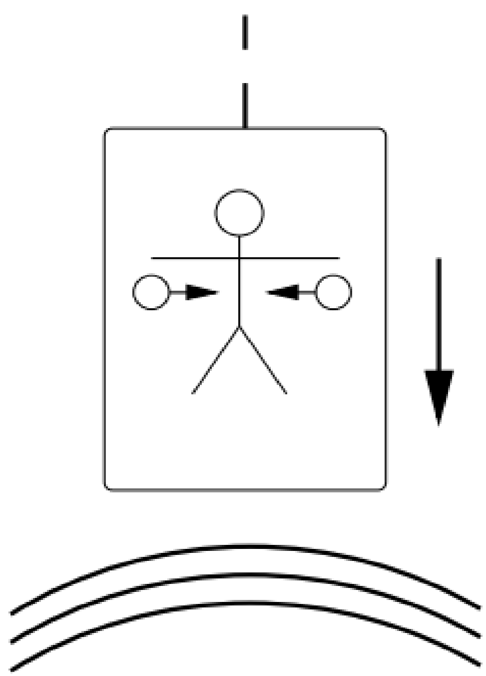

Figure 6: Experimentator and stones freely falling in a non-uniform gravitational field. The experimenter floats, so do the stones, but they move closer together, indicating the presence of some external force.

> same form as in unaccelerated Cartesian coordinate systems in the absence of gravitation.[2]

and

> Experiments in a sufficiently small freely falling laboratory, over a sufficiently short time, give results that are indistinguishable from those of the same experiments in an inertial frame in empty space.[3]

There are different versions of this principle depending on what precisely one means by 'the laws of nature'. If one just means the laws of Newtonian (or relativistic) mechanics, then this principle essentially reduces to the statement that inertial and gravitational mass are equal. Usually, however, this statement is taken to imply also Maxwell's theory, quantum mechanics etc.[4] What it pragmatically asserts in one of its stronger forms is that

> [. . .] there is no experiment that can distinguish a uniform acceleration from a uniform gravitational field. (J. Hartle, loc. cit.)

The power of the above principle, which we will regard as a heuristic guideline, rather than trying to (prematurely) give it a mathematically precise formulation, lies in the fact that we can combine it with our understanding of physics in accelerated reference systems to gain insight into the physics in a gravitational field. Two immediate consequences of this (which cannot be derived on the basis of Newtonian physics or Special Relativity alone) are

- light is deflected by a gravitational field just like material objects;
- clocks run slower in a gravitational field than in the absence of gravity.

To see the inevitability of the first assertion, imagine a light ray entering the rocket / elevator in Figure 1 horizontally through a window on the left hand side and exiting again at the same height through a window on the right. Now imagine, as in Figure 2, accelerating the elevator upwards. Then clearly the light ray that enters on the left will exit at a lower point of the elevator on the right because the elevator is accelerating upwards. By the equivalence principle one should observe exactly the same thing in a constant gravitational field (Figure 3). It follows that in a gravitational field the light

[2] S. Weinberg, *Gravitation and Cosmology*.

[3] J. Hartle, *Gravity. An Introduction to Einstein's General Relativity*.

[4] For a discussion of different formulations of the equivalence principle and the logical relations among them, see E. di Casola, S. Liberati, S. Sonego, *Nonequivalence of equivalence principles*.

ray is bent downwards, i.e. it experiences a downward acceleration with the (locally constant) gravitational acceleration g.

To understand the second assertion, one can e.g. simply appeal to the so-called "twin-paradox" of Special Relativity: the accelerated twin is younger than his unaccelerated inertial sibling. Hence accelerated clocks run slower than inertial clocks. Hence, by the equivalence principle, clocks in a gravitational field run slower than clocks in the absence of gravity.

Alternatively, one can imagine two observers at the top and bottom of the elevator, having identical clocks and sending light signals to each other at regular intervals as determined by their clocks. Once the elevator accelerates upwards, the observer at the bottom will receive the signals at a higher rate than he emits them (because he is accelerating towards the signals he receives), and he will interpret this as his clock running more slowly than that of the observer at the top. By the equivalence principle, the same conclusion now applies to two observers at different heights in a gravitational field. This can also be interpreted in terms of a gravitational redshift or blueshift (photons losing or gaining energy by climbing or falling in a gravitational field), and we will return to a more quantitative discussion of this effect in section 2.10.

Lorentz-Covariant Formulation of Special Relativity (Review)

What the equivalence principle tells us is that we can expect to learn something about the effects of gravitation by transforming the laws of nature (equations of motion) from an inertial Cartesian coordinate system to other (accelerated, curvilinear) coordinates. As a first step, we will, in section 1.3 below, discuss the above example of an observer undergoing constant acceleration in the context of special relativity.

As a preparation for this, and the remainder of the course, this section will provide a lightning review of the Lorentz-covariant formulation of special relativity, mainly to set the notation and conventions that will be used throughout, and only to the extent that it will be used in the following.

1. Minkowski space(-time)

   (a) The arena of special relativity is Minkowski space-time [henceforth *Minkowski space* for short, the union of space and time is implied by uttering the word "Minkowski"]. It is the space of events, labelled by 3 Cartesian spatial coordinates x^k and a time-coordinate t or, more usefully, by the coordinates

   $$(\xi^a) = (\xi^0 = ct, \xi^k = x^k) \ , \tag{1.15}$$

   where c is the speed of light. Typically in these notes ξ^a will indicate such a (locally) inertial coordinate system, whereas generic coordinates will be called x^μ etc. We will almost always work in units in which $c = 1$.

(b) Minkowski space is equipped with a prescription for measuring distances, encoded in a line-element which, in these coordinates, takes the form

$$ds^2 = -(d\xi^0)^2 + \sum_k (d\xi^k)^2 \quad . \tag{1.16}$$

(c) This can be written as

$$ds^2 = -(d\xi^0)^2 + \sum_k (d\xi^k)^2 \equiv \eta_{ab}\, d\xi^a d\xi^b \quad . \tag{1.17}$$

with metric $(\eta_{ab}) = \text{diag}(-1, +1, +1, +1)$ or, more explicitly,

$$\eta_{ab} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} \tag{1.18}$$

(thus we are using the "mostly plus" convention).

2. Lorentz Transformations

(a) Lorentz transformations are by definition those linear transformations

$$\xi^a \mapsto \bar{\xi}^a = L^a_{\ \beta}\xi^b \tag{1.19}$$

that leave the Minkowski line-element invariant,

$$d\bar{s}^2 \equiv \eta_{ab} d\bar{\xi}^a d\bar{\xi}^b = \eta_{ab} d\xi^a d\xi^b = ds^2 \quad \Leftrightarrow \quad \eta_{ab} L^a_{\ c} L^b_{\ d} = \eta_{cd} \quad . \tag{1.20}$$

In matrix notation this can also be written as

$$\bar{\xi} = L\xi \quad , \quad L^t \eta L = \eta \tag{1.21}$$

where L^t is the transpose of L. This is the defining condition for Lorentz transformations, and the Lorentz signature analogue of the condition $R^t \mathbb{1} R = \mathbb{1}$ for an orthogonal transformation (rotation or reflection) in Euclidean space, with metric $\eta_{ab} \to \mathbb{1}_{ik} = \delta_{ik}$.

Alternative notation:

$$\bar{\xi}^{\bar{a}} = L^{\bar{a}}_{\ b}\xi^b \quad \text{or} \quad \xi^{\bar{a}} = L^{\bar{a}}_{\ b}\xi^b \tag{1.22}$$

Strictly speaking $\bar{\xi}^a$ and $\xi^{\bar{a}}$ may be considered to refer to two different quantities, to the coordinates of the new point $\bar{\xi}$ in the old coordinate system versus the coordinates of the old point ξ in the new coordinate system. However, for many elementary purposes this difference between what is known as the active (moving points) versus the passive (relabelling points) point of view is not crucial, and one should not be hung-up on notation: coordinates are fundamentally just bookkeeping devices so use whatever is convenient for current bookkeeping or other purposes.

(b) Infinitesimal Lorentz rotations, i.e. Lorentz transformations with L of the form $L = 1 + \omega$, ω infinitesimal, are characterised by

$$(1+\omega)^t \eta (1+\omega) = \eta \quad \Rightarrow \quad (\eta\omega) + (\eta\omega)^t = 0 \ . \tag{1.23}$$

Thus the matrix $\eta\omega$ is anti-symmetric. In components, an infinitesimal Lorentz transformation therefore has the form

$$\delta\xi^a = \omega^a_{\ b}\xi^b \quad \text{with} \quad \omega_{ab} \equiv \eta_{ac}\omega^c_{\ b} = -\omega_{ba} \ . \tag{1.24}$$

(c) Poincaré transformations are those affine transformations that leave the Minkowski line-element invariant. They are composed of Lorentz transformations and arbitrary constant translations and thus have the form

$$\xi^a \mapsto \bar{\xi}^a = L^a_{\ \beta}\xi^b + \zeta^a \ , \tag{1.25}$$

infinitesimally

$$\delta\xi^a = \omega^a_{\ b}\xi^b + \epsilon^a \ . \tag{1.26}$$

Any two inertial systems in the sense of the equivalence principle of special relativity are related by a Poincaré transformation.

3. Distance & Causal Structure

(a) The Minkowski metric defines the Lorentz (and Poincaré) invariant distance

$$(\Delta\xi)^2 = \eta_{ab}(\xi^a_P - \xi^a_Q)(\xi^b_P - \xi^b_Q) \tag{1.27}$$

betwen two events P and Q with coordinates ξ^a_P and ξ^a_Q respectively.

(b) Depending on the sign of $(\Delta\xi)^2$, the two events P, Q are called, spacelike, lightlike (null) or timelike separated,

$$(\Delta\xi)^2 = \begin{cases} > 0 & \text{spacelike} \quad \text{separated} \\ = 0 & \text{lightlike} \quad \text{separated} \\ < 0 & \text{timelike} \quad \text{separated} \end{cases} \tag{1.28}$$

(c) The set of events that are lightlike separated from P define the lightcone at P. It consists of two components (joined at P), the future and the past lightcone, distinguished by the sign of $\xi^0_Q - \xi^0_P$ (positive for Q on the future lightcone, $\xi^0_Q > \xi^0_P$, negative for Q on the past lightcone).

4. Curves and Tangent Vectors

(a) A parametrised curve is given by a map $\lambda \mapsto \xi^a(\lambda)$. The tangent vector to the curve at the point $\xi(\lambda_0)$ has components

$$\xi'^a(\lambda_0) = \frac{d}{d\lambda}\xi^a(\lambda)|_{\lambda=\lambda_0} \ . \tag{1.29}$$

It is called spacelike, lightlike (null) or timelike, depending on the sign of $\eta_{ab}\xi'^a\xi'^b$,

$$\eta_{ab}\xi'^a\xi'^b \quad \begin{cases} > 0 & \text{spacelike} \\ = 0 & \text{lightlike} \\ < 0 & \text{timelike} \end{cases} \tag{1.30}$$

This sign (and hence this classification) depends only on the image of the curve, not its parametrisation.

(b) A curve whose tangent vector is everywhere timelike is called a timelike curve (and likewise for lightlike and spacelike curves). A curve whose tangent vector is everywhere timelike or null (i.e. non-spacelike) is called a causal curve. Worldlines of massive particles are timelike curves, those of massless particles (light) are null curves.

(c) A natural Lorentz-invariant parametrisation of timelike curves is provided by the Lorentz-invariant proper time τ along the curves,

$$\xi^a = \xi^a(\tau) \quad , \tag{1.31}$$

with

$$\begin{aligned} c d\tau &= \sqrt{-ds^2} = \sqrt{-\eta_{ab}d\xi^a d\xi^b} = \sqrt{-\eta_{ab}\xi'^a\xi'^b}d\lambda \\ \Rightarrow \quad & \eta_{ab}\frac{d\xi^a(\tau)}{d\tau}\frac{d\xi^b(\tau)}{d\tau} = -c^2 \quad . \end{aligned} \tag{1.32}$$

Likewise spacelike curves are naturally parametrised by proper distance ds. The derivative with respect to proper time will be denoted by an overdot,

$$\dot{\xi}^a(\tau) = \frac{d}{d\tau}\xi^a(\tau) \quad . \tag{1.33}$$

Because τ is Lorentz-invariant, $\bar{\tau} = \tau$, tangent vectors $\dot{\xi}^a$ of τ-parametrised curves transform linearly under Lorentz transformations,

$$\dot{\bar{\xi}}^a(\tau) = \frac{d}{d\tau}\bar{\xi}^a(\tau) = \frac{\partial\bar{\xi}^a}{\partial\xi^b}\frac{d}{d\tau}\xi^b(\tau) = L^a_{\ b}\dot{\xi}^b(\tau) \quad . \tag{1.34}$$

These are the prototypes of what are called Lorentz vectors or, more generally, Lorentz tensors.

5. Lorentz Vectors

(a) Lorentz vectors (or 4-vectors) are objects with components v^a which transform under Lorentz transformations with the matrix $L^a_{\ b}$ (to be thought of as the Jacobian of the transformation relating $\bar{\xi}^a$ and ξ^a),

$$\bar{v}^a = L^a_{\ b}v^b \quad . \tag{1.35}$$

(b) η_{ab} can be regarded as defining an indefinite scalar product on the space of Lorentz vectors, and the Minkowski norm $\eta_{ab}v^a v^b$ and the Minkowski scalar product $\eta_{ab}v^a w^b$ of Lorentz vectors are invariant under Lorentz transformations,

$$\eta_{ab}\bar{v}^a\bar{v}^b = \eta_{ab}v^a v^b \quad , \quad \eta_{ab}\bar{v}^a\bar{w}^b = \eta_{ab}v^a w^b \ . \tag{1.36}$$

A vector is called, spacelike, lightlike (null) or timelike depending on the sign of its Minkowski norm.

(c) One can identify Minkowski space with its "tangent space", i.e. with the vector space $\mathbb{V} = \mathbb{R}^{1,3}$ of 4-vectors equipped with the quadratic form or scalar product η_{ab} with signature (1,3).

6. Other Lorentz Tensors

(a) Lorentz scalars are objects that are invariant under Lorentz transformations. Examples are scalar products and norms of Lorentz vectors.

(b) Lorentz covectors are objects u_a that transform under Lorentz transformations with the dual (or contragredient = transpose inverse) representation

$$\Lambda = (L^t)^{-1} = \eta L \eta^{-1} \ , \tag{1.37}$$

i.e.

$$\bar{u}_a = \Lambda_a^{\ b} u_b \quad , \quad \Lambda_a^{\ b} = \eta_{ac} L^c_{\ d}\eta^{db} \ , \tag{1.38}$$

where η^{ab} denotes the components of the inverse metric η^{-1}. In terms of Λ, the condition that L is a Lorentz transformation (i.e. preserves η_{ab}) can evidently be written as

$$L^t \eta L = \eta \quad \Leftrightarrow \quad \Lambda\eta\Lambda^t = \eta \quad \Leftrightarrow \quad \Lambda_a^{\ c}\Lambda_b^{\ d}\eta_{cd} = \eta_{ab} \ . \tag{1.39}$$

Covectors can be regarded as elements of the dual $\mathbb{V}^*$ of the space $\mathbb{V}$ of 4-vectors, with u_a defining the Lorentz-invariant linear mapping

$$u: \ v \in \mathbb{V} \mapsto u(v) = u_a v^a \in \mathbb{R} \ . \tag{1.40}$$

Examples are $u_a = \eta_{ab}v^b \equiv v_a$ with v^a a Lorentz vector, the scalar product η_{ab} defining the isomorphism $\mathbb{V}^* \cong \mathbb{V}$.

(c) Lorentz (p, q)-tensors are objects that transform under Lorentz transformations like a product of p vectors and q covectors,

$$T^{a_1\ldots a_p}_{c_1\ldots c_q} \to \bar{T}^{a_1\ldots a_p}_{c_1\ldots c_q} = L^{a_1}_{\ b_1}\ldots L^{a_p}_{\ b_p}\Lambda_{c_1}^{\ d_1}\ldots\Lambda_{c_q}^{\ d_q}\, T^{b_1\ldots b_p}_{d_1\ldots d_q} \ . \tag{1.41}$$

In particular, direct products of vectors and covectors like $V^a W^b U_c$ are tensors. A special case is η_{ab}, which is a Lorentz-invariant $(0, 2)$-tensor by definition,

$$\bar{\eta}_{ab} = \Lambda_a^{\ c}\Lambda_b^{\ d}\eta_{cd} = \eta_{ab} \ . \tag{1.42}$$

Linear combinations of (p, q)-tensors are again (p, q)-tensors. Arbitrary products and contractions of Lorentz tensors are again Lorentz tensors (and the tensor type can be read off from the number and position of the "free" indices).

7. Tensor Fields

   (a) Lorentz tensor fields are assignments of Lorentz tensors to each point of Minkowski space,
   $$T : \xi \mapsto T^{a_1 \dots a_p}_{c_1 \dots c_q}(\xi) \quad . \tag{1.43}$$

   (b) Given a vector field $V^a(\xi)$, $\eta_{ab}V^a(\xi)V^b(\xi)$ is an example of a scalar field, and given a scalar field $f(\xi)$, its partial derivatives give a covector field
   $$U_a(\xi) = \partial_{\xi^a} f(\xi) \equiv \partial_a f(\xi) \tag{1.44}$$
   (providing the justification for abbreviating $\partial_{\xi^a} = \partial_a$). More generally, the partial derivatives of the components of a (p, q)-tensor,
   $$T^{a_1 \dots a_p}_{c_1 \dots c_q}(\xi) \quad \rightarrow \quad \partial_a T^{a_1 \dots a_p}_{c_1 \dots c_q}(\xi) \tag{1.45}$$
   are the components of a $(p, q+1)$-tensor, and the wave operator
   $$\Box = \eta^{ab}\partial_a \partial_b \tag{1.46}$$
   is a Lorentz-invariant differential operator mapping (p, q) tensor fields to (p, q) tensor fields.

   (c) Tensorial equations of the form
   $$T^{a_1 \dots a_p}_{c_1 \dots c_q}(\xi) = 0 \tag{1.47}$$
   are Lorentz invariant in the sense that they are satisfied in one inertial system iff they are satisfied in all inertial systems. (Here and throughout these notes "iff" is the usual mathematicians' shorthand for "if and only if".)

8. Worldlines of Massive Particles

   (a) In the covariant formulation, the timelike worldline of a massive particle is parametrised by proper time, $\xi^a = \xi^a(\tau)$. The velocity (tangent) vector
   $$u^a \equiv \dot{\xi}^a(\tau) \tag{1.48}$$
   is a Lorentz vector, normalised as
   $$u^a u_a \equiv \eta_{ab} u^a u^b = -c^2 \quad . \tag{1.49}$$

(b) The Lorentz-covariant acceleration is the 4-vector

$$a^c = \frac{d}{d\tau}u^c = \frac{d^2}{d\tau^2}\xi^c \quad , \tag{1.50}$$

and the equation of motion of a massive free particle is

$$a^c = \frac{d^2}{d\tau^2}\xi^c(\tau) = 0 \quad . \tag{1.51}$$

We will study this equation further (in any arbitrary coordinate system) in section 1.4. For observers with non-zero acceleration it follows from (1.49) by differentiation that a^c is orthogonal to u^b,

$$a^c u_c \equiv \eta_{cb}a^c u^b = 0 \quad , \tag{1.52}$$

and therefore spacelike,

$$\eta_{cb}a^c a^b \equiv \mathsf{a}^2 > 0 \quad . \tag{1.53}$$

Observers with constant acceleration will be the subject of section 1.3.

(c) The action for a free massive particle with worldline $\xi^a(\tau)$ is essentially the total proper time along the path,

$$S[\xi] = -mc^2\int d\tau = -mc\int\sqrt{-\eta_{ab}d\xi^a d\xi^b} \quad , \tag{1.54}$$

worldlines of free massive particles extremising (maximising) the proper time. In terms of an arbitrary parametrisation $\xi^a = \xi^a(\lambda)$ of the path, this action can be written as

$$S[\xi] = \int d\lambda\, L_\lambda \quad , \quad L_\lambda = -mc\left(-\eta_{ab}\frac{d\xi^a}{d\lambda}\frac{d\xi^b}{d\lambda}\right)^{1/2} \quad . \tag{1.55}$$

A special choice is $\lambda = t$, for which

$$L_t = -mc^2\sqrt{1-\vec{v}^2/c^2} \qquad \vec{v} = d\vec{\xi}/dt = d\vec{x}/dt \quad . \tag{1.56}$$

9. Energy-Momentum 4-Vector

(a) The covariant momenta p_a are defined by

$$p_a = \frac{\partial L_\lambda}{\partial(d\xi^a/d\lambda)} = m\eta_{ab}u^b \quad \Rightarrow \quad p^a = mu^a = m(d\xi^a/d\tau) \tag{1.57}$$

(independently of the choice of λ).

(b) Its components are

$$p^0 = E/c \quad , \quad p^k = p^{(c)k} \tag{1.58}$$

where $p^{(c)k}$ are the canonical momenta associated to the Lagrangian L_t,

$$p_k^{(c)} = \frac{\partial L_t}{\partial v^k} = m\gamma(v)v_k \quad , \tag{1.59}$$

with $\gamma(v) = (1-\vec{v}^2/c^2)^{-1/2}$, and $E = H$ is the corresponding energy or Hamiltonian

$$H = p_k^{(c)}v^k - L_t = m\gamma(v)c^2 \quad . \tag{1.60}$$

(c) The p^a are the components of a Lorentz vector, the energy-momentum 4-vector. It satisfies the mass-shell relation

$$\eta_{ab}p^a p^b = -m^2c^2 \quad \Leftrightarrow \quad E^2 = m^2c^4 + \vec{p}^2c^2 \ . \tag{1.61}$$

1.3 Accelerated Observers and the Rindler Metric

We return to the issue discussed in the context of the Einstein equivalence principle in section 1.1, namely physics as experienced by an observer undergoing constant acceleration (as a precursor to studying this observer in a genuine gravitational field), now specifically within the framework of special relativity.

Specialising (1.52) to an observer accelerating in the ξ^1-direction (so that in the momentary restframe of this observer one has $u^a = (1,0,0,0), a^a = (0,\mathsf{a},0,0)$), we will say that the observer undergoes constant acceleration if a is time-independent. To determine the worldline of such an observer, we note that the general solution to (1.49) with $u^2 = u^3 = 0$,

$$\eta_{ab}u^a u^b = -(u^0)^2 + (u^1)^2 = -1 \ , \tag{1.62}$$

is

$$u^0 = \cosh F(\tau) \quad , \quad u^1 = \sinh F(\tau) \tag{1.63}$$

for some function $F(\tau)$. Thus the acceleration is

$$a^a = \dot{F}(\tau)(\sinh F(\tau), \cosh F(\tau), 0, 0) \ , \tag{1.64}$$

with norm

$$\mathsf{a}^2 = \dot{F}^2 \ , \tag{1.65}$$

and an observer with constant acceleration is characterised by $F(\tau) = \mathsf{a}\tau$,

$$u^a(\tau) = (\cosh \mathsf{a}\tau, \sinh \mathsf{a}\tau, 0, 0) \ . \tag{1.66}$$

This can now be integrated, and in particular

$$\xi^a(\tau) = (\mathsf{a}^{-1}\sinh \mathsf{a}\tau, \mathsf{a}^{-1}\cosh \mathsf{a}\tau, 0, 0) \tag{1.67}$$

is the worldline of an observer with constant acceleration a and initial condition $\xi^a(\tau = 0) = (0, \mathsf{a}^{-1}, 0, 0)$. The worldlines of this observer is the hyperbola

$$\eta_{ab}\xi^a\xi^b = -(\xi^0)^2 + (\xi^1)^2 = \mathsf{a}^{-2} \tag{1.68}$$

in the quadrant $\xi^1 > |\xi^0|$ of Minkowski space-time.

We can now ask the question what the Minkowski metric or line-element looks like in the restframe of such an observer. Note that one cannot expect this to be again the constant Minkowski metric η_{ab}: the transformation to an accelerated reference system,

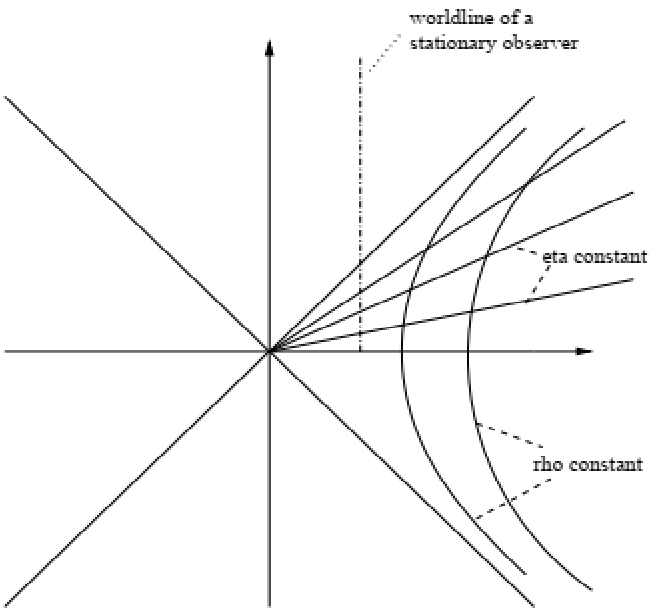

Figure 7: Rindler metric: Rindler coordinates (η, ρ) cover the first quadrant $\xi^1 > |\xi^0|$. Indicated are lines of constant ρ (hyperbolas, worldlines of constantly accelerating observers) and lines of constant η (straight lines through the origin). The quadrant is bounded by the lightlike lines $\xi^0 = \pm\xi^1 \Leftrightarrow \eta = \pm\infty$. An inertial observer reaches and crosses the line $\eta = \infty$ in finite proper time $\tau = \xi^0$.

while certainly allowed in special relativity, is not a Lorentz transformation, while η_{ab} is, by definition, invariant under Lorentz-transformations.

We are thus looking for coordinates that are adapted to these accelerated observers in the same way that the inertial coordinates are adapted to static observers (ξ^0 is proper time, and the spatial components ξ^i remain constant). In other words, we seek a coordinate transformation $(\xi^0, \xi^1) \to (\eta, \rho)$ such that the worldlines of these accelerated observers are characterised by $\rho =$ constant (this is what we mean by restframe, the observer stays at a fixed value of ρ) and ideally such that then η is proportional to the proper time of the observer.

Comparison with (1.67) suggests the coordinate transformation

$$\xi^0(\eta, \rho) = \rho \sinh \eta \qquad \xi^1(\eta, \rho) = \rho \cosh \eta \ . \tag{1.69}$$

It is now easy to see that in terms of these new coordinates the 2-dimensional Minkowski metric $ds^2 = -(d\xi^0)^2 + (d\xi^1)^2$ (we are now suppressing, here and in the remainder of this subsection, the transverse spectator dimensions 2 and 3) takes the form

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 \ . \tag{1.70}$$

This is the so-called *Rindler metric*.

Let us try to gain a better understanding of these Rindler coordinates (illustrated in Figure 7 - see also Figure 25 in section 27.4 for a *Penrose Diagram* illustration).

- The Rindler coordinates ρ and η are obvisouly in some sense hyperbolic (Lorentzian) analogues of polar coordinates ($x = r\cos\phi, y = r\sin\phi, ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\phi^2$). In particular, since

$$(\xi^1)^2 - (\xi^0)^2 = \rho^2 \quad , \quad \frac{\xi^0}{\xi^1} = \tanh\eta \ , \tag{1.71}$$

  by construction the lines of constant ρ, $\rho = \rho_0$, are hyperbolas, $(\xi^1)^2 - (\xi^0)^2 = \rho_0^2$, while the lines of constant $\eta = \eta_0$ are straight lines through the origin, $\xi^0 = (\tanh\eta_0)\xi^1$.

- The null lines $\xi^0 = \pm\xi^1$ correspond to $\eta = \pm\infty$. Thus the Rindler coordinates cover the first quadrant $\xi^1 > |\xi^0|$ of Minkowski space and can be used as coordinates there.

- The metric in these new coordinates is time-independent, where time means η, and time-independent means that the coefficients of the metric or line-element in (1.70) do not depend on η. This is due to the fact that the generator ∂_η of η-"time evolution" is actually the generator of a Lorentz boost in the (ξ^0, ξ^1)-plane in Minkowski space,

$$\partial_\eta = (\partial_\eta\xi^0)\partial_{\xi^0} + (\partial_\eta\xi^1)\partial_{\xi^1} = \xi^1\partial_{\xi^0} + \xi^0\partial_{\xi^1} \ . \tag{1.72}$$

  Since a Lorentz boost leaves the Minkowski metric invariant, the latter has to be invariant under translations in η, i.e. it has to be η-independent, as is indeed the case.

- Along the worldline of an observer with constant ρ one has $d\tau = \rho_0 d\eta$, so that his proper time parametrised path is

$$\xi^0(\tau) = \rho_0 \sinh\tau/\rho_0 \qquad \xi^1(\tau) = \rho_0\cosh\tau/\rho_0 \ , \tag{1.73}$$

  and his 4-velocity is given by

$$u^0 = \frac{d}{d\tau}\xi^0(\tau) = \cosh\tau/\rho_0 \qquad u^1 = \frac{d}{d\tau}\xi^1(\tau) = \sinh\tau/\rho_0 \ . \tag{1.74}$$

  These satisfy $-(u^0)^2 + (u^1)^2 = -1$ (as they should), and comparison with (1.66,1.67) shows that the observer's (constant) acceleration is $\mathsf{a} = 1/\rho_0$.

Even though (1.70) is just the metric of Minkowski space-time, written in accelerated coordinates, this metric exhibits a number of interesting features that are prototypical of more general metrics that one encounters in general relativity:

1. First of all, we notice that the coefficients of the line element (metric) in (1.70) are no longer constant (space-time independent). Since in the case of constant acceleration we are just describing a "fake" gravitational field, this dependence on the coordinates is such that it can be completely and globally eliminated by passing to appropriate new coordinates (namely inertial Minkowski coordinates). Since, by the equivalence principle, locally an observer cannot distinguish between a fake and a "true" gravitational field, this now suggests that a "true" gravitational field can be described in terms of a space-time coordinate dependent line-element

$$ds^2 = g_{\alpha\beta}(x)dx^\alpha dx^\beta \quad , \tag{1.75}$$

where the coordinate dependence on the x^α is now such that it cannot be eliminated globally by a suitable choice of coordinates.

2. We observe that (1.70) appears to be ill-defined at $\rho = 0$. However, in this case we already know that this is a mere *coordinate singularity* at $\rho = 0$ (akin to the coordinate singularity at the origin of standard polar coordinates in the Cartesian plane). More generally, whenever a metric written in some coordinate system appears to exhibit some singular behaviour, one needs to investigate whether this is just a coordinate singularity or a true singularity of the gravitational field itself.

3. The above coordinates do not just fail at $\rho = 0$, they actually fail to cover large parts of Minkowski space. Thus the next lesson is that, given a metric in some coordinate system, one has to investigate if the space-time described in this way needs to be extended beyond the range of the original coordinates. One way to analyse this question (which we will make extensive use of in sections 25 and 26 when trying to understand and come to terms with black holes) is to study light rays or the worldlines of freely falling (inertial) observers.

   In the present case, an example of an inertial observer is a static observer in Minkowski space, i.e. an observer at a fixed value of ξ^1, say, with $\xi^0 = \tau$ his proper time. In Rindler coordinates this is described by the condition that $\xi^1 = \rho\cosh\eta$ is a constant, so this is most certainly not a straight line in an (η, ρ)-diagram.

   Such an observer will of course "discover" that $\eta = +\infty$ is not the end of the world (indeed, he crosses this line at finite proper time $\tau = \xi^1$) and that Minkowski space continues (at the very least) into the quadrant $\xi^0 > |\xi^1|$ (see Figure 7 for an illustration of this).

4. Related to this is the behaviour of lightcones when expressed in terms of the coordinates (η, ρ) or when drawn in the (η, ρ)-plane (do this!). These lightcones satisfy $ds^2 = 0$, i.e.

$$\rho^2 d\eta^2 = d\rho^2 \quad \Rightarrow \quad d\eta = \pm\rho^{-1}d\rho \quad . \tag{1.76}$$

describing outgoing (ρ grows with η) respectively ingoing (ρ decreases with increasing η) light rays. These lightcones have the familiar Minkowskian shape at $\rho = 1$, but the lightcones open up for $\rho > 1$ and become more and more narrow for $\rho \to 0$, once again exactly as we will find for the Schwarzschild black hole metric (see Figure 16 in section 25).

5. It follows from (1.72) that the Minkowski norm of ∂_η is

$$|\partial_\eta|^2 = (\xi^0)^2 - (\xi^1)^2 \quad . \tag{1.77}$$

   Thus this generator of Rindler time-translations really is timelike in the region $\xi^1 > |\xi^0|$ covered by the Rindler coordinates, but it actually becomes lightlike on the lightlike boundary $\xi^1 = |\xi^0|$ of that region. As we will discuss in section 26.9, such a *Killing horizon* also happens to be one of the characteristic properties of a black hole.

6. Finally we note that there is a large region of Minkowski space that is "invisible" to the constantly accelerated observers. While a static observer will eventually receive information from any event anywhere in space-time (his past lightcone will eventually cover all of Minkowski space ...), the past lightcone of one of the Rindler accelerated observers (whose worldlines asymptote to the lightcone direction $\xi^0 = \xi^1$) will asymptotically only cover one half of Minkowski space, namely the region $\xi^0 < \xi^1$. Thus any event above the line $\xi^0 = \xi^1$ will forever be invisible to this class of observers. Such an observer-dependent horizon has some similarities with the *event horizon* characterising a black hole (see section 26.4 for a first encounter with such an object, and section 31 for a detailed discussion).