IFT

An Introduction to

GENERAL RELATIVITY

R. Aldrovandi and J. G. Pereira

A Preliminary Note

These notes are intended for a two-month, graduate-level course. Addressed to future researchers in a Centre mainly devoted to Field Theory, they avoid the *ex cathedra* style frequently assumed by teachers of the subject. Mainly, General Relativity is not presented as a finished theory.

Emphasis is laid on the basic tenets and on comparison of gravitation with the other fundamental interactions of Nature. Thus, a little more space than would be expected in such a short text is devoted to the equivalence principle.

The equivalence principle leads to universality, a distinguishing feature of the gravitational field. The other fundamental interactions of Nature—the electromagnetic, the weak and the strong interactions, which are described in terms of gauge theories—are not universal.

These notes, are intended as a short guide to the main aspects of the subject. The reader is urged to refer to the basic texts we have used, each one excellent in its own approach:

- L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, Oxford, 1971)

- C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (Freeman, New York, 1973)

- S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972)

- R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago, 1984)

- J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960)

Contents

Chapter 1

Introduction

1.1   General Concepts

§ **1.1** All elementary particles feel gravitation the same. More specifically, particles with different masses experience a different gravitational force, but in such a way that all of them acquire the same acceleration and, given the same initial conditions, follow the same path. Such *universality* of response is the most fundamental characteristic of the gravitational interaction. It is a unique property, peculiar to gravitation: no other basic interaction of Nature has it.

Due to universality, the gravitational interaction admits a description which makes no use of the concept of *force*. In this description, instead of acting through a force, the presence of a gravitational field is represented by a deformation of the spacetime structure. This deformation, however, preserves the pseudo-riemannian character of the *flat* Minkowski spacetime of Special Relativity, the non-deformed spacetime that represents absence of gravitation. In other words, the presence of a gravitational field is supposed to produce *curvature*, but no other kind of spacetime deformation.

A free particle in flat space follows a straight line, that is, a curve keeping a constant direction. A geodesic is a curve keeping a constant direction on a curved space. As the only effect of the gravitational interaction is to bend spacetime so as to endow it with curvature, a particle submitted exclusively to gravity will follow a geodesic of the deformed spacetime.

This is the approach of Einstein's General Relativity, according to which the gravitational interaction is described by a *geometrization* of spacetime. It is important to remark that only an interaction presenting the property of universality can be described by such a geometrization.

1.2 Some Basic Notions

§ **1.2** Before going further, let us recall some general notions taken from classical physics. They will need refinements later on, but are here put in a language loose enough to make them valid both in the relativistic and the non-relativistic cases.

Frame: a reference frame is a coordinate system for space positions, to which a clock is bound.

Inertia: a reference frame such that free (unsubmitted to any forces) motion takes place with constant velocity is an *inertial frame*; in classical physics, the force law in an inertial frame is $m\frac{dv^k}{dt} = F^k$; in Special Relativity, the force law in an inertial frame is

$$m \frac{d}{ds} U^a = F^a, \tag{1.1}$$

where U is the four-velocity $U = (\gamma, \gamma\mathbf{v}/c)$, with $\gamma = 1/\sqrt{1 - v^2/c^2}$ (as U is dimensionless, F above has not the mechanical dimension of a force — only Fc^2 has). Incidentally, we are stuck to cartesian coordinates to discuss accelerations: the second time derivative of a coordinate is an acceleration only if that coordinate is cartesian.

Transitivity: a reference frame moving with constant velocity with respect to an inertial frame is also an inertial frame;

Relativity: all the laws of nature are the same in all inertial frames; or, alternatively, the equations describing them are invariant under the transformations (of space coordinates and time) taking one inertial frame into the other; or still, the equations describing the laws of Nature in terms of space coordinates and time keep their forms in different inertial frames; this "principle" can be seen as an experimental fact; in non-relativistic classical physics, the transformations referred to belong to the Galilei group; in Special Relativity, to the Poincaré group.

Causality: in non-relativistic classical physics the interactions are given by the potential energy, which usually depends only on the space coordinates; forces on a given particle, caused by all the others, depend only on their position at a given instant; a change in position changes the force instantaneously; this instantaneous propagation effect — or action at a distance — is a typicallly classical, non-relativistic feature; it violates special-relativistic causality; Special Relativity takes into account the experimental fact that light has a finite velocity in vacuum and says that no effect can propagate faster than that velocity.

Fields: there have been tentatives to preserve action at a distance in a relativistic context, but a simpler way to consider interactions while respecting Special Relativity is of common use in field theory: interactions are mediated by a field, which has a well-defined behaviour under transformations; disturbances propagate, as said above, with finite velocities.

1.3   The Equivalence Principle

Equivalence is a guiding principle, which inspired Einstein in his construction of General Relativity. It is firmly rooted on experience.[*]

In its most usual form, the Principle includes three sub–principles: the weak, the strong and that which is called "Einstein's equivalence principle". We shall come back and forth to them along these notes. Let us shortly list them with a few comments.

§ **1.3** The **weak equivalence principle**: universality of free fall, or inertial mass = gravitational mass.

> In a gravitational field, all pointlike structureless particles follow one same path; that path is fixed once given (i) an initial position $x(t_0)$ and (ii) the correspondent velocity $\dot{x}(t_0)$.

This leads to a force equation which is a second order ordinary differential equation. No characteristic of any special particle, no particular property

---

[*] Those interested in the experimental status will find a recent appraisal in C. M. Will, *The Confrontation between General Relativity and Experiment*, arXiv:gr-qc/0103036 12 Mar 2001. Theoretical issues are discussed by B. Mashhoon, *Measurement Theory and General Relativity*, gr-qc/0003014, and *Relativity and Nonlocality*, gr-qc/0011013 v2.

appears in the equation. Gravitation is consequently universal. Being universal, it can be seen as a property of space itself. It determines geometrical properties which are common to all particles. The weak equivalence principle goes back to Galileo. It raises to the status of fundamental principle a deep experimental fact: the equality of inertial and gravitational masses of all bodies.

The **strong equivalence principle**: (Einstein's lift) says that

Gravitation can be made to vanish locally through an appropriate choice of frame.

It requires that, for any and every particle and at each point x_0, there exists a frame in which $\ddot{x}^\mu = 0$.

Einstein's equivalence principle requires, besides the weak principle, the local validity of Poincaré invariance — that is, of Special Relativity. This invariance is, in Minkowski space, summed up in the Lorentz metric. The requirement suggests that the above deformation caused by gravitation is a change in that metric.

In its complete form, the equivalence principle

1. provides an operational definition of the gravitational interaction;

2. geometrizes it;

3. fixes the equation of motion of the test particles.

§ **1.4** Use has been made above of some undefined concepts, such as "path", and "local". A more precise formulation requires more mathematics, and will be left to later sections. We shall, for example, rephrase the Principle as a prescription saying how an expression valid in Special Relativity is changed once in the presence of a gravitational field. What changes is the notion of derivative, and that change requires the concept of connection. The prescription (of "minimal coupling") will be seen after that notion is introduced.

4

§ **1.5** Now, forces equally felt by all bodies were known since long. They are the inertial forces, whose name comes from their turning up in non-inertial frames. Examples on Earth (not an inertial system !) are the centrifugal force and the Coriolis force. We shall begin by recalling what such forces are in Classical Mechanics, in particular how they appear related to changes of coordinates. We shall then show how a metric appears in an non-inertial frame, and how that metric changes the law of force in a very special way.

1.3.1 Inertial Forces

§ **1.6** In a frame attached to Earth (that is, rotating with a certain angular velocity $\boldsymbol{\omega}$), a body of mass m moving with velocity $\dot{\mathbf{X}}$ on which an external force $\mathbf{F}_{ext}$ acts will actually experience a "strange" total force. Let us recall in rough brushstrokes how that happens.

A simplified model for the motion of a particle in a system attached to Earth is taken from the classical formalism of rigid body motion.[†] It runs as follows:

The rotating Earth

Start with an inertial cartesian system, the *space system* ("inertial" means — we insist — devoid of proper acceleration). A point particle will have coordinates $\{x^i\}$, collectively written as a column vector $\mathbf{x} = (x^i)$. Under the action of a force $\mathbf{f}$, its velocity and acceleration will be, with respect to that system, $\dot{\mathbf{x}}$ and $\ddot{\mathbf{x}}$. If the particle has mass m, the force will be $\mathbf{f} = m\,\ddot{\mathbf{x}}$.

Consider now another coordinate system (the *body system*) which rotates around the origin of the first. The point particle will have coordinates $\mathbf{X}$ in this system. The relation between the coordinates will be given by a rotation matrix R,

$$\mathbf{X} = R\,\mathbf{x}.$$

The forces acting on the particle in both systems are related by the same

---

[†] The standard approach is given in H. Goldstein,*Classical Mechanics*, Addison–Wesley, Reading, Mass., 1982. A modern description can be found in J. L. McCauley,*Classical Mechanics*, Cambridge University Press, Cambridge, 1997.

5

relation,

$$\mathbf{F} = R\,\mathbf{f}.$$

We are using symbols with capitals ($\mathbf{X}$, $\mathbf{F}$, Ω, ...) for quantities referred to the body system, and the corresponding small letters ($\mathbf{x}$, $\mathbf{f}$, ω, ...) for the same quantities as "seen from" the space system.

Now comes the crucial point: as Earth is rotating with respect to the space system, a different rotation is necessary at each time to pass from that system to the body system; this is to say that the rotation matrix R is time-dependent. In consequence, the velocity and the acceleration seen from Earth's system are given by

$$\dot{\mathbf{X}} = \dot{R}\,\mathbf{x} + R\,\dot{\mathbf{x}}$$

$$\ddot{\mathbf{X}} = \ddot{R}\,\mathbf{x} + 2\dot{R}\,\dot{\mathbf{x}} + R\,\ddot{\mathbf{x}}. \tag{1.2}$$

Introduce the matrix $\omega = -\,R^{-1}\dot{R}$. It is an antisymmetric 3×3 matrix, consequently equivalent to a vector. That vector, with components

$$\omega^k = \tfrac{1}{2}\,\epsilon^k{}_{ij}\,\omega^{ij} \tag{1.3}$$

(which is the same as $\omega^{ij} = \epsilon^{ijk}\omega_k$), is Earth's angular velocity seen from the space system. ω is, thus, a matrix version of the angular velocity. It will correspond, in the body system, to

$$\Omega = R\omega R^{-1} = -\,\dot{R}\,R^{-1}.$$

Comment 1.1 Just in case, ϵ_{ijk} is the 3-dimensional Kronecker symbol in 3-dimensional space: $\epsilon_{123} = 1$; any odd exchange of indices changes the sign; $\epsilon_{ijk} = 0$ if there are repeated indices. Indices are raised and lowered with the Kronecker delta δ_{ij}, defined by $\delta_{ii} = 1$ and $\delta_{ij} = 0$ if $i \neq j$. In consequence, $\epsilon^{ijk} = \epsilon_{ijk} = \epsilon^i{}_{jk}$, etc. The usual vector product has components given by $(\mathbf{v} \times \mathbf{u})_i = (\mathbf{v} \wedge \mathbf{u})_i = \epsilon_{ijk}u_j v_k$. An antisymmetric matrix like ω, acting on a vector will give $\omega^{ij}v_j = \epsilon^{ijk}\omega_k v_j = (\boldsymbol{\omega} \times \mathbf{v})_i$.

A few relations turn out without much ado: $\Omega^2 = R\omega^2 R^{-1}$, $\dot{\Omega} = R\dot{\omega}R^{-1}$ and

$$\dot{\omega} - \omega\,\omega = -\,R^{-1}\ddot{R}\,,$$

or
$$\ddot{R} = R\,[\dot{\omega} - \omega\,\omega]\ .$$

Substitutions put then Eq. (1.2) into the form

$$\ddot{\mathbf{X}} + 2\,\Omega\,\dot{\mathbf{X}} + [\dot{\Omega} + \Omega^2]\,\mathbf{X} = R\,\ddot{\mathbf{x}}$$

The above relationship between 3×3 matrices and vectors takes matrix action on vectors into vector products: $\omega\,\mathbf{x} = \boldsymbol{\omega} \times \mathbf{x}$, etc. Transcribing into vector products and multiplying by the mass, the above equation acquires its standard form in terms of forces,

$$m\,\ddot{\mathbf{X}} = \underbrace{-\,m\,\boldsymbol{\Omega} \times \boldsymbol{\Omega} \times \mathbf{X}}_{centrifugal}\ \underbrace{-\,2m\,\boldsymbol{\Omega} \times \dot{\mathbf{X}}}_{Coriolis}\ \underbrace{-\,m\,\dot{\boldsymbol{\Omega}} \times \mathbf{X}}_{fluctuation}\ +\,\mathbf{F}_{ext}\ .$$

We have indicated the usual names of the contributions. A few words on each of them

fluctuation force: in most cases can be neglected for Earth, whose angular velocity is very nearly constant.

centrifugal force: opposite to Earth's attraction, it is already taken into account by any balance (you are fatter than you think, your mass is larger than suggested by your your weight by a few grams ! the ratio is 3/1000 at the equator).

Coriolis force: responsible for trade winds, rivers' one-sided overflows, asymmetric wear of rails by trains, and the effect shown by the Foucault pendulum.

§ **1.7** Inertial forces have once been called "ficticious", because they disappear when seen from an inertial system at rest. We have met them when we started from such a frame and transformed to coordinates attached to Earth. We have listed the measurable effects to emphasize that they are actually very real forces, though frame-dependent.

§ **1.8** The remarkable fact is that *each body feels them the same*. Think of the examples given for the Coriolis force: air, water and iron feel them, and

in the same way. Inertial forces are "universal", just like gravitation. This has led Einstein to his formidable stroke of genius, to conceive gravitation as an inertial force.

§ **1.9** Nevertheless, if gravitation were an inertial effect, it should be obtained by changing to a non-inertial frame. And here comes a problem. In Classical Mechanics, time is a parameter, external to the coordinate system. In Special Relativity, with Minkowski's invention of spacetime, time underwent a violent conceptual change: no more a parameter, it became the fourth coordinate (in our notation, the zeroth one).

Classical non-inertial frames are obtained from inertial frames by transformations which depend on time. Relativistic non-inertial frames should be obtained by transformations which depend on spacetime. Time–dependent coordinate changes ought to be special cases of more general transformations, dependent on all the spacetime coordinates. In order to be put into a position closer to inertial forces, and concomitantly respect Special Relativity, gravitation should be related to the dependence of frames on all the coordinates.

§ **1.10** Universality of inertial forces has been the first hint towards General Relativity. A second ingredient is the notion of field. The concept allows the best approach to interactions coherent with Special Relativity. All known forces are mediated by fields on spacetime. Now, if gravitation is to be represented by a field, it should, by the considerations above, be a universal field, equally felt by every particle. It should change spacetime itself. And, of all the fields present in a space the metric — the first fundamental form, as it is also called — seemed to be the basic one. The simplest way to change spacetime would be to change its metric. Furthermore, the metric does change when looked at from a non-inertial frame.

§ **1.11** The Lorentz metric η of Special Relativity is rather trivial. There is a coordinate system (the cartesian system) in which the line element of Minkowski space takes the form

$$ds^2 = \eta_{ab}dx^a dx^b = dx^0 dx^0 - dx^1 dx^1 - dx^2 dx^2 - dx^3 dx^3$$

Lorentz metric

8

$$= c^2 dt^2 - dx^2 - dy^2 - dz^2 \ . \tag{1.4}$$

Take two points P and Q in Minkowski spacetime, and consider the integral

$$\int_P^Q ds = \int_P^Q \sqrt{\eta_{ab} dx^a dx^b}.$$

Its value depends on the path chosen. In consequence, it is actually a functional on the space of paths between P and Q,

$$S[\gamma_{PQ}] = \int_{\gamma_{PQ}} ds. \tag{1.5}$$

An extremal of this functional would be a curve γ such that $\delta S[\gamma] = \int \delta ds = 0$. Now,

$$\delta ds^2 = 2 \ ds \ \delta ds = 2 \ \eta_{ab} \ dx^a \delta dx^b,$$

so that

$$\delta ds = \eta_{ab} \frac{dx^a}{ds} \ \delta dx^b = \eta_{ab} \ U^a \ \delta dx^b \ .$$

Thus, commuting d and δ and integrating by parts,

$$\delta S[\gamma] = \int_P^Q \eta_{ab} \frac{dx^a}{ds} \frac{d\delta x^b}{ds} \ ds = - \int_P^Q \eta_{ab} \frac{d}{ds} \frac{dx^a}{ds} \ \delta x^b \ ds$$

$$= - \int_P^Q \eta_{ab} \frac{d}{ds} U^a \ \delta x^b \ ds.$$

The variations δx^b are arbitrary. If we want to have $\delta S[\gamma] = 0$, the integrand must vanish. Thus, an extremal of $S[\gamma]$ will satisfy

$$\frac{d}{ds} U^a = 0. \tag{1.6}$$

This is the equation of a straight line, the force law (1.1) when $F^a = 0$. The solution of this differential equation is fixed once initial conditions are given. We learn here that a vanishing acceleration is related to an extremal of $S[\gamma_{PQ}]$.

§ **1.12** Let us see through an example what happens when a force is present. For that it is better to notice beforehand that, when considering fields, it is

in general the action which is extremal. Simple dimensional analysis shows that, in order to have a real physical action, we must take

$$S = - mc \int ds \qquad (1.7)$$

instead of the "length". Consider the case of a charged test particle. The coupling of a particle of charge e to an electromagnetic potential A is given by $A_a j^a = e\, A_a U^a$, so that the action along a curve is

$$S_{em}[\gamma] = - \frac{e}{c} \int_\gamma A_a U^a ds = - \frac{e}{c} \int_\gamma A_a dx^a.$$

The variation is

$$\delta S_{em}[\gamma] = - \frac{e}{c} \int_\gamma \delta A_a dx^a - \frac{e}{c} \int_\gamma A_a d\delta x^a = - \frac{e}{c} \int_\gamma \delta A_a dx^a + \frac{e}{c} \int_\gamma dA_b \delta x^b$$

$$= - \frac{e}{c} \int_\gamma \partial_b A_a \delta x^b dx^a + \frac{e}{c} \int_\gamma \partial_a A_b \delta x^b dx^a = - \frac{e}{c} \int_\gamma [\partial_b A_a - \partial_a A_b] \delta x^b \frac{dx^a}{ds} ds$$

$$= - \frac{e}{c} \int_\gamma F_{ba} U^a \delta x^b ds \ .$$

Combining the two pieces, the variation of the total action

$$S = -mc \int_P^Q ds - \frac{e}{c} \int_P^Q A_a dx^a \qquad (1.8)$$

is

$$\delta S = \int_P^Q \left[\eta_{ab}\, mc\, \frac{d}{ds}\, U^a - \frac{e}{c}\, F_{ba} U^a \right]\, \delta x^b ds.$$

The extremal satisfies

Lorentz force law

$$mc\, \frac{d}{ds}\, U^a = \frac{e}{c}\, F^a{}_b\, U^b, \qquad (1.9)$$

which is the Lorentz force law and has the form of the general case (1.1).

1.3.2   The Wake of Non-Trivial Metric

Let us see now — in another example — that the metric changes when viewed from a non-inertial system. This fact suggests that, if gravitation is to be related to non-inertial systems, a gravitational field is to be related to a non-trivial metric.

§ 1.13 Consider a rotating disc (details can be seen in Møller's book[‡]), seen as a system performing a uniform rotation with angular velocity ω on the x, y plane:

$$x = r\cos(\theta + \omega t) ; \quad y = r\sin(\theta + \omega t) ; \quad Z = z ;$$

$$X = R\cos\theta ; \quad Y = R\sin\theta.$$

This is the same as

$$x = X\cos\omega t - Y\sin\omega t ; \quad y = Y\cos\omega t + X\sin\omega t .$$

As there is no contraction along the radius (the motion being orthogonal to it), $R = r$. Both systems coincide at $t = 0$. Now, given the standard Minkowski line element

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

in cartesian ("space", inertial) coordinates $(x^0, x^1, x^2, x^3) = (ct, x, y, z)$, how will a "body" observer on the disk see it ?

It is immediate that

$$dx = dr\cos(\theta + \omega t) - r\sin(\theta + \omega t)[d\theta + \omega dt]$$

$$dy = dr\sin(\theta + \omega t) + r\cos(\theta + \omega t)[d\theta + \omega dt]$$

$$dx^2 = dr^2\cos^2(\theta + \omega t) + r^2\sin^2(\theta + \omega t)[d\theta + \omega dt]^2$$

$$-2rdr\cos(\theta + \omega t)\sin(\theta + \omega t)[d\theta + \omega dt] ;$$

$$dy^2 = dr^2\sin^2(\theta + \omega t) + r^2\cos^2(\theta + \omega t)[d\theta + \omega dt]^2$$

$$+2rdr\sin(\theta + \omega t)\cos(\theta + \omega t)[d\theta + \omega dt]$$

$$\therefore dx^2 + dy^2 = dR^2 + R^2(d\theta^2 + \omega^2 dt^2 + 2\omega d\theta dt).$$

It follows from

$$dX^2 + dY^2 = dR^2 + R^2 d\theta^2,$$

that

$$dx^2 + dy^2 = dX^2 + dY^2 + R^2\omega^2 dt^2 + 2\omega R^2 d\theta dt.$$

[‡] C. Møller, *The Theory of Relativity*, Oxford at Clarendon Press, Oxford, 1966, mainly in §8.9.

A simple check shows that

$$XdY - YdX = R^2 d\theta,$$

so that

$$dx^2 + dy^2 = dX^2 + dY^2 + R^2\omega^2 dt^2 + 2\omega X dY dt - 2\omega Y dX dt.$$

Thus,

$$ds^2 = (1 - \frac{\omega^2 R^2}{c^2}) \, c^2 dt^2 - dX^2 - dY^2 - 2\omega X dY dt + 2\omega Y dX dt - dZ^2 \ .$$

In the moving body system, with coordinates $(X^0, X^1, X^2, X^3) = (ct, X, Y, Z = z)$ the metric will be

$$ds^2 = g_{\mu\nu} dX^\mu dX^\nu,$$

where the only non-vanishing components of the modified metric g are:

$$g_{11} = g_{22} = g_{33} = -1; \quad g_{01} = g_{10} = \omega Y/c; \quad g_{02} = g_{20} = -\omega X/c;$$

$$g_{00} = 1 - \frac{\omega^2 R^2}{c^2} \ .$$

This is better visualized as the matrix

$$g = (g_{\mu\nu}) = \begin{pmatrix} 1 - \frac{\omega^2 R^2}{c^2} & \omega Y/c & -\omega X/c & 0 \\ \omega Y/c & -1 & 0 & 0 \\ -\omega X/c & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \qquad (1.10)$$

We can go one step further an define the body time coordinate T to be such that $dT = \sqrt{1 - \omega^2 R^2/c^2} \, dt$, that is,

$$T = \sqrt{1 - \omega^2 R^2/c^2} \ t \ .$$

This expression is physically appealing, as it is the same as $T = \sqrt{1 - v^2/c^2} \, t$, the time-contraction of Special Relativity, if we take into account the fact that a point with coordinates (R, θ) will have squared velocity $v^2 = \omega^2 R^2$. We see that, anyhow, the body coordinate system can be used only for points

satisfying the condition $\omega R < c$. In the body coordinates (cT, X, Y, Z), the line element becomes

$$ds^2 = c^2 dT^2 - dX^2 - dY^2 - dZ^2 + 2\omega[YdX - XdY]\,\frac{dT}{\sqrt{1 - \omega^2 R^2/c^2}}\;.$$
(1.11)

Time, as measured by the accelerated frame, differs from that measured in the inertial frame. And, anyhow, the metric has changed. This is the point we wanted to make: when we change to a non-inertial system the metric undergoes a significant transformation, even in Special Relativity.

Comment 1.2 Put $\beta = \omega R/c$. Matrix (1.10) and its inverse are

$$g = (g_{\mu\nu}) = \begin{pmatrix} 1-\beta^2 & \frac{\beta Y}{R} & -\frac{\beta X}{R} & 0 \\ \frac{\beta Y}{R} & -1 & 0 & 0 \\ -\frac{\beta X}{R} & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \;;\; g^{-1} = (g^{\mu\nu}) = \begin{pmatrix} 1 & \frac{\beta Y}{R} & -\frac{\beta X}{R} & 0 \\ \frac{\beta Y}{R} & \beta^2\frac{Y^2}{R^2}-1 & -\frac{\beta^2 XY}{R^2} & 0 \\ -\frac{\beta X}{R} & -\frac{\beta^2 XY}{R^2} & \beta^2\frac{X^2}{R^2}-1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \;.$$

1.3.3 Towards Geometry

§ 1.14 We have said that the only effect of a gravitational field is to bend spacetime, so that straight lines become geodesics. Now, there are two quite distinct definitions of a straight line, which coincide on flat spaces but not on spaces endowed with more sophisticated geometries. A straight line going from a point P to a point Q is

1. among all the lines linking P to Q, that with the shortest length;

2. among all the lines linking P to Q, that which keeps the same direction all along.

There is a clear problem with the first definition: length presupposes a metric — a real, positive-definite metric. The Lorentz metric does not define lengths, but pseudo-lengths. There is always a "zero-length" path between any two points in Minkowski space. In Minkowski space, $\int ds$ is actually *maximal* for a straight line. Curved lines, or broken ones, give a smaller pseudo-length. We have introduced a minus sign in Eq.(1.7) in order to conform to the current notion of "minimal action".

The second definition can be carried over to spacetime of any kind, but at a price. Keeping the same direction means "keeping the tangent velocity

vector constant". The derivative of that vector along the line should vanish. Now, derivatives of vectors on non-flat spaces require an extra concept, that of connection — which, will, anyhow, turn up when the first definition is used. We shall consequently feel forced to talk a lot about connections in what follows.

§ **1.15** Consider an arbitrary metric g, defining the interval by

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu.$$

What happens now to the integral of Eq.(1.7) with a point-dependent metric? Consider again a charged test particle, but now in the presence of a non-trivial metric. We shall retrace the steps leading to the Lorentz force law, with the action

$$S = -mc \int_{\gamma_{PQ}} ds - \frac{e}{c} \int_{\gamma_{PQ}} A_\mu dx^\mu, \tag{1.12}$$

but now with $ds = \sqrt{g_{\mu\nu}dx^\mu dx^\nu}$.

1. Take first the variation

$$\delta ds^2 = 2ds\delta ds = \delta[g_{\mu\nu}dx^\mu dx^\nu] = dx^\mu dx^\nu \delta g_{\mu\nu} + 2g_{\mu\nu}dx^\mu\delta dx^\nu$$

$$\therefore \delta ds = \tfrac{1}{2}\, \tfrac{dx^\mu}{ds}\tfrac{dx^\nu}{ds}\, \partial_\lambda g_{\mu\nu}\delta x^\lambda ds + g_{\mu\nu}\tfrac{dx^\mu}{ds}\tfrac{\delta dx^\nu}{ds}ds$$

We have conveniently divided and multiplied by ds.

2. We now insert this in the first piece of the action and integrate by parts the last term, getting

$$\delta S = -mc \int_{\gamma_{PQ}} \left[\tfrac{1}{2}\, \tfrac{dx^\mu}{ds}\tfrac{dx^\rho}{ds}\partial_\nu g_{\mu\rho} - \tfrac{d}{ds}(g_{\mu\nu}\tfrac{dx^\mu}{ds})\right] \delta x^\nu ds$$

$$- \frac{e}{c} \int_{\gamma_{PQ}} [\delta A_\mu dx^\mu + A_\mu d\delta x^\mu]. \tag{1.13}$$

3. The derivative $\tfrac{d}{ds}(g_{\mu\nu}\tfrac{dx^\mu}{ds})$ is

$$\frac{d}{ds}(g_{\mu\nu}\frac{dx^\mu}{ds}) = \frac{dx^\mu}{ds}\frac{d}{ds}g_{\mu\nu} + g_{\mu\nu}\frac{d}{ds}U^\mu = U^\mu U^\nu \partial_\nu g_{\mu\nu} + g_{\mu\nu}\frac{d}{ds}U^\mu$$

$$= g_{\mu\nu}\frac{d}{ds}U^\mu + U^\mu U^\rho \partial_\rho g_{\mu\nu} = g_{\mu\nu}\frac{d}{ds}U^\mu + \tfrac{1}{2}\, U^\sigma U^\rho[\partial_\rho g_{\sigma\nu} + \partial_\sigma g_{\rho\nu}].$$

14

4. Collecting terms in the metric sector, and integrating by parts in the electromagnetic sector,

$$\delta S = -mc \int_{\gamma_{PQ}} \left[-g_{\mu\nu} \frac{d}{ds} U^{\mu} - \tfrac{1}{2} \, U^{\sigma} U^{\rho} \left(\partial_{\rho} g_{\sigma\nu} + \partial_{\sigma} g_{\rho\nu} - \partial_{\nu} g_{\mu\rho} \right) \right] \delta x^{\nu} ds$$

$$- \frac{e}{c} \int_{\gamma_{PQ}} [\partial_{\nu} A_{\mu} \delta x^{\nu} dx^{\mu} - \delta x^{\nu} \partial_{\mu} A_{\nu} dx^{\mu}] = \tag{1.14}$$

$$-mc \int_{\gamma_{PQ}} g_{\mu\nu} \left[-\frac{d}{ds} U^{\mu} - U^{\sigma} U^{\rho} \left\{ \tfrac{1}{2} \, g^{\mu\lambda} \left(\partial_{\rho} g_{\sigma\lambda} + \partial_{\sigma} g_{\rho\lambda} - \partial_{\lambda} g_{\sigma\rho} \right) \right\} \right] \delta x^{\nu} ds$$

$$- \frac{e}{c} \int_{\gamma_{PQ}} [\partial_{\nu} A_{\mu} \delta x^{\nu} dx^{\mu} - \delta x^{\nu} \partial_{\mu} A_{\nu} dx^{\mu}]. \tag{1.15}$$

5. We meet here an important character of all metric theories. The expression between curly brackets is the *Christoffel symbol*, which will be indicated by the notation $\overset{\circ}{\Gamma}$:

> Christoffel symbol

$$\overset{\circ}{\Gamma}{}^{\mu}{}_{\sigma\rho} = \tfrac{1}{2} \, g^{\mu\lambda} \left(\partial_{\rho} g_{\sigma\lambda} + \partial_{\sigma} g_{\rho\lambda} - \partial_{\lambda} g_{\sigma\rho} \right). \tag{1.16}$$

6. After arranging the terms, we get

$$\delta S =$$
$$\int_{\gamma_{PQ}} \left[mc \, g_{\mu\nu} \left(\frac{d}{ds} U^{\mu} + \overset{\circ}{\Gamma}{}^{\mu}{}_{\sigma\rho} U^{\sigma} U^{\rho} \right) - \frac{e}{c} \, (\partial_{\nu} A_{\rho} - \partial_{\rho} A_{\nu}) U^{\rho} \right] \delta x^{\nu} ds. \tag{1.17}$$

7. The variations δx^{ν}, except at the fixed endpoints, is quite arbitrary. To have $\delta S = 0$, the integrand must vanish. Which gives, after contracting with $g^{\lambda\nu}$,

$$mc \left(\frac{d}{ds} \, U^{\lambda} + \overset{\circ}{\Gamma}{}^{\lambda}{}_{\sigma\rho} U^{\sigma} U^{\rho} \right) = \frac{e}{c} \, F^{\lambda}{}_{\rho} U^{\rho} . \tag{1.18}$$

8. This is the Lorentz law of force in the presence of a non-trivial metric. We see that what appears as acceleration is now

$$\overset{\circ}{A}{}^{\lambda} = \frac{d}{ds} \, U^{\lambda} + \overset{\circ}{\Gamma}{}^{\lambda}{}_{\sigma\rho} U^{\sigma} U^{\rho}. \tag{1.19}$$

15

The Christoffel symbol is a non-tensorial quantity, a *connection.* We shall see later that a reference frame can be always chosen in which it vanishes at a point. The law of force

$$mc \left(\frac{d}{ds} U^\lambda + \overset{\circ}{\Gamma}{}^\lambda{}_{\sigma\rho} U^\sigma U^\rho \right) = F^\lambda \qquad (1.20)$$

will, in that frame and at that point, reduce to that holding for a trivial metric, Eq. (1.1).

9. In the absence of forces, the resulting expression,

<div style="float:right; border:1px solid black; padding:2px;">geodesic
equation</div>

$$\frac{d}{ds} U^\lambda + \overset{\circ}{\Gamma}{}^\lambda{}_{\sigma\rho} U^\sigma U^\rho = 0, \qquad (1.21)$$

is the *geodesic equation,* defining the "straightest" possible line on a space in which the metric is non-trivial.

Comment 1.3 *An accelerated frame creates the illusion of a force.* Suppose a point P is "at rest". It may represent a vessel in space, far from any other body. An astronaut in the spacecraft can use gyros and accelerometers to check its state of motion. It will never be able to say that it is actually at rest, only that it has some constant velocity. Its own reference frame will be inertial. Assume another craft approaches at a velocity which is constant relative to P, and observes P. It will measure the distance from P, see that the velocity $\dot{\mathbf{x}}$ is constant. That observer will also be inertial.

Suppose now that the second vessel accelerates towards P. It will then see $\ddot{\mathbf{x}} \neq 0$, and will interpret this result in the normal way: there is a force pulling P. That force is clearly an illusion: it would have opposite sign if the accelerated observer moved away from P. No force acts on P, the force is due to the observer's own acceleration. It comes from the observer, not from P.

Comment 1.4 *Curvature creates the illusion of a force.* Two old travellers (say, Herodotus and Pausanias) move northwards on Earth, starting from two distinct points on the equator. Suppose they somehow communicate, and have a means to evaluate their relative distance. They will notice that that distance decreases with their progress until, near the pole, they will see it dwindle to nothing. Suppose further they have ancient notions, and think the Earth is flat. How would they explain it ? They would think there were some force, some attractive force between them. And what is the real explanation ? It is simply that Earth's surface is a curved space. The force is an illusion, born from the flatness prejudice.

16

§ **1.16** Gravitation is very weak. To present time, no gravitational bending in the trajectory of an *elementary* particle has been experimentally observed. Only large agglomerates of fermions have been seen to experience it. Nevertheless, an effect on the phase of the wave-function has been detected, both for neutrons and atoms.[§]

§ **1.17** Suppose that, of all elementary particles, one single existed which did not feel gravitation. That would be enough to change all the picture. The underlying spacetime would remain Minkowski's, and the metric responsible for gravitation would be a field $g_{\mu\nu}$ on that, by itself flat, spacetime.

Spacetime is a geometric construct. Gravitation should change the geometry of spacetime. This comes from what has been said above: coordinates, metric, connection and frames are part of the differential-geometrical structure of spacetime. We shall need to examine that structure. The next chapter is devoted to the main aspects of differential geometry.

[§] The so-called "COW experiment" with neutrons is described in R. Colella, A. W. Overhauser and S. A. Werner, Phys. Rev. Lett. **34**, 1472 (1975). See also U. Bonse and T. Wroblevski, Phys. Rev. Lett. **51**, 1401 (1983). Experiments with atoms are reviewed in C. J. Bordé, *Matter wave interferometers: a synthetic approach*, and in B. Young, M. Kasevich and S. Chu, *Precision atom interferometry with light pulses*, in *Atom Interferometry*, P. R. Bergman (editor) (Academic Press, San Diego, 1997).

Chapter 2

Geometry

The basic equations of Physics are differential equations. Now, not every space accepts differentials and derivatives. Every time a derivative is written in some space, a lot of underlying structure is assumed, taken for granted. It is supposed that that space is a differentiable (or smooth) manifold. We shall give in what follows a short survey of the steps leading to that concept. That will include many other notions taken for granted, as that of "coordinate", "parameter", "curve", "continuous", and the very idea of space.

2.1 Differential Geometry

Physicists work with sets of numbers, provided by experiments, which they must somehow organize. They make – always implicitly – a large number of assumptions when conceiving and preparing their experiments and a few more when interpreting them. For example, they suppose that the use of coordinates is justified: every time they have to face a continuum set of values, it is through coordinates that they distinguish two points from each other. Now, not every kind of point-set accept coordinates. Those which do accept coordinates are specifically structured sets called *manifolds*. Roughly speaking, manifolds are sets on which, at least around each point, everything looks usual, that is, *looks Euclidean*.

§ **2.1** Let us recall that a *distance function* is a function d taking any pair (p, q) of points of a set X into the real line $\mathbb{R}$ and satisfying the following four conditions : (i) $d(p, q) \geq 0$ for all pairs (p, q); (ii) $d(p, q) = 0$ if and only if $p = q$; (iii) $d(p, q) = d(q, p)$ for all pairs (p, q); (iv) $d(p, r) + d(r, q) \geq d(p, q)$ for any three points p, q, r. It is thus a mapping d: $X \times X \to \mathbb{R}_+$. A space on

distance
function

which a distance function is defined is a *metric space.* For historical reasons, a distance function is here (and frequently) called a metric, though it would be better to separate the two concepts (see in section 2.1.4, page 40, how a positive-definite metric, which is a tensor field, can define a distance).

§ **2.2** The Euclidean spaces are the basic spaces we shall start with. The 3-dimensional space $\mathbb{E}^3$ consists of the set $\mathbb{R}^3$ of ordered triples of real numbers $p = (p^1, p^2, p^3)$, $q = (q^1, q^2, q^3)$, etc, endowed with the distance function $d(p, q) = \left[\sum_{i=1}^{3}(p^i - q^i)^2\right]^{1/2}$. A r-ball around p is the set of points q such that $d(p, q) < r$, for r a positive number. These open balls define a topology, that is, a family of subsets of $\mathbb{E}^3$ leading to a well-defined concept of continuity. It was thought for much time that a topology was necessarily an offspring of a distance function. This is not true. The modern concept, presented below (§2.7), is more abstract and does without any distance function.

Non-relativistic fields live on space $\mathbb{E}^3$ or, if we prefer, on the direct–product spacetime $\mathbb{E}^3 \otimes \mathbb{E}^1$, with the extra $\mathbb{E}^1$ accounting for time. In non-relativistic physics space and time are independent of each other, and this is encoded in the direct–product character: there is one distance function for space, another for time. In relativistic theories, space and time are blended together in an inseparable way, constituting a real spacetime. The notion of spacetime was introduced by Poincaré and Minkowski in Special Relativity. Minkowski spacetime, to be described later, is the paradigm of every other spacetime.

For the n-dimensional Euclidean space $\mathbb{E}^n$, the point set is the set $\mathbb{R}^n$ of ordered n-uples p $= (p^1, p^2, ..., p^n)$ of real numbers and the topology is the ball-topology of the distance function d(p, q) $= [\sum_{i=1}^{n}(p^i - q^i)^2]^{1/2}$. $\mathbb{E}^n$ is the basic, initially assumed space, as even differential manifolds will be presently defined so as to generalize it. The introduction of coordinates on a general space S will require that S "resemble" some $\mathbb{E}^n$ around each point.

§ **2.3** When we say "around each point", mathematicians say "locally". For example, manifolds are "locally Euclidean" sets. But not every set of points can resemble, even locally, an Euclidean space. In order to do so, a point set must have very special properties. To begin with, it must have a topology. A set with such an underlying structure is a "topological space". Manifolds are

19

topological spaces with some particular properties which make them locally Euclidean spaces.

The procedure then runs as follows:

> it is supposed that we know everything on usual Analysis, that is, Analysis on Euclidean spaces. Structures are then progressively added up to the point at which it becomes possible to transfer notions from the Euclidean to general spaces. This is, as a rule, only possible locally, in a neighborhood around each point.

§ **2.4** We shall later on represent physical systems by fields. Such fields are present somewhere in space and time, which are put together in a unified spacetime. We should say what we mean by that. But there is more. Fields are idealized objects, which we represent mathematically as members of some other spaces. We talk about vectors, matrices, functions, etc. There will be spaces of vectors, of matrices, of functions. And still more: we operate with these fields. We add and multiply them, sometimes integrate them, or take their derivatives. Each one of these operations requires, in order to have a meaning, that the objects they act upon belong to spaces with specific properties.

2.1.1 Spaces

§ **2.5** Thus, first task, it will be necessary to say what we understand by "spaces" in general. Mathematicians have built up a systematic theory of spaces, which describes and classifies them in a progressive order of complexity. This theory uses two primitive notions - sets, and functions from one set to another. The elements belonging to a space may be vectors, matrices, functions, other sets, etc, but the standard language calls simply "points" the members of a generic space.

A *space S* is an organized set of points, a point set plus a structure. This structure is a division of S, a convenient family of subsets. Different purposes require different kinds of subset families. For example, in order to arrive at a well-defined notion of integration, a *measure space* is necessary, which demands a special type of sub-division called "σ-algebra". To make of

general notion

20

S a topological space, we decompose it in another peculiar way. The latter will be our main interest because most spaces used in Physics are, to start with, topological spaces.

§ 2.6 That this is so is not evident at every moment. The customary approach is just the contrary. The physicist will implant the object he needs without asking beforehand about the possibilities of the underlying space. He can do that because Physics is an experimental science. He is justified in introducing an object if he obtains results confirmed by experiment. A well-succeeded experiment brings forth evidence favoring all the assumptions made, explicit or not. Summing up: the additional objects (say, fields) defined on a certain space (say, spacetime) may serve to probe into the underlying structure of that space.

§ 2.7 Topological spaces are, thus, the primary spaces. Let us begin with them.

> Given a point set S, a *topology* is a family T of subsets of S to which belong: (a) the whole set S and the empty set $\emptyset$; (b) the intersection $\bigcap_k U_k$ of any finite sub-family of members U_k of T; (c) the union $\bigcup_k U_k$ of any sub-family (finite or infinite) of members.

<div style="text-align: right">topology</div>

> A *topological space* (S, T) is a set of points S on which a topology T is defined.

The members of the family T are, by definition, the *open sets* of (S, T). Notice that a topological space is indicated by the pair (S, T). There are, in general, many different possible topologies on a given point set S, and each one will make of S a different topological space. Two extreme topologies are always possible on any S. The *discrete space* is the topological space $(S, P(S))$, with the power set $P(S)$ — the set of all subsets of S — as the topology. For each point p, the set $\{p\}$ containing only p is open. The other extreme case is the *indiscrete* (or trivial) topology $T = \{\emptyset, S\}$.

Any subset of S containing a point p is a *neighborhood* of p. The complement of an open set is (by definition) a *closed* set. A set which is open in a

topology may be closed in another. It follows that $\emptyset$ and S are closed (and open!) sets in all topologies.

Comment 2.1 The space (S, T) is *connected* if $\emptyset$ and S are the only sets which are simultaneously open and closed. In this case S cannot be decomposed into the union of two disjoint open sets (this is different from path-connectedness). In the discrete topology all open sets are also closed, so that unconnectedness is extreme.

§ **2.8** Let $f : A \to B$ be a function between two topological spaces A (the domain) and B (the target). The inverse image of a subset X of B by f is the set $f^{<-1>}(X) = \{a \in A$ such that $f(a) \in X\}$. The function f is *continuous* if the inverse images of all the open sets of the target space B are open sets of the domain space A. It is necessary to specify the topology whenever one speaks of a continuous function. A function defined on a discrete space is automatically continuous. On an indiscrete space, a function is hard put to be continuous.

$\boxed{\text{continuity}}$

§ **2.9** A topology is a *metric topology* when its open sets are the open balls $B_r(p) = \{q \in S$ such that $d(q,p) < r\}$ of some distance function. The simplest example of such a "ball-topology" is the discrete topology $P(S)$: it can be obtained from the so-called discrete metric: $d(p, q) = 1$ if $p \neq q$, and $d(p, q) = 0$ if $p = q$. In general, however, topologies are independent of any distance function: the trivial topology cannot be given by any metric.

§ **2.10** A *caveat* is in order here. When we say "metric" we mean a positive-definite distance function as above. Physicists use the word "metrics" for some invertible bilinear forms which are not positive-definite, and this practice is progressively infecting mathematicians. We shall follow this seemingly inevitable trend, though it should be clear that only positive-definite metrics can define a topology. The fundamental bilinear form of relativistic Physics, the Lorentz metric on Minkowski space-time, does not define true distances between points.

§ **2.11** We have introduced Euclidean spaces $\mathbb{E}^n$ in §2.2. These spaces, and Euclidean half-spaces (or upper-spaces) $\mathbb{E}^n_+$ are, at least for Physics, the most important of all topological spaces. This is so because Physics deals

mostly with manifolds, and a manifold (differentiable or not) will be a space which can be approximated by some $\mathbb{E}^n$ or $\mathbb{E}^n_+$ in some neighborhood of each point (that is, "locally"). The half-space $\mathbb{E}^n_+$ has for point set $\mathbb{R}^n_+ = \{p = (p^1, p^2, ..., p^n) \in \mathbb{R}^n \text{ such that } p^n \geq 0\}$. Its topology is that "induced" by the ball-topology of $\mathbb{E}^n$ (the open sets are the intersections of $\mathbb{R}^n_+$ with the balls of $\mathbb{E}^n$). This space is essential to the definition of manifolds-with-boundary.

§ **2.12** A bijective function $f : A \rightarrow B$ will be a *homeomorphism* if it is continuous and has a continuous inverse. It will take open sets into open sets and its inverse will do the same. Two spaces are *homeomorphic* when there exists a homeomorphism between them. A homeomorphism is an equivalence relation: it establishes a complete equivalence between two topological spaces, as it preserves all the purely topological properties. Under a homeomorphism, images and pre-images of open sets are open, and images and pre-images of closed sets are closed. *Two homeomorphic spaces are just the same topological space.* A straight line and one branch of a hyperbola are the same topological space. The same is true of the circle and the ellipse. A 2-dimensional sphere S^2 can be stretched in a continuous way to become an ellipsoid or a tetrahedron. From a purely topological point of view, these three surfaces are indistinguishable. There is no homeomorphism, on the other hand, between S^2 and a torus T^2, which is a quite distinct topological space.

Take again the Euclidean space $\mathbb{E}^n$. Any isometry (distance–preserving mapping) will be a homeomorphism, in particular any translation. Also homothecies with reason $\alpha \neq 0$ are homeomorphisms. From these two properties it follows that each open ball of $\mathbb{E}^n$ is homeomorphic to the whole $\mathbb{E}^n$. Suppose a space S has some open set U which is homeomorphic to an open set (a ball) in some $\mathbb{E}^n$: there is a homeomorphic mapping $\phi : U \rightarrow$ ball, $f(p \in U) = x = (x^1, x^2, ..., x^n)$. Such a local homeomorphism ϕ, with $\mathbb{E}^n$ as target space, is called a *coordinate mapping* and the values x^k are *coordinates* of p.

§ **2.13** S is *locally Euclidean* if, for every point $p \in S$, there exists an open set U to which p belongs, which is homeomorphic to either an open set in

some $\mathbf{E}^s$ or an open set in some $\mathbf{E}^s_+$. The number s is the dimension of S at the point p.

§ **2.14** We arrive in this way at one of the concepts announced at the beginning of this chapter: a (topological) *manifold* is a connected space on which coordinates make sense.

> A *manifold* is a topological space S which is
> (i) locally Euclidean;
> (ii) has the same dimension s at all points, which is then the
> *dimension* of S, $s = \dim S$.

Points whose neighborhoods are homeomorphic to open sets of $\mathbf{E}^s_+$ and not to open sets of $\mathbf{E}^s$ constitute the *boundary* ∂S of S. Manifolds including points of this kind are "manifolds–with–boundary".

The local-Euclidean character will allow the definition of coordinates and will have the role of a "complementarity principle": in the local limit, a differentiable manifold will look still more Euclidean than the topological manifolds. Notice that we are indicating dimensions by m, n, s, etc, and manifolds by the corresponding capitals: $\dim M = m$; $\dim N = n$, $\dim S = s$, etc.

§ **2.15** Each point p on a manifold has a neighborhood U homeomorphic to an open set in some $\mathbf{E}^n$, and so to $\mathbf{E}^n$ itself. The corresponding homeomorphism

$$\phi : U \to \text{ open set in } \mathbf{E}^n$$

will give local coordinates around p. The neighborhood U is called a *coordinate neighborhood* of p. The pair (U, ϕ) is a *chart*, or *local system of coordinates* (LSC) around p.

We must be more specific. Take $\mathbf{E}^n$ itself: an open neighborhood V of a point $q \in \mathbf{E}^n$ is homeomorphic to another open set of $\mathbf{E}^n$. Each homeomorphism $u\colon V \to V'$ included in $\mathbf{E}^n$ defines a *system of coordinate functions* (what we usually call coordinate systems: Cartesian, polar, spherical, elliptic, stereographic, etc.). Take the composite homeomorphism x: $S \to \mathbf{E}^n$, x(p) $= (x^1, x^2, ..., x^n) = (u^1 \circ \phi(p), u^2 \circ \phi(p), ..., u^n \circ \phi(p))$. The functions

$x^i = u^i \circ \phi$: U $\to E^1$ will be the *local coordinates* around p. We shall use the simplified notation (U, x) for the chart. Different systems of coordinate functions require different number of charts to plot the space S. For $\mathbf{E}^2$ itself, coordinates one Cartesian system is enough to chart the whole space: V = $\mathbf{E}^2$, u = the identity mapping. The polar system, however, requires at least two charts. For the sphere S^2, stereographic coordinates require only two charts, while the cartesian system requires four.

Comment 2.2 Suppose the polar system with only one chart: $\mathbf{E}^2 \to \mathbb{R}^1_+ \times (0, 2\pi)$. Intuitively, close points $(r, 0+\epsilon)$ and $(r, 2\pi - \epsilon)$, for ϵ small, are represented by faraway points. Technically, due to the necessity of using open sets, the whole half-line $(r, 0)$ is absent, not represented. Besides the chart above, it is necessary to use $\mathbf{E}^2 \to \mathbb{R}^1_+ \times (\alpha, \alpha + 2\pi)$, with α arbitrary in the interval $(0, 2\pi)$.

Comment 2.3 Classical Physics needs coordinates to distinguish points. We see that the method of coordinates can only work on locally Euclidean spaces.

§ 2.16 As we have said, every time we write a derivative, a differential, a Laplacian we are assuming an additional underlying structure for the space we are working on: it must be a differentiable (or smooth) manifold. And manifolds and smooth manifolds can be introduced by imposing progressively restrictive conditions on the decomposition which has led to topological spaces. Just as not every space accepts coordinates (that is, not not every space is a manifold), there are spaces on which to differentiate is impossible. We arrive finally at the crucial notion by which knowledge on differentiability on Euclidean spaces is translated into knowledge on differentiability on more general spaces. We insist that knowledge of Analysis on Euclidean spaces is taken for granted.

A given point $p \in S$ can in principle have many different coordinate neighborhoods and charts. Given any two charts (U, x) and (V, y) with $U \bigcap V \neq \emptyset$, to a given point p in their intersection, $p \in U \bigcap V$, will correspond coordinates $x = x(p)$ and $y = y(p)$. These coordinates will be related by a homeomorphism between open sets of $\mathbf{E}^n$,

$$y \circ x^{<-1>} : \mathbf{E}^n \to \mathbf{E}^n$$

which is a coordinate transformation, usually written $y^i = y^i(x^1, x^2, \dots, x^n)$. Its inverse is $x \circ y^{<-1>}$, written $x^j = x^j(y^1, y^2, ..., y^n)$. Both the coordinate

transformation and its inverse are functions between Euclidean spaces. If both are C^∞ (differentiable to any order) as functions from $\mathbf{E}^n$ into $\mathbf{E}^n$, the two local systems of coordinates are said to be *differentially related*. An *atlas* on the manifold is a collection of charts $\{(U_a, y_a)\}$ such that $\bigcup_a U_a = S$. If all the charts are differentially related in their intersections, it will be a *differentiable atlas*.[*] The chain rule

$$\delta^i_k = \frac{\partial y^i}{\partial x^j}\frac{\partial x^j}{\partial y^k}$$

says that both Jacobians are $\neq 0$.[†]

An extra chart (W, x), not belonging to a differentiable atlas A, is said to be admissible to A if, on the intersections of W with all the coordinate-neighborhoods of A, all the coordinate transformations from the atlas LSC's to (W, x) are C^∞. If we add to a differentiable atlas all its admissible charts, we get a *complete atlas*, or maximal atlas, or C^∞–structure. The extension of a differentiable atlas, obtained in this way, is unique (this is a theorem).

A topological manifold with a complete differentiable atlas is a *differentiable manifold*.

> differentiable manifold

§ 2.17 A function f between two smooth manifolds is a *differentiable function* (or smooth function) when, given the two atlases, there are coordinates systems in which $y \circ f \circ x^{<-1>}$ is differentiable as a function between Euclidean spaces.

§ 2.18 A *curve* on a space S is a function $a : I \to S$, $a : t \to a(t)$, taking the interval $I = [0, 1] \subset \mathbf{E}^1$ into S. The variable $t \in I$ is the curve parameter. If the function a is continuous, then a is a *path*. If the function a is also differentiable, we have a *smooth curve*.[‡] When $a(0) = a(1)$, a is a *closed*

> curves

---

[*]This requirement of infinite differentiability can be reduced to k-differentiability (to give a "C^k–atlas").

[†]If some atlas exists on S whose Jacobians are all positive, S is *orientable*. When 2–dimensional, an orientable manifold has two faces. The Möbius strip and the Klein bottle are non-orientable manifolds.

[‡]The trajectory in a brownian motion is continuous (thus, a path) but is not differentiable (not smooth) at the turning points.

curve, or a *loop*, which can be alternatively defined as a function from the circle S^1 into S. Some topological properties of a space can be grasped by studying its possible paths.

Comment 2.4 This is the subject matter of homotopy theory. We shall need one concept — contractibility — for which the notion of homotopy is an indispensable preliminary.

Let $f, g : X \to Y$ be two continuous functions between the topological spaces X and Y. They are *homotopic* to each other ($f \approx g$) if there exists a continuous function $F : X \times I \to Y$ such that $F(p, 0) = f(p)$ and $F(p, 1) = g(p)$ for every $p \in X$. The function $F(p, t)$ is a one-parameter family of continuous functions interpolating between f and g, a *homotopy* between f and g. Homotopy is an equivalence relation between continuous functions and establishes also a certain equivalence between spaces. Given any space Z, let $id_Z : Z \to Z$ be the identity mapping on Z, $id_Z(p) = p$ for every $p \in Z$. A continuous function $f : X \to Y$ is a homotopic equivalence between X and Y if there exists a continuous function $g : Y \to X$ such that $g \circ f \approx id_X$ and $f \circ g \approx id_Y$. The function g is a kind of "homotopic inverse" to f. When such a homotopic equivalence exists, X and Y are *homotopic*. Every homeomorphism is a homotopic equivalence but not every homotopic equivalence is a homeomorphism.

Comment 2.5 A space X is *contractible* if it is homotopically equivalent to a point. More precisely, there must be a continuous function $h : X \times I \to X$ and a constant function $f : X \to X$, $f(p) = c$ (a fixed point) for all $p \in X$, such that $h(p, 0) = p = id_X(p)$ and $h(p, 1) = f(p) = c$. Contractibility has important consequences in standard, 3-dimensional vector analysis. For example, the statements that divergenceless fluxes are rotational ($div \ \mathbf{v} = 0 \Rightarrow \mathbf{v} = rot \ \mathbf{w}$) and irrotational fluxes are potential ($rot \ \mathbf{v} = 0 \Rightarrow \mathbf{v} = grad \ \phi$) are valid only on contractible spaces. These properties generalize to differential forms (see page 38).

§ **2.19** We have seen that two spaces are equivalent from a purely topological point of view when related by a homeomorphism, a topology-preserving transformation. A similar role is played, for spaces endowed with a differentiable structure, by a diffeomorphism: a *diffeomorphism* is a differentiable $\boxed{\begin{array}{c}\textbf{diffeo--}\\\textbf{morphism}\end{array}}$ homeomorphism whose inverse is also smooth. When some diffeomorphism exists between two smooth manifolds, they are said to be *diffeomorphic*. In this case, besides being topologically the same, they have equivalent differentiable structures. They are the same differentiable manifold.

§ **2.20** *Linear spaces* (or vector spaces) are spaces allowing for addition and rescaling of their members. This means that we know how to add two vectors $\boxed{\begin{array}{c}\textbf{vector}\\\textbf{space}\end{array}}$

so that the result remains in the same space, and also to multiply a vector by some number to obtain another vector, also a member of the same space. In the cases we shall be interested in, that number will be a complex number. In that case, we have a vector space V over the field C of complex numbers. Every vector space V has a dual V^*, another linear space formed by all the linear mappings taking V into C. If we indicate a vector $\in V$ by the "ket" $|v>$, a member of the dual can be indicated by the "bra" $<u|$. The latter will be a linear mapping taking, for example, $|v>$ into a complex number, which we indicate by $<u|v>$. Being linear means that a vector $a|v> + b|w>$ will be taken by $<u|$ into the complex number $a<u|v> + b<u|w>$. Two linear spaces with the same finite dimension (= maximal number of linearly independent vectors) are isomorphic. If the dimension of V is finite, V and V^* have the same dimension and are, consequently, isomorphic.

Comment 2.6 Every vector space is contractible. Many of the most remarkable properties of $\mathbf{E}^n$ come from its being, besides a topological space, a vector space. $\mathbf{E}^n$ itself and any open ball of $\mathbf{E}^n$ are contractible. This means that any coordinate open set, which is homeomorphic to some such ball, is also contractible.

Comment 2.7 A vector space V can have a norm, which is a distance function and defines consequently a certain topology called the "norm topology". In this case, V is a metric space. For instance, a norm may come from an inner product, a mapping from the Cartesian set product $V \times V$ into C, $V \times V \rightarrow C$, $(v, u) \rightarrow <v, u>$ with suitable properties. The number $\|v\| = (|<v, v>|)^{1/2}$ will be the norm of $v \in V$ induced by the inner product. This is a special norm, as norms can be defined independently of inner products. When the norm comes from an inner space, we have a Hilbert space. When not, a Banach space. When the operations (multiplication by a scalar and addition) keep a certain coherence with the topology, we have a topological vector space.

Once in possession of the means to define coordinates, we can proceed to transfer to manifolds all the (supposedly well–known) results of usual vector and tensor analysis on Euclidean spaces. Because a manifold is equivalent to an Euclidean space only locally, this will be possible only in a certain neighborhood of each point. This is the basic difference between Euclidean spaces and general manifolds: properties which are "global" on the first hold only locally on the latter.

2.1.2 Vector and Tensor Fields

§ 2.21 The best means to transfer the concepts of vectors and tensors from Euclidean spaces to general differentiable manifolds is through the mediation of spaces of functions. We have talked on function spaces, such as Hilbert spaces separable or not, and Banach spaces. It is possible to define many distinct spaces of functions on a given manifold M, differing from each other by some characteristics imposed in their definitions: square–integrability for example, or different kinds of norms. By a suitable choice of conditions we can actually arrive at a space of functions containing every information on M. We shall not deal with such involved subjects. At least for the time being, we shall need only spaces with poorly defined structures, such as the space of real functions on M, which we shall indicate by $R(M)$.

§ 2.22 Of the many equivalent notions of a vector on $\mathbb{E}^n$, the directional derivative is the easiest to adapt to differentiable manifolds. Consider the set $R(\mathbb{E}^n)$ of real functions on $\mathbb{E}^n$. A vector $V = (v^1, v^2, \ldots, v^n)$ is a linear operator on $R(\mathbb{E}^n)$: take a point $p \in \mathbb{E}^n$ and let $f \in R(\mathbb{E}^n)$ be differentiable in a neighborhood of p. The vector V will take f into the real number

$$V(f) = v^1 \left[\frac{\partial f}{\partial x^1} \right]_p + v^2 \left[\frac{\partial f}{\partial x^2} \right]_p + \cdots + v^n \left[\frac{\partial f}{\partial x^n} \right]_p.$$

This is the directional derivative of f along the vector V at p. This action of V on functions respects two conditions:

1. linearity: $V(af + bg) = aV(f) + bV(g)$, $\forall a, b \in \mathbb{E}^1$ and $\forall f, g \in R(\mathbb{E}^n)$;

2. Leibniz rule: $V(f \cdot g) = f \cdot V(g) + g \cdot V(f)$.

§ 2.23 This conception of vector – an operator acting on functions – can be defined on a differential manifold N as follows. First, introduce a curve through a point $p \in N$ as a differentiable curve $a : (-1, 1) \to N$ such that $a(0) = p$ (see page 27). It will be denoted by $a(t)$, with $t \in (-1, 1)$. When t varies in this interval, a 1-dimensional continuum of points is obtained on N. In a chart (U, x) around p, these points will have coordinates $a^i(t) = x^i(t)$.

Consider now a function $f \in R(N)$. The vector V_p tangent to the curve $a(t)$ at p is given by

$$V_p(f) = \left[\frac{d}{dt}(f \circ a)(t) \right]_{t=0} = \left[\frac{dx^i}{dt} \right]_{t=0} \frac{\partial}{\partial x^i} f \ .$$

V_p is independent of f, which is arbitrary. It is an operator $V_p : R(N) \to \mathbb{E}^1$. Now, any vector V_p, tangent at p to some curve on N, is a *tangent vector to N at p*. In the particular chart used above, $\frac{dx^k}{dt}$ is the k-th component of V_p. The components are chart-dependent, but V_p itself is not. From its very definition, V_p satisfies the conditions (1) and (2) above. A *tangent vector* on N at p is just that, a mapping $V_p : R(N) \to \mathbb{E}^1$ which is linear and satisfies the Leibniz rule.

§ 2.24 The vectors tangent to N at p constitute a linear space, the *tangent space T_pN* to the manifold N at p. Given some coordinates $x(p) = (x^1, x^2, \ldots, x^n)$ around the point p, the operators $\{\frac{\partial}{\partial x^i}\}$ satisfy conditions (1) and (2) above. More than that, they are linearly independent and consequently constitute a basis for the linear space: any vector can be written in the form

$$V_p = V_p^i \frac{\partial}{\partial x^i}.$$

The V_p^i's are the components of V_p in this basis. Notice that each coordinate x^j belongs to $R(N)$. The basis $\{\frac{\partial}{\partial x^i}\}$ is the *natural, holonomic,* or *coordinate* basis associated to the coordinate system $\{x^j\}$. Any other set of n vectors $\{e_i\}$ which are linearly independent will provide a base for T_pN. If there is *no* coordinate system $\{y^k\}$ such that $e_k = \frac{\partial}{\partial y^k}$, the base $\{e_i\}$ is *anholonomic* or *non-coordinate*.

§ 2.25 T_pN and $\mathbb{E}^n$ are finite vector spaces of the same dimension and are consequently isomorphic. The tangent space to $\mathbb{E}^n$ at some point will be itself an $\mathbb{E}^n$. Euclidean spaces are diffeomorphic to their own tangent spaces, and that explains in part their simplicity — in equations written on such spaces, one can treat indices related to the space itself and to the tangent spaces on the same footing. This cannot be done on general manifolds.

These tangent vectors are called simply *vectors*, or *contravariant vectors*. The members of the dual *cotangent space T_p^*N*, the linear mappings $\omega_p: T_pN \to \mathbb{E}^n$, are *covectors*, or *covariant vectors*.

§ **2.26** Given an arbitrary basis $\{e_i\}$ of T_pN, there exists a unique basis $\{\alpha^j\}$ of T_p^*N, its dual basis, with the property $\alpha^j(e_i) = \delta_i^j$. Any $\omega_p \in T_p^*N$, is written $\omega_p = \omega_p(e_i)\alpha^i$. Applying V_p to the coordinates x^i, we find $V_p^i = V_p(x^i)$, so that $V_p = V_p(x^i)\frac{\partial}{\partial x^i} = \alpha(V_p)e_i$. The members of the basis dual to the natural basis $\{\frac{\partial}{\partial x^i}\}$ are indicated by $\{dx^i\}$, with $dx^j(\frac{\partial}{\partial x^i}) = \delta_i^j$. This notation is justified in the usual cases, and extended to general manifolds (when f is a function between general differentiable manifolds, df takes vectors into vectors). The notation leads also to the reinterpretation of the usual expression for the differential of a function, $df = \frac{\partial f}{\partial x^i}dx^i$, as a linear operator:

$$df(V_p) = \frac{\partial f}{\partial x^i}dx^i(V_p).$$

In a natural basis,

$$\omega_p = \omega_p(\frac{\partial}{\partial x^i})dx^i.$$

§ **2.27** The same order of ideas can be applied to tensors in general: a $\boxed{\text{tensors}}$ tensor at a point p on a differentiable manifold M is defined as a tensor on T_pM. The usual procedure to define tensors – covariant and contravariant – on Euclidean vector spaces can be applied also here. A covariant tensor of order s, for example, is a multilinear mapping taking the Cartesian product $T_p^{\times s}M = T_pM \times T_pM \cdots \times T_pM$ of T_pM by itself s-times into the set of real numbers. A contravariant tensor of order r will be a multilinear mapping taking the the Cartesian product $T_p^{*\times r}M = T_p^*M \times T_p^*M \cdots \times T_p^*M$ of T_p^*M by itself r-times into $\mathbb{E}^1$. A mixed tensor, s-times covariant and r-times contravariant, will take the Cartesian product $T_p^{\times s}M \times T_p^{*\times r}M$ multilinearly into $\mathbb{E}^1$. Basis for these spaces are built as the direct product of basis for the corresponding vector and covector spaces. The whole lore of tensor algebra is in this way transmitted to a point on a manifold. For example, a symmetric covariant tensor of order s applies to s vectors to give a real number, and is indifferent to the exchange of any two arguments:

$$T(v_1, v_2, \ldots, v_k, \ldots, v_j, \ldots, v_s) = T(v_1, v_2, \ldots, v_j, \ldots, v_k, \ldots, v_s).$$

An antisymmetric covariant tensor of order s applies to s vectors to give a real number, and change sign at each exchange of two arguments:

$$T(v_1, v_2, \ldots, v_k, \ldots, v_j, \ldots, v_s) = - T(v_1, v_2, \ldots, v_j, \ldots, v_k, \ldots, v_s).$$

§ **2.28** Because they will be of special importance, let us say a little more on such antisymmetric covariant tensors. At each fixed order, they constitute a vector space. But the tensor product $\omega \otimes \eta$ of two antisymmetric tensors ω and η of orders p and q is a $(p+q)$-tensor which is not antisymmetric, so that the antisymmetric tensors do not constitute a subalgebra with the tensor product.

§ **2.29** The wedge product is introduced to recover a closed algebra. First we define the alternation Alt(T) of a covariant tensor T, which is an anti-symmetric tensor given by

$$Alt(T)(v_1, v_2, \ldots, v_s) = \frac{1}{s!} \sum_{(P)} (sign\ P) T(v_{p_1}, v_{p_2}, \ldots, v_{p_s}),$$

the summation taking place on all the permutations $P = (p_1, p_2, \ldots, p_s)$ of the numbers (1,2,. . . , s) and (sign P) being the parity of P. Given two antisymmetric tensors, ω of order p and η of order q, their *exterior product*, or wedge product, indicated by $\omega \wedge \eta$, is the (p+q)-antisymmetric tensor

$$\omega \wedge \eta = \frac{(p+q)!}{p!\ q!} Alt(\omega \otimes \eta).$$

With this operation, the set of antisymmetric tensors constitutes the *exterior algebra*, or *Grassmann algebra*, encompassing all the vector spaces of antisymmetric tensors. The following properties come from the definition:

$$(\omega + \eta) \wedge \alpha = \omega \wedge \alpha + \eta \wedge \alpha; \qquad (2.1)$$

$$\alpha \wedge (\omega + \eta) = \alpha \wedge \omega + \alpha \wedge \eta; \qquad (2.2)$$

$$a(\omega \wedge \eta) = (a\omega) \wedge \eta = \omega \wedge (a\eta), \ \ \forall\ a \in \mathbb{R}; \qquad (2.3)$$

$$(\omega \wedge \eta) \wedge \alpha = \omega \wedge (\eta \wedge \alpha); \qquad (2.4)$$

$$\omega \wedge \eta = (-)^{\partial_\omega \partial_\eta} \eta \wedge \omega . \qquad (2.5)$$

In the last property, ∂_ω and ∂_η are the respective orders of ω and η. If $\{\alpha^i\}$ is a basis for the covectors, the space of s-order antisymmetric tensors has a basis

$$\{\alpha^{i_1} \wedge \alpha^{i_2} \wedge \cdots \wedge \alpha^{i_s}\}, 1 \leq i_1, i_2, \ldots, i_s \leq \dim\ T_pM, \qquad (2.6)$$

Grassmann algebra

in which an antisymmetric covariant s-tensor will be written

$$\omega = \frac{1}{s!}\,\omega_{i_1 i_2 \ldots i_s}\alpha^{i_1} \wedge \alpha^{i_2} \wedge \cdots \wedge \alpha^{i_s}\ .$$

In a natural basis $\{dx^j\}$,

$$\omega = \frac{1}{s!}\,\omega_{i_1 i_2 \ldots i_s}dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_s}\ .$$

§ **2.30** Thus, a tensor at a point $p \in M$ is a tensor defined on the tangent space T_pM. One can choose a chart around p and use for T_pM and T_p^*M the natural bases $\{\frac{\partial}{\partial x^i}\}$ and $\{dx^j\}$. A general tensor will be written

$$T = T^{i_1 i_2 \ldots i_r}_{j_1 j_2 \ldots j_s}\,\frac{\partial}{\partial x^{i_1}} \otimes \frac{\partial}{\partial x^{i_2}} \otimes \cdots \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes dx^{j_2} \otimes \cdots \otimes dx^{j_s}\ .$$

In another chart, with natural bases $\{\frac{\partial}{\partial x^{i'}}\}$ and $(dx^{j'})$, the same tensor will be written

$$T = T^{i'_1 i'_2 \ldots i'_r}_{j'_1 j'_2 \ldots j'_s}\,\frac{\partial}{\partial x^{i'_1}} \otimes \frac{\partial}{\partial x^{i'_2}} \otimes \cdots \frac{\partial}{\partial x^{i'_r}} \otimes dx^{j'_1} \otimes dx^{j'_2} \otimes \cdots \otimes dx^{j'_s}$$

$$= T^{i'_1 i'_2 \ldots i'_r}_{j'_1 j'_2 \ldots j'_s}\,\frac{\partial x^{i_1}}{\partial x^{i'_1}} \otimes \frac{\partial x^{i_2}}{\partial x^{i'_2}} \otimes \cdots \frac{\partial x^{i_r}}{\partial x^{i'_r}} \otimes \frac{\partial x^{j'_1}}{\partial x^{j_1}} \otimes \frac{\partial x^{j'_2}}{\partial x^{j_2}} \otimes \cdots \frac{\partial x^{j'_s}}{\partial x^{j_s}}$$

$$\otimes \frac{\partial}{\partial x^{i_1}} \otimes \frac{\partial}{\partial x^{i_2}} \otimes \cdots \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes dx^{j_2} \otimes \cdots \otimes dx^{j_s}\ , \qquad (2.7)$$

which gives the transformation of the components under changes of coordinates in the charts' intersection. We find frequently tensors defined as entities whose components transform in this way, with one Lamé coefficient $\frac{\partial x^{j'_r}}{\partial x^{j_r}}$ for each index. It should be understood that a tensor is always a tensor with respect to a given group. Just above, the group of coordinate transformations was involved. General base transformations constitute another group.

§ **2.31** Vectors and tensors have been defined at a fixed point p of a differentiable manifold M. The natural basis we have used is actually $\{[\frac{\partial}{\partial x^i}]_p\}$. A vector at $p \in M$ has been defined as the tangent to a curve $a(t)$ on M, with $a(0) = p$. We can associate a vector to each point of the curve by allowing the variation of the parameter t: $X_{a(t)}(f) = \frac{d}{dt}(f \circ a)(t)$. $X_{a(t)}$ is then the *tangent field* to $a(t)$, and $a(t)$ is the *integral curve* of X through p. In general, this only makes sense locally, in a neighborhood of p. When X is tangent to a curve globally, X is a *complete field*.

vector fields

§ 2.32 Let us, for the sake of simplicity take a neighborhood U of p and suppose $a(t) \in U$, with coordinates $(a_1(t), a_2(t), \cdots, a_m(t))$. Then, $X_{a(t)} = \frac{da^i}{dt}\frac{\partial}{\partial a^i}$, and $\frac{da^i}{dt}$ is the component $X^i_{a(t)}$. In this sense, the field whose integral curve is $a(t)$ is given by the "velocity" $\frac{da}{dt}$. Conversely, if a field is given by its components X^k $(x^1(t), x^2(t), \dots, x^m(t))$ in some natural basis, its integral curve $x(t)$ is obtained by solving the system of differential equations $X^k = \frac{dx^k}{dt}$. Existence and uniqueness of solutions for such systems hold in general only locally, as most fields exhibit singularities and are not complete. Most manifolds accept no complete vector fields at all. Those which do are called *parallelizable*. Toruses are parallelizable, but, of all the spheres S^n, only S^1, S^3 and S^7 are parallelizable. S^2 is not.[§]

§ 2.33 At a point p, V_p takes a function belonging to $R(M)$ into some real number, $V_p : R(M) \to \mathbf{R}$. When we allow p to vary in a coordinate neighborhood, the image point will change as a function of p. By using successive cordinate transformations and as long as singularities can be surounded, V can be extended to M. Thus, a vector field is a mapping $V : R(M) \to R(M)$. In this way we arrive at the formal definition of a field:

a vector field V on a smooth manifold M is a linear mapping $V : R(M) \to R(M)$ obeying the Leibniz rule:

$$X(f \cdot g) = f \cdot X(g) + g \cdot X(f), \forall f, g \in R(M).$$

We can say that a vector field is a differentiable choice of a member of T_pM at each p of M. An analogous reasoning can be applied to arrive at tensors fields of any kind and order.

§ 2.34 Take now a field X, given as $X = X^i\frac{\partial}{\partial x^i}$. As $X(f) \in R(M)$, another field as $Y = Y^i\frac{\partial}{\partial x^i}$ can act on X(f). The result,

$$YXf = Y^j\,\frac{\partial X^i}{\partial x^j}\frac{\partial f}{\partial x^i} + Y^jX^i\,\frac{\partial^2 f}{\partial x^j\partial x^i},$$

does not belong to the tangent space because of the last term, but the commutator

$$[X, Y] := (XY - YX) = \left(X^i\,\frac{\partial Y^j}{\partial x^i} - Y^i\,\frac{\partial X^j}{\partial x^i}\right)\frac{\partial}{\partial x^j}$$

[§]This is the hedgehog theorem: you cannot comb a hedgehog so that all its prickles stay flat; there will be always at least one singular point, like the head crown.

does, and is another vector field. The operation of commutation defines a **Lie algebra**
linear algebra. It is also easy to check that

$$[X, X] = 0, \tag{2.8}$$

$$[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0, \tag{2.9}$$

the latter being the Jacobi identity. An algebra satisfying these two conditions is a *Lie algebra*. Thus, the vector fields on a manifold constitute, with the operation of commutation, a Lie algebra.

2.1.3 Differential Forms

§ 2.35 Differential forms[¶] are antisymmetric covariant tensor fields on differentiable manifolds. They are of extreme interest because of their good behavior under mappings. A smooth mapping between M and N take differential forms on N into differential forms on M (yes, in that inverse order) while preserving the operations of exterior product and exterior differentiation (to be defined below). In Physics they have acquired the status of a new vector calculus: they allow to write most equations in an invariant (coordinate- and frame-independent) way. The covector fields, or Pfaffian forms, or still 1-forms, provide basis for higher-order forms, obtained by exterior product [see eq. (2.6)]. The exterior product, whose properties have been given in eqs.(2.1)-(2.5), generalizes the vector product of $\mathbb{E}^3$ to spaces of any dimension and thus, through their tangent spaces, to general manifolds.

§ 2.36 The exterior product of two members of a basis $\{\omega^i\}$ is a 2-form, typical member of a basis $\{\omega^i \wedge \omega^j\}$ for the space of 2-forms. In this basis, a 2-form F, for instance, will be written $F = \frac{1}{2}F_{ij}\omega^i \wedge \omega^j$. The basis for the m-forms on an m-dimensional manifold has a unique member, $\omega^1 \wedge \omega^2 \cdots \omega^m$. The nonvanishing m-forms are called volume elements of M, or volume forms.

[¶] On the subject, a beginner should start with H. Flanders, *Differential Forms*, Academic Press, New York, l963; and then proceed with C. Westenholz, *Differential Forms in Mathematical Physics*, North-Holland, Amsterdam, l978; or W. L. Burke, *Applied Differential Geometry*, Cambridge University Press, Cambridge, l985; or still with R. Aldrovandi and J. G. Pereira, *Geometrical Physics*, World Scientific, Singapore, l995.

§ **2.37** The name "differential forms" is misleading: most of them are not differentials of anything. Perhaps the most elementary form in Physics is the mechanical work, a Pfaffian form in $\mathbb{E}^3$. In a natural basis, it is written $W = F_k dx^k$, with the components F_k representing the force. The total work realized in taking a particle from a point a to point b along a line γ is

$$W_{ab}[\gamma] = \int_\gamma W = \int_\gamma F_k dx^k,$$

and in general depends on the chosen line. It will be path-independent only when the force comes from a potential U as a gradient, $F_k = - (grad\ U)_k$. In this case $W = -dU$, truly the differential of a function, and $W_{ab} = U(a) - U(b)$. An integrability criterion is: $W_{ab}[\gamma] = 0$ for γ any closed curve. Work related to displacements in a non-potential force field is a typical non-differential 1-form. Another well-known example is heat exchange.

§ **2.38** In a more geometric mood, the form appearing in the integrand of the arc length $\int_a^x ds$ is not the differential of a function, as the integral obviously depends on the trajectory from a to x, and is a multi-valued function of x. The elementary length ds is a prototype form which is not a differential, despite its conventional appearance. A 1-form is *exact* if it is a gradient, like $\omega = dU$. Being exact is not the same as being integrable. Exact forms are integrable, but non-exact forms may also be integrable if they are of the form fdU.

§ **2.39** The 0-form f has the differential $df = \frac{\partial f}{\partial x^i}dx^i = \frac{\partial f}{\partial x^i} \wedge dx^i$, which is a 1-form. The generalization of this differential of a function to forms of any order is the differential operator d with the following properties:

1. when applied to a k-form, d gives a (k+1)-form;

2. $d(\alpha + \beta) = d\alpha + d\beta$;

3. $d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-)^{\partial_\alpha} \alpha \wedge d(\beta)$, where ∂_α is the order of α;

4. $d^2\alpha = dd\alpha \equiv 0$ for any form α.

§ 2.40 The invariant, basis-independent definition of the differential of a k-form is given in terms of vector fields:

$$d\alpha(X_0, X_1, \ldots, X_k) = \sum_{i=0}^{k} (-)^i X_i \left[\alpha(X_0, X_1, \ldots, X_{i-1}, \hat{X}_i, X_{i+1} \ldots, X_k) \right]$$

$$+ \sum_{i<j} (-)^{i+j} \alpha([X_i, X_j], X_0, X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j \ldots, X_k) \tag{2.10}$$

Wherever it appears, the notation $\hat{X}_n$ means that X_n is absent. From this definition, or from the systematic use of the defining conditions, we can obtain the first examples of derivatives:

- if f is a function (0-form), $df = \partial_i f\ dx^i$ (gradient) ;

- if $A = A_i\ dx^i$ is a covector (1-form), then $dA = \frac{1}{2}(\partial_i A_j - \partial_j A_i)\ dx^i \wedge dx^j$ (rotational)

Comment 2.8 To grasp something about the meaning of

$$d^2\alpha \equiv 0,$$

which is usually called the Poincaré lemma, let us examine the simplest case, a 1-form α in a natural basis: $\alpha = \alpha_i dx^i$. Its differential is $d\alpha = (d\alpha_i) \wedge dx^i + \alpha_i \wedge d(dx^i) = \frac{\partial \alpha_i}{\partial x^j} dx^i \wedge dx^j = \frac{1}{2} \left(\frac{\partial \alpha_i}{\partial x^j} - \frac{\partial \alpha_j}{\partial x^i} \right) dx^i \wedge dx^j$. If α is exact, $\alpha = df$ (in components, $\alpha_i = \partial_i f$) then

$$d\alpha = d^2 f = \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial^2 f}{\partial x^j \partial x^i} \right] dx^i \wedge dx^j$$

and the property $d^2 f \equiv 0$ is just the symmetry of the mixed second derivatives of a function. Along the same lines, if α is not exact, we can consider

$$d^2\alpha = \frac{\partial^2 \alpha_i}{\partial x^j \partial x^k}\ dx^j \wedge dx^k \wedge dx^i = \frac{1}{2} \left[\frac{\partial^2 \alpha_i}{\partial x^j \partial x^k} - \frac{\partial^2 \alpha_i}{\partial x^k \partial x^j} \right] dx^j \wedge dx^k \wedge dx^i = 0.$$

Thus, the condition $d^2 \equiv 0$ comes from the equality of mixed second derivatives of the functions α_i, related to integrability conditions.

§ 2.41 A form α such that $d\alpha = 0$ is said to be *closed*. A form α which can be written as $\alpha = d\beta$ for some β is said to be *exact*.

Comment 2.9 It is natural to ask whether every closed form is exact. The answer, given by the *inverse Poincaré lemma*, is: yes, but only locally. It is yes in Euclidean spaces, and differentiable manifolds are locally Euclidean. Every closed form is locally exact. The precise meaning of "locally" is the following: if $d\alpha = 0$ at the point $p \in M$, then there exists a *contractible* (see below) neighborhood of p in which there exists a form β (the "local integral" of α) such that $\alpha = d\beta$. But attention: if γ is another form of the same order of β and satisfying $d\gamma = 0$, then also $\alpha = d(\beta + \gamma)$. There are infinite forms of which an exact form is the differential.

The inverse Poincaré lemma gives an expression for the local integral of $\alpha = d\beta$. In order to state it, we have to introduce still another operation on forms. Given in a natural basis the p-form

$$\alpha(x) = \alpha_{i_1 i_2 i_3 \ldots i_p}(x) dx^{i_1} \wedge dx^{i_2} \wedge dx^{i_3} \wedge \cdots \wedge dx^{i_p}$$

the *transgression* of α is the (p-1)-form

$$T\alpha = \sum_{j=1}^{p} (-)^{j-1} \int_0^1 dt t^{p-1} x^{i_j} \alpha_{i_1 i_2 i_3 \ldots i_p}(tx)$$
$$dx^{i_1} \wedge dx^{i_2} \ldots \wedge dx^{i_{j-1}} \wedge dx^{i_{j+1}} \wedge \cdots \wedge dx^{i_p}. \tag{2.11}$$

Notice that, in the x-dependence of α, x is replaced by (tx) in the argument. As t ranges from 0 to 1, the variables are taken from the origin up to x. This expression is frequently referred to as the *homotopy formula*.

The operation T is meaningful only in a star-shaped region, as x is linked to the origin by the straight line "tx", but can be generalized to a contractible region. Contractibility has been defined in Comment 2.5. Consider the interval $I = [0, 1]$. A space or domain X is *contractible* if there exists a continuous function $h : X \times I \to X$ and a constant function $f : X \to X$, $f(p) = c$ (a fixed point) for all $p \in X$, such that $h(p, 0) = p = id_X(p)$ and $h(p, 1) = f(p) = c$. Intuitively, X can be continuously contracted to one of its points. $\mathbb{E}^n$ is contractible (and, consequently, any coordinate neighborhood), but spheres S^n and toruses T^n are not. The limitation to the result given below comes from this strictly local property. Well, the lemma then says that, in a contractible region, any form α can be written in the form

$$\alpha = dT\alpha + Td\alpha. \tag{2.12}$$

When

$$d\alpha = 0 , \tag{2.13}$$

$$\alpha = dT\alpha, \tag{2.14}$$

so that α is indeed exact and the integral looked for is just $\beta = T\alpha$, always up to γ's such that $d\gamma = 0$. Of course, the formulae above hold globally on Euclidean spaces, which are

contractible. The condition for a closed form to be exact on the open set V is that V be contractible (say, a coordinate neighborhood). On a smooth manifold, every point has an Euclidean (consequently contractible) neighborhood — and the property holds at least locally. The sphere S^2 requires at least two neighborhoods to be charted, and the lemma holds only on each of them. The expression stating the closedness of α, $d\alpha = 0$ becomes, when written in components, a system of differential equations whose integrability (i.e., the existence of a unique integral β) is granted locally. In vector analysis on $\mathbb{E}^3$, this includes the already mentioned fact that an irrotational flux ($d\mathbf{v} = rot\ \mathbf{v} = 0$) is potential ($\mathbf{v} = grad\ U = dU$). If one tries to extend this from one of the S^2 neighborhoods, a singularity inevitably turns up.

§ **2.42** Let us finally comment on the mappings between differential manifolds and the announced good–behavior of forms. A C^∞ function $f : M \to N$ between differentiable manifolds M and N induces a mapping between the tangent spaces:

$$f_* : T_p M \to T_{f(p)} N.$$

If g is an arbitrary real function on N, $g \in R(N)$, this mapping is defined by

$$[f_*(X_p)](g) = X_p(g \circ f) \tag{2.15}$$

for every $X_p \in T_p M$ and all $g \in R(N)$. When $M = \mathbb{E}^m$ and $N = \mathbb{E}^n$, f_* is the jacobian matrix. In the general case, f_* is a homomorphism (a mapping which preserves the algebraic structure) of vector spaces, called the *differential* of f. It is also frequently written "df". When f and g are diffeomorphisms, then $(f \circ g)_* X = f_* \circ g_* X$. Still more important, a diffeomorphism f preserves the commutator:

$$f_*[X, Y] = [f_* X, f_* Y]. \tag{2.16}$$

Consider now an antisymmetric s-tensor $w_{f(p)}$ on the vector space $T_{f(p)} N$. Then f determines a tensor on $T_p M$ by

$$(f^* \omega)_p(v_1, v_2, \dots, v_s) = \omega_{f(p)}(f_* v_1, f_* v_2, \dots, f_* v_s). \tag{2.17}$$

Thus, the mapping f induces a mapping f^* between the tensor spaces, working however in the inverse sense. f^* is called a *pull-back* and f_*, by extension, *push-forward*. The pull-back has some wonderful properties:

pull–
back

- f^* is linear;

- $(f \circ g)^* = g^* \circ f^*$.

- f^* preserves the exterior product: $f^* (\omega \wedge \eta) = f^* \omega \wedge f^* \eta$;

- f^* preserves the exterior derivative: $f^*(d\omega) = d(f^* \omega)$.

Mappings between differential manifolds preserve, consequently, the most important aspects of differential exterior algebra. The well-defined behavior when mapped between different manifolds makes of the differential forms the most interesting of all tensors. Notice that all these properties apply also when f is simply some differentiable transformation mapping M into itself.

2.1.4 Metrics

§ **2.43** The Euclidean space $\mathbb{E}^3$ consists of the set of triples $\mathbb{R}^3$ with the ball-topology. The balls come from the Euclidean metric, a symmetric second-order positive-definite tensor g whose components are, in global Cartesian coordinates, given by $g_{ij} = \delta_{ij}$. Thus, $\mathbb{E}^3$ is $\mathbb{R}^3$ plus the Euclidean metric. We use this metric to measure lengths in our everyday life, but it happens frequently that another metric is simultaneously at work on the same $\mathbb{R}^3$. Suppose, for example, that the space is permeated by a medium endowed with a point-dependent isotropic refractive index $n(p)$. Light rays will "feel" the metric $g'_{ij} = n^2(p)\delta_{ij}$. To "feel" means that they will bend, acquire a "curved" aspect. Fermat's principle says simply that light rays will become geodesics of the new metric, the straightest possible curve if measurements are made using g'_{ij} instead of g_{ij}. As long as we proceed to measurements using only light rays, distances – optical lengths – will be different from those given by the Euclidean metric. Suppose further that the medium is some compressible fluid, with temperature gradients and all which is necessary to render point-dependent the derivative of the pressure with respect to the fluid density at fixed entropy, $c_s = \left(\frac{\partial p}{\partial \rho}\right)_S$. In that case, sound propagation will be governed by still another metric, $g''_{ij} = \frac{1}{c_s}\delta_{ij}$. Nevertheless, in both cases we use also the Euclidean metric to make measurements, and much of geometric optics and acoustics comes from comparing the results in both

metrics involved. These words are only to call attention to the fact that there is no such a thing like *the* metric of a space. It happens frequently that more than one is important in a given situation.

§ 2.44 Bilinear forms are covariant tensors of second order. The tensor product of two linear forms w and z is defined by $(w \otimes z)(X,Y) = w(X) \cdot z(Y)$. The most fundamental bilinear form appearing in Physics is the Lorentz metric on $\mathbf{R}^4$ (see the final of this section).

Given a basis $\{\omega^j\}$ for the space of 1-forms, the products $w^i \otimes w^j$, with $i, j = 1, 2, \ldots, m$, constitute a basis for the space of covariant 2-tensors, in terms of which a bilinear form g is written $g = g_{ij} \, w^i \otimes w^j$. In a natural basis, $g = g_{ij} \, dx^i \otimes dx^j$.

A metric on a smooth manifold is a bilinear form, denoted $g(X, Y)$, $X \cdot Y$ or $< X, Y >$, satisfying the following conditions:

1. it is indeed bilinear:
$$X \cdot (Y + Z) = X \cdot Y + X \cdot Z$$
$$(X + Y) \cdot Z = X \cdot Z + Y \cdot Z;$$

2. it is symmetric:
$$X \cdot Y = Y \cdot X;$$

3. it is non-singular: if $X \cdot Y = 0$ for every field Y, then $X = 0$.

In a basis, $g(X_i, X_j) = X_i \cdot X_j = g_{mn} \, \omega^m(X_i) \, \omega^n(X_j)$, so that

$$g_{ij} = g_{ji} = g(X_i, X_j) = X_i \cdot X_j.$$

It is standard notation to write simply $w^i w^j = w^{(i} \otimes w^{j)} = \frac{1}{2}(w^i \otimes w^j + w^j \otimes w^i)$ for the symmetric part of the bilinear basis, so that

$$g = g_{ij} \, w^i w^j$$

or, in a natural basis,

$$g = g_{ij} \, dx^i dx^j.$$

§ **2.45** Given a field $Y = Y^i e_i$ and a form $z = z_j w^j$ in the dual basis, $z(Y) = \; < z, Y > \; = z_j Y^j$. A metric establishes a relation between vector and covector fields: Y is said to be the contravariant image of a form z if, for every X, $g(X,Y) = z(X)$. Then $g_{ij} Y^j = z_i$. In this case, we write simply $z_j = Y_j$. This is the usual role of the covariant metric, to lower indices, taking a vector into the corresponding covector. If the mapping $Y \to z$ so defined is onto, the metric is non-degenerate. This is equivalent to saying that the matrix (g_{ij}) is invertible. A contravariant metric $\hat{g}$ can then be introduced whose components (denoted by g^{rs}) are the elements of the matrix inverse to (g_{ij}). If w and z are the covariant images of X and Y, defined in a way inverse to the image given above, then $\hat{g}(w,z) = g(X,Y)$. All this defines on the spaces of vector and covector fields an internal product $(X,Y) := (w,z) := g(X,Y) = \hat{g}(w,z)$. Invertible metrics are called *semi-Riemannian*. Although physicists usually call them just Riemannian, mathematicians more frequently reserve this denomination to non-degenerate positive-definite metrics, with values in the positive real line $\mathbf{R}_+$.

As the Lorentz metric is not positive definite, it does not define balls and is consequently unable to provide for a topology on Minkowski space-time (whose topology is, by the way, unknown). A Riemannian manifold is a smooth manifold on which a Riemannian metric is defined. A theorem due to Whitney states that it is always possible to define at least one Riemannian metric on an arbitrary differentiable manifold.

A positive definite metric is presupposed in any measurement: lengths, angles, volumes, etc. The length of a vector X is introduced as

$$\|X\| = (X,X)^{1/2}.$$

A metric is *indefinite* when $\|X\| = 0$ does not imply $X = 0$. It is the case of Lorentz metric, which attributes zero length to vectors on the light cone.

The length of a curve $\gamma : (a,b) \to M$ is then defined as

$$L_\gamma = \int_a^b \|\frac{d\gamma}{dt} \; dt\|.$$

Given two points p, q on a Riemannian manifold M, consider all the piecewise differentiable curves γ with $\gamma(a) = p$ and $\gamma(b) = q$. The distance between

p and q is defined as the infimum of the lengths of all such curves between them:

$$d(p, q) = \inf_{\gamma(t)} \int_a^b \|\frac{dC}{dt} \ dt\|. \qquad (2.18)$$

In this way a metric tensor defines a distance function on M.

§ 2.46 The metrics referred to in the introduction of this section, concerned with simplified models for the behavior of light rays and sound waves, are both obtained by multiplying all the components of the Euclidean metric by a given function. A transformation like $g_{ij} \to g'_{ij} = f(p) \ g_{ij}$ is called a *conformal transformation*. In angle measurements, the metric appears in a numerator and in a denominator and in consequence two metrics differing by a conformal transformation will give the same angles. Conformal transformations preserve the angles, or the cones.

Comment 2.10 To find the angle made by two vector fields U and V at each point, calculate $\|U - V\|^2 = \|U\|^2 + \|V\|^2 - 2 \ U \cdot V = U \cdot U - V \cdot V - 2 \ \|U\|\|V\| \cos \theta_{UV}$, that is, $g_{\mu\nu}(U - V)^\mu(U - V)^\nu = g_{\mu\nu}U^\mu U^\nu + g_{\rho\sigma}V^\rho V^\sigma - \sqrt{g_{\mu\nu}U^\mu U^\nu}\sqrt{g_{\rho\sigma}V^\rho V^\sigma} \cos \theta_{UV}$. Then,

$$\theta_{UV} = \arccos \frac{g_{\mu\nu}U^\mu U^\nu + g_{\rho\sigma}V^\rho V^\sigma - g_{\mu\nu}(U - V)^\mu(U - V)^\nu}{\sqrt{g_{\mu\nu}U^\mu U^\nu}\sqrt{g_{\rho\sigma}V^\rho V^\sigma}} \ ,$$

which does not change if $g_{\mu\nu}$ is replaced by $g'_{\mu\nu} = f(p) \ g_{\mu\nu}$.

§ 2.47 Geometry has had a very strong historical bond to metric. "Geometries" have been synonymous of "kinds of metric manifolds". This comes from the impression that we measure something (say, distance from the origin) when attributing coordinates to a point. We do not. Only homeomorphisms are needed in the attribution, and they have nothing to do with metrics. We hope to have made it clear that a metric on a differentiable manifold is chosen at convenience.

§ 2.48 Minkowski spacetime is a 4-dimensional connected manifold on which a certain indefinite metric (the "Lorentz metric") is defined. Points on space- | **Minkowski spacetime** |
time are called "events". Being indefinite, the metric defines only a pseudo-distance function for any pair of points. Given two events x and y with Cartesian coordinates (x^0, x^1, x^2, x^3) and (y^0, y^1, y^2, y^3), their pseudo-distance will

be

$$s^2 = \eta_{\alpha\beta}\, x^\alpha x^\beta = (x^0 - y^0)^2 - (x^1 - y^1)^2 - (x^2 - y^2)^2 - (x^3 - y^3)^2 \ . \quad (2.19)$$

This pseudo-distance is called the "interval" between x and y. Notice the usual practice of attributing the first place, with index zero, to the time–related coordinate. The Lorentz metric does not define a topology, but establishes a partial ordering of the events: causality. A general spacetime will be any differentiable manifold S such that, at each point p, the tangent space $T_p S$ is a Minkowski spacetime. This will induce on S another metric, with the same set of signs (+,-,-,-) in the diagonalized form. A metric with that set of signs is said to be "Lorentzian".

In General Relativity, Einstein's equations determine a Lorentzian metric which will be "felt" by any particle or wave travelling in spacetime. They are non-linear second-order differential equations for the metric, with as source an energy-momentum density of the other fields in presence. Thus, the metric depends on the source and on the assumed boundary conditions.

2.2   Pseudo-Riemannian Metric

§ 2.49 Each spacetime is a 4–dimensional pseudo–Riemannian manifold. Its main character is the fundamental form, or metric

$$g(x) = g_{\mu\nu} dx^\mu dx^\nu. \quad (2.20)$$

This metric has signature 2. Being symmetric, the matrix $g(x) = (g_{\mu\nu})$ can be diagonalized. Signature concerns the signs of the eigenvalues: it is the numbers of eigenvalues with one sign minus the number of eigenvalues with the opposite sign. It is important because it is an invariant under changes of coordinates and vector bases. In the convention we shall adopt this means that, at any selected point P, it is possible to choose coordinates $\{x^\mu\}$ in terms of which $g_{\mu\nu}$ takes the form

$$g(P) = \begin{pmatrix} +|g_{00}| & 0 & 0 & 0 \\ 0 & -|g_{11}| & 0 & 0 \\ 0 & 0 & -|g_{22}| & 0 \\ 0 & 0 & 0 & -|g_{33}| \end{pmatrix} \ . \quad (2.21)$$

§ 2.50 The first example has been given above: it is the Lorentz metric of Minkowski space, for which we shall use the notation

$$\eta(x) = \eta_{ab}dx^a dx^b. \tag{2.22}$$

We are using indices $\mu, \nu, \lambda, \ldots$ for Riemannian spacetime, and $a, b, c, \ldots$ for Minkowski spacetime.

Minkowski space is the simplest, standard spacetime. Up to the signature, is an Euclidean space, and as such can be covered by a single, global coordinate system. This system — the cartesian system — is the father of all coordinate systems and just puts η in the diagonal form

$$g(P) = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{2.23}$$

Comment 2.11 A metric is a real symmetric non–singular bilinear form. A bilinear form takes two vectors into a real number: $g(V, U) = g_{\mu\nu}V^\mu U^\nu \in \mathbb{R}$. When symmetric, the numbers $g(V, U)$ and $g(U, V)$ are the same. A metric defines orthogonality. Two vectors U and V are orthogonal to each other by g if $g(V, U) = 0$. The metric components can be disposed in a matrix $(g_{\mu\nu})$. It will have, consequently, ten components on a 4–dimensional spacetime.

§ 2.51 Given the metric, a vector field V is timelike, spacelike or a null vector, depending on the sign of $g_{\mu\nu}V^\mu V^\nu$:

$$g_{\mu\nu}V^\mu V^\nu \begin{cases} > 0 & \text{timelike} \\ < 0 & \text{spacelike} \\ = 0 & \text{null} \end{cases} \tag{2.24}$$

V^μ are the components of the vector V in the coordinate system $\{x^\lambda\}$. Higher indices indicate a contravariant vector. The metric can be used to lower indices. From V, obtain a covariant vector, or covector, whose components are $V_\mu = g_{\mu\nu}V^\nu$. Einstein's summation convention is being used: repeated higher–lower indices are summed over everytime they appear, without the summation symbol.

§ 2.52 As presented above, with lower indices, the metric is sometimes called the *contravariant* metric. The elements of its inverse matrix are indicated by $g^{\mu\nu}$. Thus, with the Kronecker delta δ^μ_ν (which is $= 1$ when $\mu = \nu$ and zero otherwise), we have

$$g_{\mu\lambda}g^{\lambda\nu} = g^{\nu\lambda}g_{\lambda\mu} = \delta^\nu_\mu .$$

The set $\{g^{\mu\nu}\}$ is then called the *covariant* metric, and can be used to raise indices, or to get a contravariant vector from a covariant vector: $V^\mu = g^{\mu\nu} V_\nu$. The same holds for indices of general tensors. Frequently used notations are $g = |g| = \det(g_{\mu\nu})$. Of course, $\det(g^{\mu\nu}) = g^{-1}$.

§ 2.53 The norm ds of the infinitesimal displacement dx^μ, whose square is

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{2.25}$$

is the *infinitesimal interval*, or simply *interval*. Given a curve γ with extreme points a and b, the integral

$$L[\gamma] = \int_\gamma ds = \int_a^b ds \tag{2.26}$$

along γ is a function(al) of γ, called its "length".

Comment 2.12 Again a name used by extension of the strictly Riemannian case, in which this integrals is a true length. A strictly Riemannian metric determines a true distance between two points. In the case above, the distance between a and b would be the infimum of $L[\gamma]$, all curves considered.

Comment 2.13 The determinant of a metric matrix like (2.23) is always negative. In cartesian coordinates, an integration over 4-space has the form

$$\int_{V^4} d^4 x.$$

In another coordinate system, a Jacobian turns up. We recall that what appears in integration measures is the Jacobian up to the sign. Integration using a coordinate system in which the infinitesimal length takes the form (2.25), and in which the metric has a negative determinant, is given by the expression

$$\int_{V^4} \sqrt{-g}\, d^4 x,$$

which holds in every case.

2.3   The Notion of Connection

Besides well–behaved entities like tensors (including metrics, vectors and functions), a manifold contains other, not so well-behaved objects. The most important are connections, essential to the notion of parallelism. We proceed now to present the physicits' approach to connections.

§ 2.54 What we understand by good behavior is: covariance under change of coordinates. A scalar field is invariant under change of coordinates. Take the next case in complexity, a vector field V. It will have components V^α in a coordinate system $\{x^\alpha\}$ and components V^μ in a coordinate system $\{y^\mu\}$. The two sets are related by

$$V^\mu = \frac{\partial y^\mu}{\partial x^\alpha} \, V^\alpha. \tag{2.27}$$

This is the standard behavior. It defines a vector by the group of coordinate transformations. Tensors of any order reproduce it, index by index. The metric, for example, has its components changed according to

$$g_{\mu\nu} = \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} \, g_{\alpha\beta}. \tag{2.28}$$

Notice by the way that, contracting (2.27) with the gradient operator $\frac{\partial}{\partial y^\mu}$,

$$V^\mu \frac{\partial}{\partial y^\mu} = V^\alpha \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial}{\partial y^\mu} = V^\alpha \frac{\partial}{\partial x^\alpha} \ .$$

Thus, the expression

$$V = V^\alpha \frac{\partial}{\partial x^\alpha} \tag{2.29}$$

is invariant under change of coordinates. The vector field V, once conceived as such a directional derivative, is an invariant concept. This is the notion of vector field used by the mathematicians: a directional derivative acting on the functions defined on the manifold.

Take now the derivative of (2.27):

$$\frac{\partial}{\partial y^\lambda} \, V^\mu = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial}{\partial y^\lambda} \, V^\alpha + \frac{\partial}{\partial y^\lambda} \left[\frac{\partial y^\mu}{\partial x^\alpha} \right] \, V^\alpha$$

$$= \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial}{\partial x^\beta} \, V^\alpha + \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial}{\partial x^\beta} \left[\frac{\partial y^\mu}{\partial x^\alpha} \right] \, V^\alpha$$

$$= \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial}{\partial x^\beta} \, V^\alpha + \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial^2 y^\mu}{\partial x^\beta \partial x^\gamma} \, V^\gamma$$

$$= \frac{\partial x^\beta}{\partial y^\lambda} \left[\frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial}{\partial x^\beta} \, V^\alpha + \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial y^\rho} \frac{\partial^2 y^\rho}{\partial x^\beta \partial x^\gamma} \, V^\gamma \right] \ .$$

$$\therefore \quad \frac{\partial}{\partial y^\lambda} V^\mu = \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial y^\mu}{\partial x^\alpha} \left[\frac{\partial}{\partial x^\beta} V^\alpha + \frac{\partial x^\alpha}{\partial y^\rho} \frac{\partial^2 y^\rho}{\partial x^\beta \partial x^\gamma} V^\gamma \right]. \qquad (2.30)$$

If alone, the first term in the right–hand side would confer to the derivative a good tensor status. The second term breaks that behavior: the derivative of a vector is not a tensor. In other words, the derivative is not covariant. On a manifold, it is impossible to tell, for example, whether a vector is constant or not.

The solution is to change the very definition of derivative by adding another structure. We add to each derivative an extra term involving a new object Γ:

covariant derivative

$$\frac{\partial}{\partial y^\lambda} V^\mu \Rightarrow \frac{D}{Dy^\lambda} V^\mu = \frac{\partial}{\partial y^\lambda} V^\mu + \Gamma^\mu{}_{\nu\lambda} V^\nu$$

$$\frac{\partial}{\partial x^\beta} V^\alpha \Rightarrow \frac{D}{Dx^\beta} V^\alpha = \frac{\partial}{\partial x^\beta} V^\alpha + \Gamma^\alpha{}_{\gamma\beta} V^\gamma \qquad (2.31)$$

and impose good behavior of the modified derivative:

$$\frac{D}{Dy^\lambda} V^\mu = \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial y^\mu}{\partial x^\alpha} \frac{D}{Dx^\beta} V^\alpha,$$

or

$$\frac{\partial}{\partial y^\lambda} V^\mu + \Gamma^\mu{}_{\nu\lambda} V^\nu = \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial y^\mu}{\partial x^\alpha} \left[\frac{\partial}{\partial x^\beta} V^\alpha + \Gamma^\alpha{}_{\gamma\beta} V^\gamma \right]. \qquad (2.32)$$

We then compare with (2.30) and look for conditions on the object Γ. These conditions fix the behavior of Γ under coordinate transformations. Γ must transform according to

$$\Gamma^\mu{}_{\nu\lambda} = \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial x^\gamma}{\partial y^\nu} \left[\Gamma^\alpha{}_{\gamma\beta} + \frac{\partial x^\alpha}{\partial y^\rho} \frac{\partial^2 y^\rho}{\partial x^\beta \partial x^\gamma} \right],$$

or

$$\Gamma^\mu{}_{\nu\lambda} = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial x^\gamma}{\partial y^\nu} \Gamma^\alpha{}_{\gamma\beta} \frac{\partial x^\beta}{\partial y^\lambda} + \frac{\partial x^\beta}{\partial y^\lambda} \frac{\partial x^\gamma}{\partial y^\nu} \frac{\partial^2 y^\mu}{\partial x^\beta \partial x^\gamma}. \qquad (2.33)$$

This non–covariant behavior of the *connection* Γ makes of (2.31) a well–behaved, *covariant derivative*. We have used a vector field to find how Γ should behave but, once Γ is known, covariant derivatives can be defined on general tensors.

There are actually infinite objects satisfying conditions (2.33), that is, there are infinite connections. Take another one, $\Gamma'^{\mu}{}_{\nu\lambda}$. It is immediate that the difference $\Gamma'^{\mu}{}_{\nu\lambda} - \Gamma^{\mu}{}_{\nu\lambda}$ is a tensor under coordinate transformations.

The covariant derivative of a function (tensor of zero degree) is the usual derivative, which in the case is automatically covariant. Take a third order mixed tensor T. Its covariant derivative will be given by

$$D_\mu T^\nu{}_{\rho\sigma} = \partial_\mu T^\nu{}_{\rho\sigma} + \Gamma^\nu{}_{\lambda\mu} T^\lambda{}_{\rho\sigma} - \Gamma^\lambda{}_{\rho\mu} T^\nu{}_{\lambda\sigma} - \Gamma^\lambda{}_{\sigma\mu} T^\nu{}_{\rho\lambda}. \qquad (2.34)$$

The rules to calculate, involving terms with contractions for each original index, are fairly illustrated in this example. Notice the signs: positive for upper indices, negative for lower indices. The metric tensor, in particular, will have the covariant derivative

$$D_\mu g_{\rho\sigma} = \partial_\mu g_{\rho\sigma} - \Gamma^\lambda{}_{\rho\mu} g_{\lambda\sigma} - \Gamma^\lambda{}_{\sigma\mu} g_{\rho\lambda}. \qquad (2.35)$$

When the covariant derivative of T is zero on a domain, T is "self–parallel" on the domain, or *parallel–transported*. An intuitive view of this notion will be given soon (see below, Fig.2.1, page 51). It exactly translates to curved space the idea of a straight line as a curve with maintains its direction along all its length.

If the metric is parallel–transported, the equation above gives the *metricity condition*

$$\partial_\mu g_{\rho\sigma} = \Gamma^\lambda{}_{\rho\mu} g_{\lambda\sigma} + \Gamma^\lambda{}_{\sigma\mu} g_{\rho\lambda} = \Gamma_{\sigma\rho\mu} + \Gamma_{\rho\sigma\mu} = 2\,\Gamma_{(\rho\sigma)\mu}, \qquad (2.36)$$

where the symbol with lowered index is defined by $\Gamma_{\rho\sigma\mu} = g_{\rho\lambda}\Gamma^\lambda{}_{\sigma\mu}$ and the compact notation for the symmetrized part

$$\Gamma_{(\rho\sigma)\mu} = \tfrac{1}{2}\,\{\Gamma_{\rho\sigma\mu} + \Gamma_{\sigma\rho\mu}\}, \qquad (2.37)$$

has been introduced. The analogous notation for the antisymmetrized part

$$\Gamma_{[\rho\sigma]\mu} = \tfrac{1}{2}\,\{\Gamma_{\rho\sigma\mu} - \Gamma_{\sigma\rho\mu}\} \qquad (2.38)$$

is also very useful.

Another convenient notational device: it it usual to indicate the common derivative by a comma. We shall adopt also one of the two current notations

for the covariant derivative, the semi-colon. The metricity condition, for example, will have the expressions

$$g_{\rho\sigma;\mu} = g_{\rho\sigma,\mu} - 2\ \Gamma_{(\rho\sigma)\mu} = 0. \tag{2.39}$$

The bar notation for the covariant derivative, $g_{\rho\sigma|\mu} = g_{\rho\sigma;\mu}$, is also found in the literature.

2.4 The Levi–Civita Connection

There are, actually, infinite connections on a manifold, infinite objects behaving according to (2.33). And, given a metric, there are infinite connections satisfying the metricity condition. One of them, however, is special. It is given by

$$\overset{\circ}{\Gamma}{}^{\lambda}{}_{\rho\sigma} = \tfrac{1}{2}\ g^{\lambda\mu}\{\partial_{\rho}g_{\sigma\mu} + \partial_{\sigma}g_{\rho\mu} - \partial_{\mu}g_{\rho\sigma}\}\ . \tag{2.40}$$

It is the single connection satisfying metricity and which is symmetric in the last two indices. This symmetry has a deep meaning. The torsion of a connection of components $\Gamma^{\lambda}{}_{\rho\sigma}$ is a tensor T of components

$$T^{\lambda}{}_{\rho\sigma} = \Gamma^{\lambda}{}_{\sigma\rho} - \Gamma^{\lambda}{}_{\rho\sigma} = -\ 2\ \Gamma^{\lambda}{}_{[\rho\sigma]}. \tag{2.41}$$

Connection (2.40) is called the *Levi–Civita connection*. Its components are just the *Christoffel symbols* we have met in §1.15. It has, as said, a special relationship to the metric and is the only metric–preserving connection with zero torsion. Standard General Relativity works only with such connections. We can now give a clear image of parallel-transport. If the connection is related to a definite-positive metric — so that angles can be measured — a parallel-transported vector keeps the angle with the curve the same all along (Fig.2.1).

Notice that, with the notations introduced above, a general connection will have components of the form

$$\Gamma^{\lambda}{}_{\rho\sigma} = \Gamma^{\lambda}{}_{(\rho\sigma)} - 2\ T^{\lambda}{}_{\rho\sigma}. \tag{2.42}$$

Connections with zero torsion, $\Gamma^{\lambda}{}_{\rho\sigma} = \Gamma^{\lambda}{}_{(\rho\sigma)}$ are usually called *symmetric connections*.

We have said that a curve on a manifold M is a mapping from the real line on M. The mapping will be continuous, differentiable, etc, depending
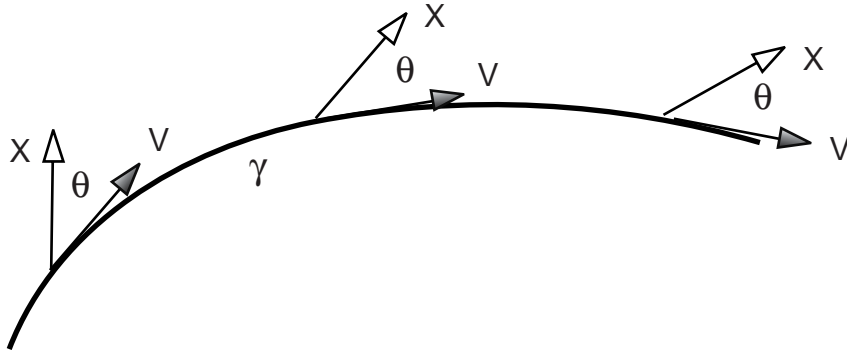
Figure 2.1: *If the connection is related to a definite-positive metric, a parallel-transported vector keeps the angle with the curve all along.*

on the point of interest. We shall be mostly interested in curves with fixed endpoints. A curve with initial endpoint p_0 and final endpoint p_1 is better defined as a mapping from the closed interval $I = [0, 1]$ into M:

$$\gamma : [0, 1] \to M \; ; \; u \to \gamma(u) \; ; \quad \gamma(0) = p_0 \; , \gamma(1) = p_1 \; . \qquad (2.43)$$

The variable $u \in [0, 1]$ is the curve parameter. A loop, or closed curve, with $p_0 = p_1$, can be alternatively defined as a mapping of the circle S^1 into M. We shall be primarily concerned with differentiable, or smooth curves, defined as those for which the above mapping is a differentiable function. A smooth curve has always a tangent vector at each point. Suppose the tangent vectors at all the points is time-like (see Eq. (2.24)). In that case the curve itself is said to be time-like. In the same way, the curve is said to be space-like or light-like when the tangent vectors are all respectively space-like and null.

We can now consider derivatives along a curve γ, whose points have coordinates $x^\alpha(u)$. The usual derivative along γ is defined as

$$\frac{d}{du} = \frac{dx^\alpha}{du} \frac{\partial}{\partial x^\alpha}$$

The covariant derivative along γ – also called absolute derivative — is defined by

$$\frac{D}{Du} = \frac{dx^\alpha}{du} D_\alpha \; , \qquad (2.44)$$

and apply to tensors just in the same way as covariant derivatives. Thus, for example,

$$\frac{DV^\alpha}{Du} = \frac{dV^\alpha}{du} + \Gamma^\alpha{}_{\gamma\beta} V^\gamma \frac{dx^\beta}{du}. \tag{2.45}$$

Let us go back to the invariant expression of a vector field, Eq.(2.29). It is an operator acting on the functions defined on the manifold. If the manifold is a differentiable manifold, and $\{x^\alpha\}$ is a coordinate system, the set of operators $\{\frac{\partial}{\partial x^\alpha}\}$ is linearly independent, and $\frac{\partial x^\beta}{\partial x^\alpha} = \delta^\beta_\alpha$. We say then that the set of derivative operators $\{\frac{\partial}{\partial x^\alpha}\}$ constitute a *base* for vector fields. Such a base is called natural, or coordinate base, as it is closely related to the coordinate system $\{x^\alpha\}$. Any vector field can be expressed as in Eq.(2.29), with components V^α. Under a change of coordinate system, $\frac{\partial}{\partial x^\alpha} = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial}{\partial y^\mu}$. The base member $e_\alpha = \frac{\partial}{\partial x^\alpha}$ is in this way expressed in terms of another base, the set $\{\frac{\partial}{\partial y^\mu}\}$. The latter constitutes another coordinate base, naturally related to the coordinate system $\{y^\mu\}$. The components V^α transform in the converse way, so that V is, as already said, an invariant object. Vector bases on 4–dimensional spacetime are usually called tetrads (also vierbeine, and four–legs).

Tetrad fields are actually much more general – they need not be related to any coordinate system. Put $h_\alpha{}^\mu = \frac{\partial y^\mu}{\partial x^\alpha}$, so that $e_\alpha = \frac{\partial}{\partial x^\alpha} = h_\alpha{}^\mu \frac{\partial}{\partial y^\mu}$. Of course, the members of the base $\{e_\alpha\}$ commute with each other, $[e_\alpha, e_\beta] = 0$. This is also the sufficient the condition for a base to be natural in some coordinate system: if $[e_\alpha, e_\beta] = 0$, then there exists a coordinate system $\{x^\alpha\}$ such that $e_\alpha = \frac{\partial}{\partial x^\alpha}$ for $\alpha = 0, 1, 2, 3$. These things are simpler in the so–called dual formalism, which uses differential forms.

As said in §2.20, to every real vector space $\mathcal{V}$ corresponds another vector space $\mathcal{V}^*$, which is its dual. This $\mathcal{V}^*$ is defined as the set of linear mappings from $\mathcal{V}$ into the real line $\mathbb{R}$. Given a base $\{e_\alpha\}$ of $\mathcal{V}$, there exists always a base $\{e^\alpha\}$ of $\mathcal{V}^*$ which is its dual, by which we mean that $e^\alpha(e_\beta) = \delta^\alpha_\beta$. The base dual to $\{\frac{\partial}{\partial x^\alpha}\}$ is the set of differential forms $\{dx^\alpha\}$, each dx^α being understood as a linear mapping satisfying $dx^\alpha(\frac{\partial}{\partial^\beta}) = \delta^\alpha_\beta$.

Take the differential of a function f: $df = \frac{\partial f}{\partial x^\alpha} dx^\alpha$. Applying the above rule, it follows that $df(\frac{\partial}{\partial x^\beta}) = \frac{\partial f}{\partial x^\beta}$. An arbitrary differential 1–form (also called a covector field) is written $\omega = \omega_\alpha dx^\alpha$ in base $\{dx^\alpha\}$. The above dual base will have members $e^\alpha = dx^\alpha = h^\alpha{}_\mu dy^\mu$, with $h^\alpha{}_\mu = \frac{\partial x^\alpha}{\partial y^\mu}$. As it always happen that $d^2 \equiv 0$, the condition for a covector base to be naturally related to a coordinate system is that $de^\alpha = 0$. Each dx^α transforms according to $dx^\alpha = \frac{\partial x^\alpha}{\partial y^\mu} dy^\mu$, and the above expression for a covector ω is invariant. We shall later examine general tetrads in detail.

dual
space
again

2.5 Curvature Tensor

§ **2.55** A connection defines covariant derivatives of general tensorial objects. It goes actually a little beyond tensors. A connection Γ defines a covariant derivative of itself. This gives, rather surprisingly, a tensor, the Riemann curvature tensor of the connection:

$$R^{\kappa}{}_{\lambda\rho\sigma} = \partial_\rho \Gamma^{\kappa}{}_{\lambda\sigma} - \partial_\sigma \Gamma^{\kappa}{}_{\lambda\rho} + \Gamma^{\kappa}{}_{\nu\rho}\Gamma^{\nu}{}_{\lambda\sigma} - \Gamma^{\kappa}{}_{\nu\sigma}\Gamma^{\nu}{}_{\lambda\rho} . \qquad (2.46)$$

It is important to notice the position of the indices in this definition. Authors differ in that point, and these differences can lead to differences in the signs (for example, in the scalar curvature defined below). We are using all along notations consistent with the differential forms. There is a clear antisymmetry in the last two indices,

$$R^{\kappa}{}_{\lambda\rho\sigma} = - R^{\kappa}{}_{\lambda\sigma\rho} .$$

§ **2.56** Notice that what exists is the curvature of a connection. Many connections are defined on a given space, each one with its curvature. It is common language to speak of "the curvature of space", but this only makes sense if a certain connection is assumed to be included in the very definition of that space.

§ **2.57** The above formulas hold on spaces of any dimension. The meaning of the curvature tensor can be understood from the diagram of Figure 2.2.

- First, build an infinitesimal parallelogram formed by pieces of geodesics, indicated by dx^λ, dx^ν, dx'^λ and dx'^ν.

- Take a vector field X with components X^μ at a corner point P. Parallel-transport X along dx^λ. At its extremity, X will have the value $X^\mu - \Gamma^{\mu}{}_{\nu\lambda}X^\nu dx^\lambda$

- Parallel-transport that value along dx'^ν. It will lead to a value which we denote X'^μ.

- Back at the starting point, parallel-transport X now first along dx^ν and then along dx'^λ. This will lead to a field value which we denote X''^μ.

- In a flat case, $X''^\mu = X'^\mu$. On a curved case, the difference between them is non-vanishing, and given by

$$\delta X^\mu = X''^\mu - X'^\mu = -\, R^\mu{}_{\rho\lambda\nu} X^\rho \, dx^\lambda dx^\nu \ . \qquad (2.47)$$

- This is the infinitesimal case, in the limit of vanishing parallelogram and with the value of $R^\mu{}_{\rho\lambda\nu}$ at the left-lower corner.
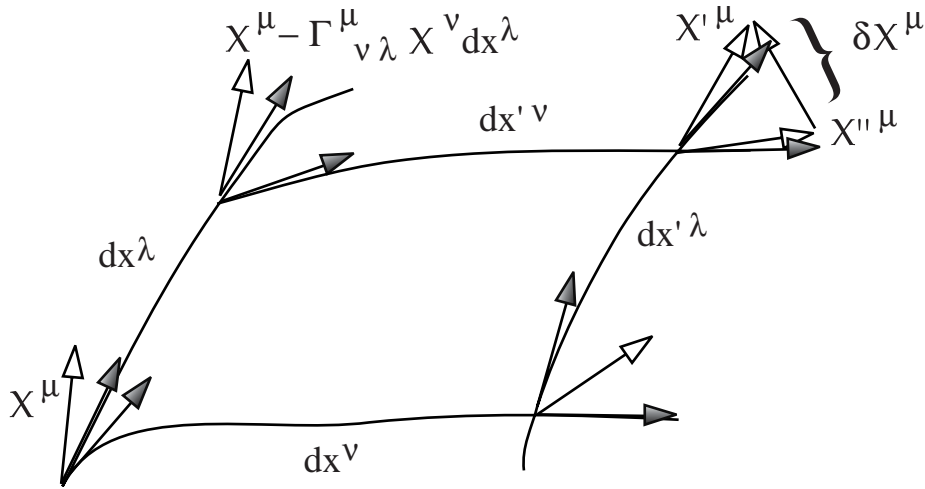
Figure 2.2: *The meaning of the Riemann tensor.*

§ **2.58** Other tensors can be obtained from the Riemann curvature tensor by contraction. The most important is the Ricci tensor

$$R_{\lambda\sigma} = R^\rho{}_{\lambda\rho\sigma} = \partial_\rho \Gamma^\rho{}_{\lambda\sigma} - \partial_\sigma \Gamma^\rho{}_{\lambda\rho} + \Gamma^\rho{}_{\nu\rho}\Gamma^\nu{}_{\lambda\sigma} - \Gamma^\rho{}_{\nu\sigma}\Gamma^\nu{}_{\lambda\rho} \ . \qquad (2.48)$$

This tensor is symmetric in the case of the Levi-Civita connection:

$$\mathring{R}_{\mu\nu} = \mathring{R}_{\nu\mu}. \qquad (2.49)$$

In this case, which has a special relation to the metric, the contraction with it gives the scalar curvature

$$\mathring{R} = g^{\mu\nu}\mathring{R}_{\mu\nu}. \qquad (2.50)$$

54

2.6 Bianchi Identities

§ 2.59 Take the definition (2.46) in the case of the Levi-Civita connection,

$$\mathring{R}^{\kappa}{}_{\lambda\rho\sigma} = \partial_\rho \mathring{\Gamma}^{\kappa}{}_{\lambda\sigma} - \partial_\sigma \mathring{\Gamma}^{\kappa}{}_{\lambda\rho} + \mathring{\Gamma}^{\kappa}{}_{\nu\rho}\mathring{\Gamma}^{\nu}{}_{\lambda\sigma} - \mathring{\Gamma}^{\kappa}{}_{\nu\sigma}\mathring{\Gamma}^{\nu}{}_{\lambda\rho} \; . \qquad (2.51)$$

The metric can be used to lower the index κ. Calculation shows that

$$\mathring{R}_{\kappa\lambda\rho\sigma} = g_{\kappa\mu}\mathring{R}^{\mu}{}_{\kappa\lambda\rho\sigma} = \partial_\rho \mathring{\Gamma}_{\kappa\lambda\sigma} - \partial_\sigma \mathring{\Gamma}_{\kappa\lambda\rho} + \mathring{\Gamma}_{\kappa\nu\rho}\mathring{\Gamma}^{\nu}{}_{\lambda\sigma} - \mathring{\Gamma}_{\kappa\nu\sigma}\mathring{\Gamma}^{\nu}{}_{\lambda\rho} \; , \qquad (2.52)$$

where $\mathring{\Gamma}_{\mu\rho\sigma} = g_{\mu\nu}\mathring{\Gamma}^{\nu}{}_{\rho\sigma}$. The curvature of a Levi-Civita connection has some special symmetries in the indices, which can be obtained from the detailed expression in terms of the metric:

$$\mathring{R}_{\kappa\lambda\rho\sigma} = -\mathring{R}_{\kappa\lambda\sigma\rho} = \mathring{R}_{\lambda\kappa\rho\sigma} = \mathring{R}_{\lambda\kappa\sigma\rho} \; ; \qquad (2.53)$$

$$\mathring{R}_{\kappa\lambda\rho\sigma} = \mathring{R}_{\rho\sigma\kappa\lambda} \; . \qquad (2.54)$$

In consequence of these symmetries, the Ricci tensor (2.48) is essentially the only contracted second-order tensor obtained from the Riemann tensor, and the scalar curvature (2.50) is essentially the only scalar.

§ 2.60 A detailed calculation gives the simplest way to exhibit curvature. Consider a vector field U with components U^α and take twice the covariant derivative, getting $U^{\alpha}{}_{;\beta;\gamma}$. Reverse then the order to obtain $U^{\alpha}{}_{;\gamma;\beta}$ and compare. The result is

$$U^{\alpha}{}_{;\beta;\gamma} - U^{\alpha}{}_{;\gamma;\beta} = - \mathring{R}^{\alpha}{}_{\epsilon\beta\gamma}U^\epsilon \; . \qquad (2.55)$$

Curvature turns up in the commutator of two covariant derivatives.

§ 2.61 Detailed calculations lead also to some identities. Of of them is

$$\mathring{R}^{\kappa}{}_{\lambda\rho\sigma} + \mathring{R}^{\kappa}{}_{\sigma\lambda\rho} + \mathring{R}^{\kappa}{}_{\rho\sigma\lambda} = 0 \; . \qquad (2.56)$$

Another is

$$\mathring{R}_{\kappa\lambda\rho\sigma;\mu} + \mathring{R}_{\kappa\lambda\mu\rho;\sigma} + \mathring{R}_{\kappa\lambda\sigma\mu;\rho} = 0 \qquad (2.57)$$

(notice, in both cases, the cyclic rotation of the three last indices). The last expression is called the Bianchi identity. As the metric has zero covariant derivative, it can be inserted in this identity to contract indices in a convenient way. Contracting with $g^{\kappa\rho}$, it comes out

$$\overset{\circ}{R}_{\lambda\sigma;\mu} - \overset{\circ}{R}_{\lambda\mu;\sigma} + \overset{\circ}{R}^{\rho}{}_{\lambda\sigma\mu;\rho} = 0.$$

Further contraction with $g^{\lambda\sigma}$ yields

$$\overset{\circ}{R}_{;\mu} - \overset{\circ}{R}^{\sigma}{}_{\mu;\sigma} - \overset{\circ}{R}^{\rho}{}_{\mu;\rho} = 0,$$

which is the same as

$$\overset{\circ}{R}_{;\mu} - 2\overset{\circ}{R}^{\sigma}{}_{\mu;\sigma} = 0$$

or

$$\left[\overset{\circ}{R}^{\mu}{}_{\nu} - \tfrac{1}{2}\,\delta^{\mu}_{\nu}\,\overset{\circ}{R}\right]_{;\mu} = 0. \tag{2.58}$$

This expression is the "contracted Bianchi identity". The tensor thus "covariantly conserved" will have an important role. Its totally covariant form,

$$G_{\mu\nu} = \overset{\circ}{R}_{\mu\nu} - \tfrac{1}{2}\,g_{\mu\nu}\,\overset{\circ}{R}\;, \tag{2.59}$$

is called the Einstein tensor. Its contraction with the metric gives the scalar curvature (up to a sign).

$$g^{\mu\nu}G_{\mu\nu} = -\,\overset{\circ}{R}\;. \tag{2.60}$$

§ **2.62** When the Ricci tensor is related to the metric tensor by

$$\overset{\circ}{R}_{\mu\nu} = \lambda\,g_{\mu\nu}\;, \tag{2.61}$$

where λ is a constant, It is usual to say that we have an *Einstein space*. In that case, $\overset{\circ}{R} = 4\,\lambda$ and $G_{\mu\nu} = -\,\lambda\,g_{\mu\nu}$. Spaces in which $\overset{\circ}{R}$ is a constant are said to be *spaces of constant curvature*. This is the standard language. We insist that there is no such a thing as the curvature of space. Curvature is a characteristic of a connection, and many connections are defined on a given space.

Some formulas above hold only in a four-dimensional space. We shall in the following give some lower-dimensional examples. It should be kept in mind that, in a d-dimensional space, $g^{\mu\nu}\,g_{\mu\nu} = d$. When $d = 2$ for example, $g^{\mu\nu}\,G_{\mu\nu} \equiv 0$. On a two-dimensional Einstein space, $G_{\mu\nu} \equiv 0$.

56

2.6.1 Examples

Calculations of the above objects will be necessary to write the dynamical equations of General Relativity. Such calculations are rather tiresome. In order to get a feeling, let us examine some low-dimensional examples. We shall look at the 2-dimensional sphere and at two 2-dimensional hyperboloids. In effect, two surfaces of revolution can be got from the hyperbola shown in Figure 2.3: one by rotating around the vertical axis, the other by rotating around the horizontal axis (Figure 2.4). The first will have two separate sheets, the other only one. They are obtained from each other by exchanging the squared vertical coordinate (below, z^2) by the squared coordinates of the resulting surface (below, $x^2 + y^2$). The spacetimes of de Sitter are higher-dimensional versions of these hyperbolic surfaces.
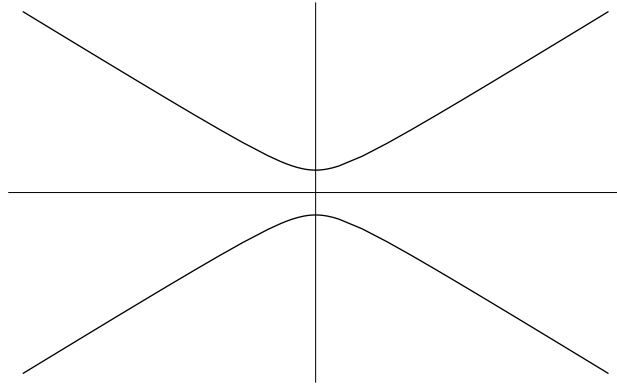
Figure 2.3: *Hyperbola. Two distinct surfaces of revolution can be obtained, by rotating either around the vertical or the horizontal axis.*

§ **2.63 The sphere** S^2 is defined as the set of points of $\mathbb{E}^3$ satisfying $x^2 + y^2 + z^2 = a^2$ in cartesian coordinates. This means $z = \pm\sqrt{a^2 - x^2 - y^2}$, and consequently

$$dz = -\frac{dx\,x}{\sqrt{a^2 - x^2 - y^2}} - \frac{dy\,y}{\sqrt{a^2 - x^2 - y^2}} \quad .$$
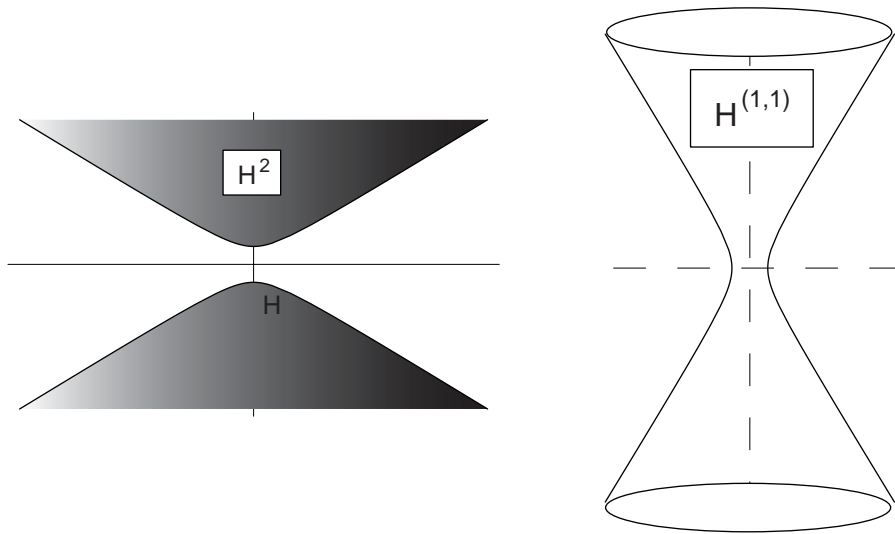
Figure 2.4: *The two hyperbolic surfaces. The second is the complement of the first, rotated of* 90°.

The interval $dl^2 = dx^2 + dy^2 + dz^2$ becomes then

$$dl^2 = \frac{a^2\,dx^2 + a^2\,dy^2 - dy^2\,x^2 + 2\,dx\,dy\,x\,y - dx^2\,y^2}{a^2 - x^2 - y^2}$$

It is convenient to change to spherical coordinates

$$x = a\,\sin\theta\,\cos\phi\,; \quad y = a\,\sin\theta\,\sin\phi\,; \quad z = a\cos\theta\,.$$

The interval becomes

$$dl^2 = a^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\,.$$

The corresponding metric is given by

$$g = (g_{\mu\nu}) = \begin{pmatrix} a^2 & 0 \\ 0 & a^2\,\sin^2\theta \end{pmatrix}\,,$$

with obvious inverse. The only non-vanishing Christoffel symbols are $\overset{\circ}{\Gamma}{}^1{}_{22} = -\sin\theta\cos\theta$ and $\overset{\circ}{\Gamma}{}^2{}_{12} = \overset{\circ}{\Gamma}{}^2{}_{21} = \cot\theta$. The only non-vanishing components of

58

the Riemann tensor are (up to symmetries in the indices) $\overset{\circ}{R}{}^1{}_{212} = \sin^2\theta$ and $\overset{\circ}{R}{}^2{}_{112} = -1$. The Ricci tensor has $\overset{\circ}{R}_{11} = 1$ and $\overset{\circ}{R}_{22} = \sin^2\theta$, so that it can be represented by the matrix

$$(\overset{\circ}{R}_{\mu\nu}) = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} = \tfrac{1}{a^2}\,(g_{\mu\nu})\ .$$

Consequently, the sphere is an Einstein space, with $\lambda = \tfrac{1}{a^2}$. Finally, the scalar curvature is

$$\overset{\circ}{R} = \frac{2}{a^2}\ .$$

Not surprisingly, a sphere is a space of constant curvature. As previously said, the Einstein tensor is of scarce interest for two-dimensional spaces. Though we shall not examine the geodesics, we note that their equations are

$$\ddot{\theta} = \sin\theta\cos\theta\,\dot{\phi}^2\ ;\quad \ddot{\phi} = -2\,\cot\theta\,\dot{\theta}\dot{\phi}\ ,$$

and that constant θ and ϕ are obvious particular solutions. Actually, the solutions are the great arcs.

§ 2.64 The **two-sheeted hyperboloid** H^2 is defined as the locus of those points of $\mathbb{E}^3$ satisfying $x^2 + y^2 - z^2 = -a^2$ in cartesian coordinates. The line element has the form

$$dx^2 + dy^2 - dz^2 = \frac{a^2\,dx^2 + a^2\,dy^2 - dy^2\,x^2 + 2\,dx\,dy\,x\,y - dx^2\,y^2}{a^2 - x^2 - y^2}\ .$$

Changing to coordinates

$$x = a\,\sinh\theta\,\cos\phi\ ;\quad y = a\,\sinh\theta\,\sin\phi\ ;\quad z = a\cosh\theta$$

the interval becomes

$$dl^2 = a^2\left(d\theta^2 + \sinh^2\theta\,d\phi^2\right)\ .$$

The metric will be

$$g = (g_{\mu\nu}) = \begin{pmatrix} a^2 & 0 \\ 0 & a^2\,\sinh^2\theta \end{pmatrix}\ .$$

59

The only non-vanishing Christoffel symbols are $\overset{\circ}{\Gamma}{}^1{}_{22} = -\sinh\theta\cosh\theta$ and $\overset{\circ}{\Gamma}{}^2{}_{12} = \overset{\circ}{\Gamma}{}^2{}_{21} = \coth\theta$. The only non-vanishing components of the Riemann tensor are (up to symmetries in the indices) $\overset{\circ}{R}{}^1{}_{212} = -\sinh^2\theta$ and $\overset{\circ}{R}{}^2{}_{112} = 1$. The Ricci tensor constitutes the matrix

$$(\overset{\circ}{R}_{\mu\nu}) = \begin{pmatrix} -1 & 0 \\ 0 & -\sinh^2\theta \end{pmatrix} .$$

We see that H^2 is an Einstein space, with $\lambda = -\frac{1}{a^2}$. The scalar curvature is

$$\overset{\circ}{R} = -\frac{2}{a^2} .$$

H^2 is a space of constant, but negative, curvature. It is similar to the sphere, but with a imaginary angle. An old name for it is "pseudo-sphere".

§ 2.65 The **one-sheeted hyperboloid** $H^{(1,1)}$ is defined as the locus of those points of $\mathbb{E}^3$ satisfying $x^2 + y^2 - z^2 = a^2$ in cartesian coordinates. The line element has the form

$$dz^2 - dx^2 - dy^2 = \frac{a^2\,dx^2 + a^2\,dy^2 - dy^2\,x^2 + 2\,dx\,dy\,x\,y - dx^2\,y^2}{a^2 - x^2 - y^2}.$$

Changing to coordinates

$$x = a\,\cosh\theta\,\cos\phi\,; \quad y = a\,\cosh\theta\,\sin\phi\,; \quad z = a\,\sinh\theta$$

the interval becomes

$$dl^2 = a^2\left(-d\theta^2 + \sinh^2\theta\,d\phi^2\right) .$$

The metric will be

$$g = (g_{\mu\nu}) = \begin{pmatrix} -a^2 & 0 \\ 0 & a^2\,\cosh^2\theta \end{pmatrix} .$$

The only non-vanishing Christoffel symbols are $\overset{\circ}{\Gamma}{}^1{}_{22} = \sinh\theta\cosh\theta$ and $\overset{\circ}{\Gamma}{}^2{}_{12} = \overset{\circ}{\Gamma}{}^2{}_{21} = \tanh\theta$. The only non-vanishing components of the Riemann tensor are (up to symmetries in the indices) $\overset{\circ}{R}{}^1{}_{212} = \cosh^2\theta$ and $\overset{\circ}{R}{}^2{}_{112} = 1$. The Ricci tensor constitutes the matrix

$$(\overset{\circ}{R}_{\mu\nu}) = \begin{pmatrix} -1 & 0 \\ 0 & \cosh^2\theta \end{pmatrix} .$$

We see that $H^{(1,1)}$ is an Einstein space, with $\lambda = \frac{1}{a^2}$. The scalar curvature is

$$\overset{\circ}{R} = \frac{2}{a^2} \ .$$

$H^{(1,1)}$ is a space of constant positive curvature, like the sphere.

§ 2.66 The rotating disk We shall now consider the rotating disk of §1.13 from two points of view: in $(2+1)$ dimensions (two for the rotation plane, one for time); and only the 3-dimensional space section. Curiously, the two cases will have quite different results: the $(2,1)$ case gives zero curvature, while the 3-dimensional space is curved.

Take first the $(2,1)$ case, with coordinates (ct, R, θ). The metric will be

$$g = (g_{\mu\nu}) = \begin{pmatrix} 1 - \frac{\omega^2 R^2}{c^2} & 0 & -\frac{\omega R^2}{c} \\ 0 & -1 & 0 \\ -\frac{\omega R^2}{c} & 0 & -R^2 \end{pmatrix} \ .$$

The only non-vanishing Christoffel symbols are $\overset{\circ}{\Gamma}{}^2{}_{11} = -\frac{\omega^2 R}{c^2}$, $\overset{\circ}{\Gamma}{}^2{}_{13} = \overset{\circ}{\Gamma}{}^2{}_{31} = -\frac{\omega R}{c}$, $\overset{\circ}{\Gamma}{}^2{}_{33} = -R$, $\overset{\circ}{\Gamma}{}^3{}_{12} = \overset{\circ}{\Gamma}{}^3{}_{21} = \frac{\omega}{cR}$, and $\overset{\circ}{\Gamma}{}^3{}_{23} = \overset{\circ}{\Gamma}{}^3{}_{32} = \frac{1}{R}$. All the components of the Riemman tensor vanish, so that the 3-dimensional spacetime is flat.

Take now the 3-dimensional space, with coordinates (x, y, z). The metric will be

$$g = (g_{\mu\nu}) = \begin{pmatrix} 1 + fy^2 & -fxy & 0 \\ -fxy & 1 + fx^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ ,$$

with $f = \frac{\omega^2 c^2}{c^2 - \omega^2 (x^2 + y^2)^2}$. All the Christoffel symbols of type $\overset{\circ}{\Gamma}{}^3{}_{ij}$ are zero, but the other have some rather lengthy expressions. The same is true of the Ricci tensor. We shall only quote the scalar curvature, which is (with $r^2 = x^2 + y^2$)

$$R = \frac{2\,c^4\,r^4\,\omega^6 - 2\,c^8\,\omega^2\,(3 + 2\,r^2\,\omega^2) + c^6\,(4\,r^2\,\omega^4 - 2\,r^4\,\omega^6)}{(c^2 - r^2\,\omega^2)^2\,(r^2\,\omega^2 - c^2\,(1 + r^2\,\omega^2))^2} \ .$$

This means that the space is curved. A simpler expression turns up for the particular values $\omega = 1$, $c = 1$:

$$R = -\frac{6}{1 - r^2} \ .$$

61

We see that there is a singularity when $r \to \omega c$. Exactly the same results come out if we consider the pure 2-dimensional case, the plane rotating disk. This corresponds to dropping the last column and row of the metric above.

§ **2.67** In all these examples,

1. a set point S is first defined by a constraint on the points of an ambient space E;

2. the metric is defined by the restriction, on the subset, of the metric of the ambient space;

3. such a metric is said to be *induced* by the imbedding of S in E.

Chapter 3

Dynamics

3.1 Geodesics

§ **3.1** Curves defined on a manifold provide tests for many of its properties. In Physics, they do still more: any observer will be ultimately represented by some special curve on spacetime. We have obtained the geodesic equation before. We had then used implicitly the assumption that $ds \neq 0$. The approach which follows* though more involved, has the advantage of including the null geodesics, for which $ds = 0$.

Our aim is to obtain certain priviledged curves as extremals of certain functionals, or functions defined on the space of curves. Such functional involve integrals along each curve, like

$$F[\gamma] = \int_{\gamma} f \ .$$

A simple stratagem allows to manipulate such functionals as if they were functions. To do it, it is necessary to give a label to each curve, so that varying the curve becomes a simple change of label. This can be done by considering, beyond the initial family of possible, curves, another family of "transversal" curves.

§ **3.2** We shall be interested in families of curves, like the curves α, β and γ in Figure 3.1. We have indicated by a_0 and a_1 the initial and the final

congru−
ences

---

* See J. L. Synge, *Relativity: The general theory*, North–Holland, Amsterdam, 1960.
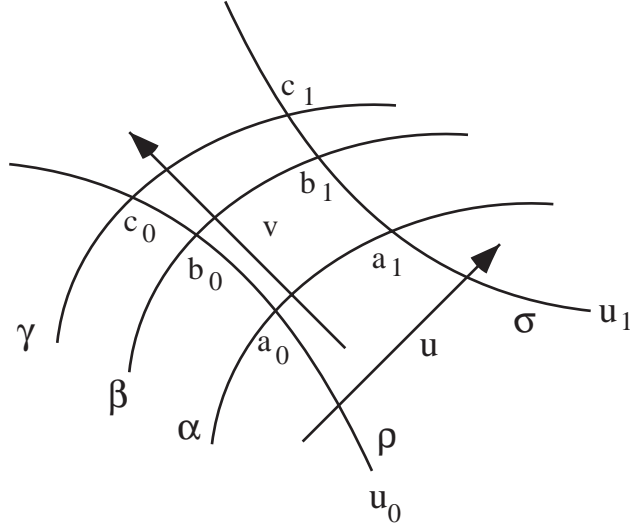
Figure 3.1: *The family of curves α, β, γ is parametrized by u. The crossed family of curves ρ, σ is parametrized by v. The second family is a "variation" of the first.*

endpoints of the piece of α which will be of interest, and analogously for the other curves. The curve parameter u is indicated as "going along" them, with u_0 and u_1 the initial and final endpoints. We say that the curves form a *congruence*. The variation leading from α to β and to γ can be represented by another parameter v. It is useful to consider v as the parameter related to another set of curves, such that each curve of the second congruence intersects every curve of the first, and vice versa. We have drawn only ρ and σ, which go through the endpoints of the first family of curves. The points on all the curves constitute a double continuum, a two–parameter domain which we shall indicate by $\gamma(u, v)$. It will have coordinates $x^\alpha(u, v) = \gamma^\alpha(u, v)$. For a fixed value of v, $\gamma_v(u) = \gamma(u, v = \text{constant})$ will describe a curve of the first family. For a fixed value of u, $\gamma_u(v) = \gamma(u = \text{constant}, v)$ will describe a curve of the second family. The second family is, by the way, called a "variation" of any member of the first family. There will be vectors along both families, such as

$$U^\alpha = \frac{dx^\alpha}{du} \quad \text{and} \quad V^\alpha = \frac{dx^\alpha}{dv}.$$

If the connection is symmetric, the crossed absolute derivatives coincide:

$$\frac{DV^\alpha}{Du} = \frac{DU^\alpha}{Dv} \; . \tag{3.1}$$

Let us first fix v at some value, thereby choosing a fixed curve of the first family and consider, along that curve, the function(al)

$$\mathcal{I}[v] = \tfrac{1}{2} \, (u_1 - u_0) \, \int_{u_0}^{u_1} g_{\alpha\beta} \frac{dx^\alpha}{du} \frac{dx^\beta}{du} \; du = \tfrac{1}{2} \, (u_1 - u_0) \, \int_{u_0}^{u_1} g_{\alpha\beta} \, U^\alpha \, U^\beta \; du \; . \tag{3.2}$$

On the two-parameter domain, variations of that curve become simple derivatives with respect to v. By conveniently adding antisymmetric and symmetric pieces, we have the series of steps

$$\frac{d}{dv}\mathcal{I}[v] = (u_1 - u_0) \, \int_{u_0}^{u_1} g_{\alpha\beta} U^\alpha \frac{D}{Dv} U^\beta \; du = (u_1 - u_0) \, \int_{u_0}^{u_1} g_{\alpha\beta} U^\alpha \frac{DV^\beta}{Du} \; du$$

$$= (u_1 - u_0) \, \int_{u_0}^{u_1} \frac{d}{du} \left[g_{\alpha\beta} U^\alpha \, V^\beta \right] \; du - (u_1 - u_0) \, \int_{u_0}^{u_1} g_{\alpha\beta} \, V^\beta \frac{D}{Du} U^\alpha du$$

$$= (u_1 - u_0) \, \left[g_{\alpha\beta} U^\alpha \, V^\beta \right]_{u_0}^{u_1} \; du - (u_1 - u_0) \, \int_{u_0}^{u_1} g_{\alpha\beta} \, V^\beta \frac{D}{Du} U^\alpha du \tag{3.3}$$

We are interested in curves with the same endpoints, and shall now collapse the parameter v. In Figure 3.1, $a_0 = b_0 = c_0$ and $a_1 = b_1 = c_1$. This corresponds to putting $V^\beta = 0$ at the endpoints. The first contribution in the right–hand side above vanishes and

$$\frac{d}{dv} \, \mathcal{I}[v] = -(u_1 - u_0) \, \int_{u_0}^{u_1} g_{\alpha\beta} \, V^\beta \frac{D}{Du} U^\alpha du \; . \tag{3.4}$$

We now call geodesics the curves with fixed endpoints for which $\mathcal{I}$ is stationary. As in the last integrand V^β is arbitrary except at the endpoints, such curves must satisfy

$$\frac{D}{Du} \, U^\alpha = \frac{dU^\alpha}{du} + \overset{\circ}{\Gamma}{}^\alpha{}_{\gamma\beta} U^\gamma U^\beta = 0 \; , \tag{3.5}$$

which is the same as

$$\boxed{\begin{array}{l} \textbf{geodesic} \\ \textbf{equation} \end{array}}$$

$$\frac{d^2 x^\alpha}{du^2} + \overset{\circ}{\Gamma}{}^\alpha{}_{\gamma\beta} \frac{dx^\beta}{du} \frac{dx^\gamma}{du} = 0. \tag{3.6}$$

This is the geodesic equation we have met before. It says that the vector $U^\alpha = \frac{dx^\alpha}{du}$ has vanishing absolute derivative along the curve. This means that it is parallel-transported along the curve. The only direction that can be attributed to a curve is, at each point, that of its "velocity" U^α. This leads to a much better name for such a curve: *self-parallel curve*. It keeps the same direction all along. Geodesics (solutions of the geodesic equation) play on curved spaces the role the straight lines have on flat spaces.

Comment 3.1 Why is the name "self-parallel" better ? There are at least two reasons:

1. the word "geodesic" has a very strong metric conotation; its original meaning was that of a shortest length curve; but length, real length, only is defined by a positive-definite metric, not our case;

2. the concept has actually no relation to metric at all; such curves can be defined for any connection; connection is a concept quite independent of metric, and defines parallelism; for general connections, only "self-parallel" makes sense.

§ 3.3 The geodesic equation has the first integral

$$g_{\alpha\beta}\, U^\alpha\, U^\beta = C. \tag{3.7}$$

This is to say that, along the curve,

$$\frac{d}{du}\left[g_{\alpha\beta}\, U^\alpha\, U^\beta\right] = 0.$$

Comment 3.2 Prove it for the connection given by (2.40).

§ 3.4 Comparison with the interval expression (2.25) shows that

$$ds = C^{1/2}du.$$

Eq.(3.6) is invariant under parameter changes of type

$$u \to u' = au + b\ . \tag{3.8}$$

The constant C can, consequently, be rescaled to have only the values 1 and 0. This means that, unless $C = 0$, the interval parameter s can be used as the curve parameter. That parameter is the proper time, and we recall that

$U^\alpha = \frac{dx^\alpha}{ds}$ is the the four–velocity. The choice of $C = 1$ allows contact to be made with Special Relativity, as (3.7) becomes

$$U^2 = U \cdot U = U_\alpha U^\alpha = g_{\alpha\beta} U^\alpha \, U^\beta = 1. \tag{3.9}$$

The case $C = 0$ includes trajectories of particles with vanishing masses, in special light–rays. As long as we keep in mind this exception, we can rewrite the geodesic equation in the forms

$$\frac{D}{Ds} \, U^\alpha = \frac{dU^\alpha}{ds} + \Gamma^\alpha{}_{\beta\gamma} U^\beta U^\gamma = \frac{d^2 x^\alpha}{ds^2} + \Gamma^\alpha{}_{\beta\gamma} \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} = 0. \tag{3.10}$$

Once things have been interpreted in this way, we say that the velocity is covariantly derived along γ. This is supported by Eq.(2.44), which now shows the absolute derivative as the covariant derivative projected, at each point, on the velocity.

§ **3.5** In the case of massive particles, we can use the freedom allowed by (3.8) to choose another parameter. Introduce s by

$$u - u_0 = (u_1 - u_0)^{1/2} s/L \ .$$

Then, comparison with(2.26) shows that $\mathcal{I}(v) = \frac{1}{2} L^2$. This leads to the variational principle

$$\delta L = \delta \int ds = 0 \ , \tag{3.11}$$

which we have found before.

§ **3.6** If we look back at what has been done in §1.15, we see that we have there got (3.10) from (3.11). A massive particle, without additional structure (for instance, supposing that the effect of its spin is negligible, or zero) will follow the geodesic equation. That is actually the standard approach. It is enough to replace, in all the discussion, the expression "in the presence of a metric" by the expression "in the presence of a gravitational field". It holds, wee see now, for massive particles, for which $ds \neq 0$.

§ 3.7 Eqs.(3.6) and (3.10) are second–order ordinary differential equations. Existence and unicity theorems of the theory of differential equations state that, given starting coordinates x^i and "velocities" $\frac{dx^i}{du}$ at a point P, there will be a curve $\gamma(u)$, with $-\epsilon < u < \epsilon$ for some $\epsilon > 0$, which goes through P ($\gamma(P) = 0$) and which is unique.

§ 3.8 The expression

$$A^\alpha = \frac{dU^\alpha}{ds} + \Gamma^\alpha{}_{\beta\gamma}U^\beta U^\gamma \tag{3.12}$$

is the covariant acceleration, the only to have a meaning in an arbitrary coordinate system. The existence of the above first integral is equivalent to

$$U \cdot A = g_{\alpha\beta}\, U^\alpha A^\beta = 0 \ . \tag{3.13}$$

As in Special Relativity the acceleration is, at each point of γ, orthogonal to the velocity. As the velocity is parallel (or tangent) to γ at each point, we can say that the acceleration is orthogonal to γ. A curve is self–parallel when its acceleration vanishes.

§ 3.9 We have above supposed a Levi–Civita connection. All the terminology comes from the first historical case, the Levi–Civita connection of a strictly Riemannian metric. We have said that any vector which is parallel–transported along a curve keeps the same angle with respect to the velocity all along it. The concept of self–parallel curve keeps its meaning for a general linear connection. A self–parallel curve is in that case *defined* as a curve satisfying (3.10).

§ 3.10 Let us go back to the first integral given in Eq.(3.7). It has a very deep meaning. As the momentum of a particle of mass m is $P^\mu = mc\, U^\mu$, we rewrite it as

$$m^2c^2\, g_{\mu\nu}\, U^\mu\, U^\nu = g_{\mu\nu}\, P^\mu\, P^\nu = m^2c^2\, C \ . \tag{3.14}$$

Write now

$$mc\, g_{\mu\nu}\, U^\nu = \partial_\mu\, S. \tag{3.15}$$

This means $dS = mc \, g_{\mu\nu} \, U^\nu dx^\mu$. This gradient form, applied to the tangent vector $U = U^\lambda \frac{\partial}{\partial x^\lambda}$, gives $dS(U) = mc \, C$. That reveals the meaning of S for the present case of a massive particle: with the choice $C = 1$ announced in § 3.4, it is the action written in terms of the sole coordinates along the curve, as it appears in Hamilton-Jacobi theory:

$$P_\mu = \partial_\mu \, S \, . \tag{3.16}$$

The geodesic curve is, at each point, orthogonal to the surface $S = \text{con-}$stant passing through that point. In effect, $dS = mc \, g_{\mu\nu} \, U^\nu dx^\mu$ is (up to the sign) just the action we have used before, but here along the curve. The expression

<div style="float:right; border:1px solid black; padding:4px; text-align:center">
Hamilton Jacobi equation
</div>

$$g_{\mu\nu} \, \partial^\mu S \, \partial^\nu S = m^2 c^2 \tag{3.17}$$

is the relativistic *Hamilton-Jacobi* equation for the free particle. It is possible to recover the geodesic equation from it. Let us see how.

Consider both sides of the expression

$$\frac{d}{du} [g_{\mu\nu} \, U^\nu] = \frac{d}{du} U_\mu = U^\rho \partial_\rho U_\mu \, .$$

The left-hand side (LHS) is

$$\frac{d}{du} (g_{\mu\nu} \, U^\nu) = g_{\mu\nu} \frac{d}{du} U^\nu + U^\nu \left(\frac{d}{du} g_{\mu\nu} \right) = g_{\mu\nu} \frac{d}{du} U^\nu + U^\nu U^\rho \partial_\rho g_{\mu\nu} \, .$$

The last piece is symmetric in the indices ν and ρ, so that we can write

$$LHS = \frac{d}{du} (g_{\mu\nu} \, U^\nu) = g_{\mu\nu} \frac{d}{du} U^\nu + \tfrac{1}{2} \, U^\nu U^\rho (\partial_\rho g_{\mu\nu} + \partial_\nu g_{\mu\rho}) \, .$$

Now to the right-hand side:

$$RHS = U^\rho \partial_\rho U_\mu = U^\rho \partial_\rho \partial_\mu S = U^\rho \partial_\mu \partial_\rho S = U^\rho \partial_\mu U_\rho = - \, U_\rho \partial_\mu U^\rho \, ,$$

using $\partial_\mu [g_{\rho\nu} U^\nu U^\rho] = \partial_\mu [\text{constant}] = 0$ in the last step. Notice that here we have the only contribution of Hamilton-Jacobi theory: we have used the fact (3.16) that the momentum is a derivative to justify the exchange $\partial_\rho U_\mu = \partial_\mu U_\rho$. Now, $\partial_\mu [g_{\rho\nu} U^\nu U^\rho] = 0$ is also

$$0 = (\partial_\mu g_{\rho\nu}) \, U^\nu U^\rho + 2 \, g_{\rho\nu} U^\nu \partial_\mu U^\rho \quad \therefore \quad U_\rho \partial_\mu U^\rho = - \, \tfrac{1}{2} \, (\partial_\mu g_{\rho\nu}) \, U^\nu U^\rho \, ,$$

69

so that

$$RHS = -\,U_\rho \partial_\mu U^\rho = \tfrac{1}{2}\left(\partial_\mu g_{\rho\nu}\right) U^\nu U^\rho \ .$$

Now, LHS = RHS gives

$$g_{\mu\nu}\,\frac{d}{du}U^\nu + \tfrac{1}{2}\ \left(\partial_\rho g_{\mu\nu} + \partial_\nu g_{\mu\rho} - \partial_\mu g_{\rho\nu}\right) U^\rho U^\nu = 0\ ,$$

or

$$\frac{d}{du}U^\lambda + \tfrac{1}{2}g^{\lambda\mu}\left[\partial_\tau g_{\mu\sigma}\ + \partial_\sigma g_{\mu\tau} - \partial_\mu g_{\tau\sigma}\right] U^\sigma U^\tau = 0\ ,$$

the announced result. All this can be repeated step by step, but starting from

$$g_{\mu\nu}\ \partial^\mu S\ \partial^\nu S = 0 \ . \tag{3.18}$$

This equation has a different meaning: it is the *eikonal equation*, with S now in the role of the eikonal. The geodesic, in this case, is the light-ray equation.

The Hamilton-Jacobi and the eikonal equations allow thus a unified view of particle trajectories and light rays. The geodesic equation comes

- from the Hamilton-Jacobi equation, as the equation of motion of a massive particle;

- from the eikonal equation, as the trajectory of a light ray.

§ 3.11 As we have said, curves are of fundamental importance. They not only allow testing many properties of a given space. In spacetime, every (ideal) observer is ultimately a time–like curve.

The nub of the equivalence principle is the concept of observer:

An observer is a timelike curve on spacetime, a world–line.

Such a curve represents a point-like object in 3-space, evolving in the time-like 4-th "direction". An object extended in 3-space would be necessarily represented by a bunch of world–lines, one for each one of its points. This mesh of curves will be necessary if, for example, the observer wishes to do some experiment. For the time being, let us take the simplifying assumption above, and consider only one world–line. This is an ideal, point-like observer. If free from external forces, this line will be a geodesic.

And here comes the crucial point. Given a geodesic γ going through a point P ($\gamma(0) = P$), there is always a very special system of coordinates (Riemannian normal coordinates) in a neighborhood U of P in which the components of the Levi-Civita connection vanish at P. The geodesic is, in this system, a straight line: $y^a = c^a s$. This means that, as long as γ traverses U, the observer will not feel gravitation: the geodesic equation reduces to the forceless equation $\frac{du^a}{ds} = \frac{d^2 y^a}{ds^2} = 0$. This is an inertial observer in the absence of external forces. If $\Gamma = 0$, covariant derivatives reduce to usual derivatives. If external forces are present, they will have the same expressions they have in Special Relativity. Thus, the inertial observer will see the force equation $\frac{du^a}{ds} = F^a$ of Special Relativity (see Section 3.7).

§ **3.12** There is actually more. Given any curve γ, it is possible to find a local frame in which the components of the Levi-Civita connection vanish along γ. That observer would not feel the presence of gravitation.

§ **3.13** How point–like is a real observer ? We are used to say that an observer can always know whether he/she is accelerated or not, by making experiments with accelerometers and gyroscopes. The point is that all such apparatuses are extended objects. We shall see later that a gravitational field is actually represented by curvature and that two geodesics are enough to denounce its presence (§ 3.46).

§ **3.14** As repeatedly said, the principle of equivalence is a heuristic guiding precept. It states that, as long as the dimensions involved in the definition of an observer are negligible, an observer can choose his/hers coordinates so that everything (s)he experiences is described by the laws of Special Relativity.

3.2   The Minimal Coupling Prescription

§ **3.15** The equivalence principle has been used up to now in a one-way trip, from General Relativity to Special Relativity. Some frame exists in which the connection vanishes, so that covariant derivatives reduce to simple derivatives. This can be used in the opposite sense. Given a special-relativistic expression, how to get its version in the presence of a gravitational field ?

The answer is now very simple: replace common derivatives on any tensorial object by covariant derivatives. Symbolically, this is represented by a rule,

$$\partial_\mu \Rightarrow \partial_\mu + \Gamma_\mu. \tag{3.19}$$

This comma $\Rightarrow$ semi-colon rule is the minimal coupling prescription. To this $\boxed{\text{rules}}$ must be added the already discussed passage from flat to curved metric,

$$\eta_{ab} \Rightarrow g_{\mu\nu}. \tag{3.20}$$

Once acccepted, these two rules allow to translate special-relativistic laws, or equations, into expressions which hold in the presence of a gravitational field.

§ **3.16 Conservation of energy** is one of the most important laws of Physics. Its special-relativistic version states that the energy-momentum tensor has vanishing divergence:

$$\partial_\nu T^{\mu\nu} = T^{\mu\nu}{}_{,\nu} = 0. \tag{3.21}$$

In the presence of a gravitational field, this becomes

$$\partial_\nu T^{\mu\nu} + \Gamma^\mu{}_{\rho\nu} T^{\rho\nu} + \Gamma^\nu{}_{\rho\nu} T^{\mu\rho} = 0, \tag{3.22}$$

being understood that every other derivative and any metric factor appearing in $T^{\mu\nu}$ are equally replaced according to (3.19) and (3.20). Notice that, because of the metricity condition (2.39), the metric can be inserted in or extracted from covariant derivatives at will. Equation (3.22) is usually represented in the comma-notation, as

$$T^{\mu\nu}{}_{;\nu} = 0 \ . \tag{3.23}$$

§ **3.17** An exercise: A **dust cloud** (or incoherent fluid) is a fluid formed by massive particles ignoring each other. It is a gas without pressure — only $\boxed{\substack{\text{dust}\\\text{cloud}}}$ energy is present. Special Relativity gives its energy–momentum the form

$$T^{\mu\nu} = \epsilon \, U^\mu U^\nu, \tag{3.24}$$

where ϵ is the energy density. The 4–vector U is a field representing the velocities of the fluid stream–lines.

Comment 3.3 The following results are immediate:

$$T^{\mu\nu}U_\nu = \epsilon\, U^\mu \; ; \quad T^{\mu\nu}U_\nu U_\mu = \epsilon \; ; \quad g_{\mu\nu}T^{\mu\nu} = \epsilon \, .$$

The covariant divergence of $T^{\mu\nu}$ must vanish:

$$D_\nu T^{\mu\nu} = U^\mu\, D_\nu(\epsilon\, U^\nu) + \epsilon\, U^\nu D_\nu U^\mu = U^\mu\, D_\nu(\epsilon\, U^\nu) + \epsilon\, \frac{D}{Ds}\, U^\mu = 0.$$
$$(3.25)$$

Contract this expression with U_μ: as $U_\mu U^\mu = 1$,

$$D_\nu(\epsilon U^\nu) + \epsilon\, U_\mu \frac{D}{Ds} U^\mu = 0.$$

As the metric can be inserted into or extracted from the covatiant derivative without any modification (because it is preserved by the Levi-Civita connection), it follows that $U_\mu \frac{D}{Ds} U^\mu = 0$. We get in consequence a continuity equation, or energy flux "conservation":

$$D_\nu(\epsilon U^\nu) = 0.$$

Taken into (3.25), this leads to the geodesic equation for the stream–lines:

$$\frac{D}{Ds}\, U^\mu = 0 \, .$$

§ 3.18 The covariant derivative $D_\mu\phi$ of a scalar field ϕ is just the usual derivative, $D_\mu\phi = \partial_\mu\phi$. But the derivative $\partial_\mu\phi$ is by itself a vector. The Laplacian operator, or (its usual name in 4 dimensions) D'Alembertian, is $D_\mu\partial^\mu\phi = \partial_\mu\partial^\mu\phi + \Gamma^\mu{}_{\rho\mu}\partial^\rho\phi$. The contracted form $\Gamma^\mu{}_{\rho\mu}$ of the Levi-Civita connection has a special expression in terms of the metric. From Eq. (2.40),

$$\Gamma^\mu{}_{\rho\mu} = \tfrac{1}{2}\, g^{\mu\nu}\{\partial_\rho g_{\mu\nu} + \partial_\mu g_{\rho\nu} - \partial_\nu g_{\mu\rho}\} = \tfrac{1}{2}\, g^{\mu\nu}\partial_\rho g_{\mu\nu} = \tfrac{1}{2}\, tr[\mathbf{g}^{-1}\partial_\rho\mathbf{g}],$$

where the matrix $\mathbf{g} = (g_{\mu\nu})$ and its inverse have been introduced. This is

$$\Gamma^\mu{}_{\rho\mu} = \tfrac{1}{2}\, tr[\partial_\rho \ln\mathbf{g}] = \tfrac{1}{2}\, \partial_\rho\, tr[\ln\mathbf{g}].$$

For any matrix $\mathbf{M}$, $tr[\ln\mathbf{M}] = \ln\det\mathbf{M}$, so that

$$\Gamma^\mu{}_{\rho\mu} = \tfrac{1}{2}\, \partial_\rho\, \ln[\det\mathbf{g}] = \partial_\rho\, \ln[-g]^{1/2} = \tfrac{1}{\sqrt{-g}}\, \partial_\rho\, \sqrt{-g}, \qquad (3.26)$$

where $g = |\det \mathbf{g}|$. The D'Alembertian becomes

$$\Box\, \phi = \partial_\mu \partial^\mu \phi + \frac{1}{\sqrt{-g}}\,[\partial_\mu\,\sqrt{-g}]\partial^\mu\,\phi,$$

or

$$\Box\, \phi = \frac{1}{\sqrt{-g}}\,\partial_\mu\,[\sqrt{-g}\,\partial^\mu\,\phi]. \qquad (3.27)$$

Laplaceans on curved spaces are known since long, and are called Laplace-Beltrami operators.

§ **3.19** Another example: the **electromagnetic field**. To begin with, we notice that, due to (3.18) and using (3.26),

$$\partial_\mu A^\mu \Rightarrow A^\mu{}_{;\mu} = \frac{1}{\sqrt{-g}}\,\partial_\mu\,[\sqrt{-g}\,A^\mu]\,. \qquad (3.28)$$

This is, by the way, the general form of the covariant divergence of a four-vector. The field strength $F_{\mu\nu}$ is antisymmetric and, due to the symmetry of the Christofell symbols, remain the same:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [\overset{\circ}{\Gamma}{}^\rho{}_{\mu\nu} - \overset{\circ}{\Gamma}{}^\rho{}_{\nu\mu}]A_\rho = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

In consequence, only the metric changes in the Lagrangean $L_{em} = -\frac{1}{4}F_{ab}F^{ab}$ $= -\frac{1}{4}\eta^{ac}\eta^{bd}F_{ab}F_{cd}$ of Special Relativity: it becomes

$$L_{em} = -\tfrac{1}{4}g^{\mu\rho}\eta^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma}. \qquad (3.29)$$

Maxwell's equations, which in Special Relativity are written

$$\partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} + \partial_\mu F_{\nu\lambda} = F_{\mu\nu,\lambda} + F_{\lambda\mu,\nu} + F_{\nu\lambda,\mu} = 0 \qquad (3.30)$$

$$\partial_\mu F^{\mu\nu} = F^{\mu\nu}{}_{,\mu} = j^\nu, \qquad (3.31)$$

take the form

$$F_{\mu\nu;\lambda} + F_{\lambda\mu;\nu} + F_{\nu\lambda;\mu} = 0 \qquad (3.32)$$

$$F^{\mu\nu}{}_{;\mu} = j^\nu. \qquad (3.33)$$

74

The latter deserves to be seen in some detail. Let us first examine the derivative

$$F^{\mu\nu}{}_{;\mu} = \partial_\mu F^{\mu\nu} + \mathring{\Gamma}^\mu{}_{\rho\mu} F^{\rho\nu} + \mathring{\Gamma}^\nu{}_{\rho\mu} F^{\mu\rho}.$$

The last term vanishes, again because $\mathring{\Gamma}^\nu{}_{\rho\mu}$ is symmetric. If another connection were at work, a coupling with torsion would turn up: the last term would be $(-\frac{1}{2} T^\nu{}_{\rho\mu} F^{\mu\rho})$. The first two terms give Maxwell's equations in the form

$$F^{\mu\nu}{}_{;\mu} = \frac{1}{\sqrt{-g}} \partial_\mu \left[\sqrt{-g} \ F^{\mu\nu} \right] = j^\nu. \tag{3.34}$$

The energy-momentum tensor of the electromagnetic field, which is

$$T_{ab} = -\eta^{cd} F_{ac} F_{bd} + \tfrac{1}{4}\eta_{ab}\eta^{ec}\eta^{fd} F_{ef} F_{cd}$$

in Special Relativity, becomes

$$T_{\mu\nu} = -g^{\rho\sigma} F_{\mu\rho} F_{\nu\sigma} + \tfrac{1}{4}g_{\mu\nu}g^{\kappa\rho}g^{\lambda\sigma} F_{\kappa\lambda} F_{\rho\sigma} \ . \tag{3.35}$$

Comment 3.4 Notice

$$g_{\mu\nu}T^{\mu\nu} = 0.$$

§ 3.20 A fluid of pressure p and energy density ϵ has an energy-momentum tensor generalizing (3.24):

$$T^{ab} = (p + \epsilon) \ U^a U^b - p\eta^{ab}.$$

This changes in a subtle way. Its form is quite analogous,

energy
momentum

$$T^{\mu\nu} = (p + \epsilon) \ U^\mu U^\nu - pg^{\mu\nu} \ , \tag{3.36}$$

but the 4-velocities are $U^\mu = \frac{dx^\mu}{ds}$, with ds the modified interval.

Comment 3.5 With respect to the dust cloud, only the trace changes:

$$T^{\mu\nu}U_\nu = \epsilon \ U^\mu \ ; \ \ T^{\mu\nu}U_\nu U_\mu = \epsilon \ ; \ \ T = g_{\mu\nu}T^{\mu\nu} = \epsilon - 3 \ p \ .$$

If the gas is ultrarelativistic, the equation of state is $\epsilon = 3 \ p$ and, consequently, $T = 0$, as for the electromagnetic field.

§ 3.21 We have seen in § 3.17 that the streamlines in a dust cloud are geodesics. The energy-momentum density (3.36) differs from that of dust by the presence of pressure. We can repeat the procedure of that paragraph in order to examine its effect. Here, $T^{\mu\nu}{}_{;\nu} = 0$ leads to

$$T^{\mu\nu}{}_{;\nu} = U^\mu \frac{D}{Ds}(p+\epsilon) + (p+\epsilon)\frac{DU^\mu}{Ds} + (p+\epsilon)U^\mu U^\nu{}_{;\nu} - \partial^\mu p = 0.$$

Contraction with U_μ leads now to

$$(\epsilon U^\nu)_{;\nu} = -\ p\ U^\nu{}_{;\nu}\ ,$$

so that the energy flux is no more conserved. Taking this expression back into that of $T^{\mu\nu}{}_{;\nu} = 0$ implies

$$(p+\epsilon)\ \frac{DU^\mu}{Ds} = \partial^\mu p -\ U^\mu \frac{Dp}{Ds} = (g^{\mu\nu} - U^\mu U^\nu)\partial_\nu p\ . \qquad (3.37)$$

More will be said on this force equation in § 3.50.

§ 3.22 Suppose that $T^{\mu\nu}$ is symmetric, as in the examples above. Define the quantity

$$L^{\mu\nu\lambda} = x^\lambda T^{\mu\nu} - x^\nu T^{\mu\lambda}.$$

It is immediate that

$$L^{\mu\nu\lambda}{}_{;\mu} = 0$$

and

$$U_\mu\ L^{\mu\nu\lambda} = \epsilon(x^\lambda U^\nu - x^\nu U^\lambda).$$

3.3   Einstein's Field Equations

§ 3.23 The Einstein tensor (2.59) is a purely geometrical second-order tensor which has vanishing covariant derivative. It is actually possible to prove that it is the only one. The energy-momentum tensor is a physical object with the same property. The next stroke of genius comes here. Einstein was convinced that some physical chacteristic of the sources of a gravitational field should engender the deformation in spacetime, that is, in its geometry. He looked for a dynamical equation which gave, in the non-relativistic, classical limit,

the newtonian theory. This means that he had to generalize the Poisson equation

$$\Delta V = 4\pi G\rho \qquad (3.38)$$

within riemannian geometry. The $G_{\mu\nu}$ has second derivatives of the metric, and the energy-momentum tensor contains, as one of its components, the energy density.

He took then the bold step of equating them to each other, obtaining what we know nowadays to be the simplest possible generalization of the Poisson equation in a riemannian context:

<div style="float:right; border:1px solid black; padding:4px; text-align:center;">**field equation**</div>

$$R_{\mu\nu} - \tfrac{1}{2}\, g_{\mu\nu} R = \tfrac{8\pi G}{c^4}\, T_{\mu\nu}. \qquad (3.39)$$

This is the *Einstein equation*, which fixes the dynamics of a gravitational field. The constant in the right-hand side was at first unknown, but he fixed it when he obtained, in the due limit, the Poisson equation of the newtonian theory (as will be seen in §3.31 below).

Comment 3.6 A text on gravitation has always a place for the value of G. At present time, the best experimental value for the gravitational constant is

$$G = 6.67390 \times 10^{-11} m^3/kg/sec^2.$$

The uncertainty is 0.0014%. From that Earth's and Sun's masses can be obtained. The values are $M_\oplus = 5.97223(\pm 0.00008) \times 10^{24} kg$ and $M_\odot = 1.98843(\pm 0.00003) \times 10^{30} kg$. The apparatus used (by Jens H. Gundlach *et al*, University of Washington, 2000) is a moderm version of the Cavendish torsion balance.

§ 3.24 Contracting (3.39) with $g^{\mu\nu}$, we find

$$R = -\, \tfrac{8\pi G}{c^4}\, T \, , \qquad (3.40)$$

where $T = g^{\mu\nu} T_{\mu\nu}$. This result can be inserted back into the Einstein equation, to give it the form

$$R_{\mu\nu} = \tfrac{8\pi G}{c^4}\, \left[T_{\mu\nu} - \tfrac{1}{2}\, g_{\mu\nu} T\right]. \qquad (3.41)$$

§ 3.25 Consider the sourceless case, in which $T_{\mu\nu} = 0$. It follows from the above equation that $R_{\mu\nu} = 0$ and, therefore, that $R = 0$. Notice that

this does *not* imply $R^\rho{}_{\sigma\mu\nu} = 0$. The Riemann tensor can be nonvanishing even in the absence of source. Einstein's equations are non-linear and, in consequence, the gravitational field can engender itself. Absence of gravitation is signalled by $R^\rho{}_{\sigma\mu\nu} = 0$, which means a flat spacetime. This case — Minkowski spacetime — is a particular solution of the sourceless equations. Beautiful examples of solutions without any source at all are the de Sitter spaces (see subsection 4.3.2 below).

§ 3.26 It is usual to introduce test particles, and test fields, to probe a gravitational field, as we have repeatedly done when discussing geodesics. They are meant to be just that, test objects. They are supposed to have no influence on the gravitational field, which they see as a background. They do not contribute to the source.

§ 3.27 In reality, the Einstein tensor (2.59) is not the most general parallel-transported purely geometrical second-order tensor which has vanishing co-variant derivative. The metric has the same property. Consequently, it is in principle possible to add a term $\Lambda g_{\mu\nu}$ to $G_{\mu\nu}$, with Λ a constant. Equation (3.39) becomes

$$R_{\mu\nu} - (\tfrac{1}{2}\,R + \Lambda)g_{\mu\nu} = \tfrac{8\pi G}{c^4}\,T_{\mu\nu}. \tag{3.42}$$

From the point of view of covariantly preserved objects, this equation is as valid as (3.39). In his first trial to apply his theory to cosmology, Einstein looked for a static solution. He found it, but it was unstable. He then added the term $\Lambda g_{\mu\nu}$ to make it stable, and gave to Λ the name *cosmological constant*. Later, when evidence for an expanding universe became clear, he called this "the biggest blunder in his life", and dropped the term. This is the same as putting $\Lambda = 0$. It was not a blunder: recent cosmological evidence claims for $\Lambda \neq 0$. Equation (3.42) is the *Einstein's equation with a cosmological term*. With this extra term, Eq,(3.41) becomes

$$R_{\mu\nu} = \frac{8\pi G}{c^4}\,\left[T_{\mu\nu} - \tfrac{1}{2}\,g_{\mu\nu}T\right] - \Lambda g_{\mu\nu}. \tag{3.43}$$

3.4   Action of the Gravitational Field

§ 3.28 Einstein's equations can be derived from an action functional, the Hilbert-Einstein action

$$S[g] = \int \sqrt{-g}\ R\ d^4x\ . \tag{3.44}$$

It is convenient to separate the metric as soon as possible, as in $R = g^{\mu\nu} R_{\mu\nu}$. Variations in the integration measure, which is metric-dependent, are concentrated in the Jacobian $\sqrt{-g}$. Taking variations with respect to the metric,

$$\frac{\delta S[g]}{\delta g^{\rho\sigma}} = \int \frac{\delta\sqrt{-g}}{\delta g^{\rho\sigma}}\ g^{\mu\nu} R_{\mu\nu}\ d^4x + \int \sqrt{-g}\ \frac{\delta g^{\mu\nu}}{\delta g^{\rho\sigma}} R_{\mu\nu}\ d^4x + \int \sqrt{-g}\ g^{\mu\nu} \frac{\delta R_{\mu\nu}}{\delta g^{\rho\sigma}}\ d^4x.$$

The first term is

$$\frac{\delta\sqrt{-g}}{\delta g^{\rho\sigma}} =$$

$$\frac{1}{2\sqrt{-g}}\ \frac{\delta(-g)}{\delta g^{\rho\sigma}} = \frac{1}{2\sqrt{-g}}\ \frac{\delta(\exp[tr\ln\mathbf{g}])}{\delta g^{\rho\sigma}} = \frac{1}{2\sqrt{-g}}\ \exp[tr\ln\mathbf{g}]\ \frac{\delta tr\ln\mathbf{g}}{\delta g^{\rho\sigma}}$$

$$= \frac{1}{2\sqrt{-g}}\ (-g)\ tr\frac{\delta\ln\mathbf{g}}{\delta g^{\rho\sigma}} = \frac{1}{2\sqrt{-g}}\ (-g)\ tr\left[\mathbf{g}^{-1}\ \frac{\delta\mathbf{g}}{\delta g^{\rho\sigma}}\right]$$

$$= \frac{1}{2\sqrt{-g}}\ (-g)\ \left[g^{\mu\nu}\ \frac{\delta g_{\mu\nu}}{\delta g^{\rho\sigma}}\right] = -\frac{1}{2\sqrt{-g}}\ (-g)\ \left[g_{\mu\nu}\ \frac{\delta g^{\mu\nu}}{\delta g^{\rho\sigma}}\right] = -\tfrac{1}{2}\ \sqrt{-g}\ g_{\rho\sigma}.$$

The last term can be shown to produce a total divergence and can be dropped. The first two terms give

$$\frac{\delta S[g]}{\delta g^{\rho\sigma}} = \sqrt{-g}\ \left[R_{\rho\sigma} - \tfrac{1}{2}\ g_{\rho\sigma}\ R\right].$$

This is the left-hand side of (3.39). In the absence of sources, Einstein's equation reduces to

$$\sqrt{-g}\ \left[R_{\mu\nu} - \tfrac{1}{2}\ g_{\mu\nu}R\right] = 0. \tag{3.45}$$

§ 3.29 A few comments:

1. variations have been taken with respect to the metric, which is the fundamental field

2. it is perhaps strange that the variation of $R_{\mu\nu}$, which encapsulates the gravitational field-strength, gives no contribution

3. If a source is present, with Lagragian density $\mathcal{L}$ and action

$$S_{source} = \int \sqrt{-g}\ \mathcal{L}\ d^4x, \qquad (3.46)$$

its contribution to the field equation, that is, its modified energy-momentum tensor density, will be given by

$$- \sqrt{-g}\ T_{\rho\sigma} = \frac{\delta}{\delta g^{\rho\sigma}}\ S_{source} = \int \frac{\delta}{\delta g^{\rho\sigma}} \left[\sqrt{-g}\ \mathcal{L} \right]\ d^4x$$

$$= \int \sqrt{-g}\ \frac{\delta \mathcal{L}}{\delta g^{\rho\sigma}}\ d^4x + \int \mathcal{L}\ \frac{\delta \sqrt{-g}}{\delta g^{\rho\sigma}}\ d^4x = \sqrt{-g}\ \left[\frac{\delta \mathcal{L}}{\delta g^{\rho\sigma}} - \tfrac{1}{2} g_{\rho\sigma} \mathcal{L} \right].$$

Consequently,

$$T_{\rho\sigma} = - \frac{\delta \mathcal{L}}{\delta g^{\rho\sigma}} + \tfrac{1}{2} g_{\rho\sigma} \mathcal{L}\ . \qquad (3.47)$$

4. Einstein's equation (3.42) with a cosmological term comes, in and analogous way, from the action

$$S[g] = \int \sqrt{-g}\ (R + 2\Lambda)\ d^4x. \qquad (3.48)$$

§ **3.30** We have used in §1.15 a dimensional argument to write the action of a test particle and of its coupling with an electromagnetic potential. Naïve dimensional analysis can be very useful in field theories. Not only do they help keeping trace of factors in calculations, but also lead to deep questions in the problem of quantization. Actually, naïve dimensional considerations are enough to exhibit a fundamental difference between gravitation and all the other basic interactions. Indeed, all the coupling constants are dimensionless quantities (think of the electric charge e), except that of gravitation. This is related to the Lagrangean density $\sqrt{-g}R$, whose dimension differs from those of the other theories. Let us be naïve for a while, and consider usual mechanical dimensions in terms of mass (M), length (L) and time (T). The dimension of a velocity will be represented as $[v] = [c] = [LT^{-1}]$; that of a force as $[F] = [MLT^{-2}]$; an energy will have $[E] = [FL] = [ML^2T^{-2}]$; a

80

pressure, $[p] = [FL^{-2}] = [EL^{-3}]$; an action, $[S] = [\hbar] = [ET] = ML^2T^{-1}$. Metric is dimensionless, so that $[ds] = [dx] = [L]$. The fact that a quantity is dimensionless will be represented as in $[e] = [g] = [0]$.

Field theory makes use of natural units (seemingly forbidden by international law), in which $\hbar = 1$ and $c = 1$. In that scheme, which greatly simplify discussions, actions and velocities are dimensionless. In consequence, $[L] = [T] = [M^{-1}]$ and only one mechanical dimension remains. Usually, the mass M is taken as the standard reference. A new set turns out. As examples, $[F] = [2]$, $[E] = [M]$, $[ds] = [M^{-1}]$. We say then that force has "numeric" dimension zero, energy has dimension one, length has dimension minus one.

quantity	usual dimension	natural dimension	numeric
mass	M	M	+ 1
length	L	M^{-1}	- 1
time	T	M^{-1}	- 1
velocity	LT^{-1}	M^0	0
acceleration	LT^{-2}	M	+ 1
force	MLT^{-2}	M^2	+ 2
Newton's G	$M^{-1}L^3T^{-2}$	M^{-2}	- 2
energy	ML^2T^{-2}	M^1	+ 1
action	ML^2T^{-1}	M^0	0
pressure	$ML^{-1}T^{-2}$	M^{-2}	- 2
$g_{\mu\nu}$	$M^0L^0T^0$	M^0	0
ds	L	M^{-1}	- 1
charge e	$M^0L^0T^0$	M^0	0
A_μ	ML^2T^{-2}	M	+ 1
$F_{\mu\nu}$, E, B	MLT^{-2}	M^2	+ 2
$T_{\mu\nu}$	$ML^{-1}T^{-2}$	M^{-2}	- 2
$\int \sqrt{-g} F^{\mu\nu} F_{\mu\nu} d^4x$	$M^0L^0T^0$	M^0	0
$\Gamma^\lambda{}_{\mu\nu}$	L^{-1}	M	+ 1
$R^\lambda{}_{\mu\nu\rho}$, R	L^{-2}	M^2	+ 2
$\int \sqrt{-g} \, R \, d^4x$	L^2	M^{-2}	- 2

Fields representing elementary particles have, in general, natural dimension + 1. Notice that $\int \sqrt{-g}\ R\ d^4x$ has not the dimension of an action. Actually, in order to give coherent results, it must be multiplied by a constant of dimension MT^{-1}, actually $-\frac{c^3}{16\pi G}$.

3.5   Non-Relativistic Limit

§ **3.31** A massive particle follows the geodesic equation (3.10),

$$\frac{dU^\alpha}{ds} + \Gamma^\alpha{}_{\beta\gamma} U^\beta U^\gamma = \frac{d^2 x^\alpha}{ds^2} + \Gamma^\alpha{}_{\beta\gamma} \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} = 0, \qquad (3.49)$$

which comes from the first term in action (1.7),

$$S = -\ mc \int ds \ . \qquad (3.50)$$

To compare with the non-relativistic case, we recall that the motion of a particle in a gravitational field is in that case described by the Lagrangian

$$L = -\ mc^2 + \frac{mv^2}{2} -\ mV \ .$$

This means the action

$$S = -\ \int mc \left[c - \frac{v^2}{2c} + \frac{V}{c} \right] dt \ . \qquad (3.51)$$

Comparison with (3.50) shows that necessarily

$$ds = \left[c - \frac{v^2}{2c} + \frac{V}{c} \right] dt$$

which, neglecting all the smaller terms, gives

$$ds^2 = \left[1 + \frac{2V}{c^2} \right] c^2 dt^2 - d\mathbf{x}^2 \ . \qquad (3.52)$$

We see that, in the non-relativistic limit,

<div style="text-align:right">

non−
retativistic
metric

</div>

$$g_{11} = g_{22} = g_{33} = -1 \ ; \quad g_{00} = 1 + \frac{2V}{c^2} \ . \qquad (3.53)$$

According to Special Relativity, the energy-momentum density tensor (3.36) reduces to the sole component

$$T_{00} = \rho c^2, \tag{3.54}$$

where ρ is the mass density. The trace has the same value, $T = \rho c^2$. If we use Eq.(3.41), we have

$$R_{\mu\nu} = \frac{8\pi G}{c^2} \, \rho \, \left[\delta_{0\mu}\delta_{\nu 0} - \tfrac{1}{2} \, g_{\mu\nu}\right] .$$

In particular,

$$R_{00} = \frac{4\pi G}{c^2} \, \rho \, . \tag{3.55}$$

The other cases are $R_{i\neq j} = 0$ and $R_{ii} = 4\pi G\rho V/c^4 \approx =$ in our approximation. We can then proceed to a careful calculation of R_{00} as given by (2.48), using (3.53) and (2.40). It turns out that the terms in $\Gamma\,\Gamma$ are of at least second order in v/c. The derivatives with respect to $x^0 = ct$ are of lower order if compared with the derivatives with respect to the space coordinates x^k, due to the presence of the factors $1/c$. What remains is $R_{00} = \frac{\partial\Gamma^k{}_{00}}{\partial x^k}$. It is also found that $\Gamma^k{}_{00} \approx -\tfrac{1}{2} \, g^{kj}\frac{\partial g_{00}}{\partial x^j}$. But this is $= \frac{1}{c^2} \, \frac{\partial V}{\partial x^k}$. Thus,

$$R_{00} = \frac{1}{c^2} \, \frac{\partial}{\partial x^k} \, \frac{\partial V}{\partial x^k} \; = \frac{1}{c^2} \, \Delta V \, . \tag{3.56}$$

Comparison with (3.55) leads then to

Poisson
recovered

$$\Delta V = 4\pi G \, \rho \, . \tag{3.57}$$

This shows how Einstein's equation (3.39) reduces to the Poisson equation (3.38) in the non-relativistic limit. By the way, the above result confirms the value of the constant introduced in (3.39).

§ **3.32** Notice that the non-relativistic limit corresponds to a weak gravitational field. A strong field would accelerate the particles so that soon the small-velocity approximation would fail.

§ **3.33** If the cosmological constant Λ is nonvanishing, then we should use (3.43) to obtain $R_{\mu\nu}$, instead of Eq.(3.41) as above. An extra term Λg_{00}

appears in R_{00}. The Poisson equation becomes $\Delta V = 4\pi G \rho - \Lambda c^2$. To examine the effect of this term in the non-relativistic limit, we can separate the potential in two pieces, $V = V_1 + V_2$, with $\Delta V_1 = 4\pi G \rho$. Then, $\Delta V_2 = -\Lambda c^2$ has for solution $V_2 = -\Lambda c^2 r^2/6$, leading to a force $F_\Lambda(r) = \Lambda c^2 r/3$. As $\Lambda > 0$ by present-day evidence, this harmonic oscillator-type force is repulsive. This is a universal effect: the cosmological constant term produces repulsion between any two bodies.

§ 3.34 Let us now examine what happens to the geodesic equation. First of all, the four-velocity U has the components $(\gamma, \gamma \mathbf{v}/c)$ in Special Relativity, with $\gamma = 1/\sqrt{1 - \mathbf{v}^2/c^2}$. In the non-relativistic limit,

$$\gamma \approx 1 + \tfrac{1}{2} \tfrac{\mathbf{v}^2}{c^2} \; ;$$

$$U \approx (1 + \tfrac{1}{2} \tfrac{\mathbf{v}^2}{c^2}, \mathbf{v}/c) \; .$$

Concerning the interval ds, all 3-space distances $|d\mathbf{x}|$ are negligible if compared with cdt, so that $ds^2 \approx c^2 dt^2$. Consequently,

$$\frac{dU}{ds} \approx \left(\frac{d}{d(ct)} \left[\tfrac{1}{2} \tfrac{\mathbf{v}^2}{c^2} \right], \frac{1}{c^2} \frac{d}{dt} \mathbf{v} \right) \; .$$

Now, in order to use the geodesic equation, we have to calculate the components of the connection given in Eq.(2.40),

$$\overset{\circ}{\Gamma}{}^\lambda{}_{\rho\sigma} = \tfrac{1}{2} g^{\lambda\mu} \{ \partial_\rho g_{\sigma\mu} + \partial_\sigma g_{\rho\mu} - \partial_\mu g_{\rho\sigma} \} \; .$$

To begin with, we notice that $g^{11} = g^{22} = g^{33} = -1$ and, to the order we are considering, $g^{00} = 1 - 2V/c$. Given the metric (3.53), only the derivatives with respect to the space variables of g_{00} are $\neq 0$. We arrive at the general expression

$$\overset{\circ}{\Gamma}{}^\rho{}_{\mu\nu} = \tfrac{1}{c^2} \left[(\delta^k_\mu \delta^0_\nu + \delta^k_\nu \delta^0_\mu) \, \delta^{\rho 0} - \delta^0_\nu \delta^0_\mu \delta^{\rho k} \right] \partial_k V \; .$$

The only non-vanishing components are

$$\overset{\circ}{\Gamma}{}^k{}_{00} = \overset{\circ}{\Gamma}{}^0{}_{0k} = \overset{\circ}{\Gamma}{}^0{}_{k0} = \tfrac{1}{c^2} \, \partial_k V \; .$$

The geodesic equations are:

1. for the time-like component,

$$0 = \frac{dU^0}{ds} + \overset{\circ}{\Gamma}{}^0{}_{\mu\nu} U^\mu U^\nu \approx \frac{d}{d(ct)} \left[\tfrac{1}{2} \, \tfrac{\mathbf{v}^2}{c^2} \right] + 2 \, \overset{\circ}{\Gamma}{}^0{}_{k0} U^k U^0$$

$$\approx \frac{d}{dt} \left[\tfrac{1}{2} \, \tfrac{\mathbf{v}^2}{c^3} \right] + 2 \, \tfrac{1}{c^3} \, \partial_k V v^k \ .$$

Both terms are of order $1/c^3$, equally negligible in our approximation.

2. the space-like components are more informative:

$$0 = \frac{dU^k}{ds} + \overset{\circ}{\Gamma}{}^k{}_{\mu\nu} U^\mu U^\nu \approx \frac{1}{c^2} \, \frac{d}{dt} \, v^k + \overset{\circ}{\Gamma}{}^k{}_{00} U^0 U^0 = \tfrac{1}{c^2} \, \frac{d}{dt} \, v^k + \tfrac{1}{c^2} \, \partial_k V \ ,$$

which is the force equation

<div style="float:right; border:1px solid black; padding:2px;">**force equation**</div>

$$\frac{d}{dt} \, v^k = - \, \partial^k V \ . \tag{3.58}$$

§ 3.35 In the non-relativistic limit, only g_{00} remains, as well as R_{00}. Einstein's equation greatly enlarge the the scope of the problem. What they achieve is better stateid in the words of Mashhoon (reference of the footnote in page 3):

> "the newtonian potential V is generalized to the ten components of the metric; the acceleration of gravity is replaced by the Christoffel connection; and the tidal matrix $\frac{\partial^2 V}{\partial x^i \partial x^j}$ is replaced by the curvature tensor; the trace of the tidal matrix is related to the local density of matter ρ; the Riemann tensor is related to the energy–momentum tensor"

3.6   About Time, and Space

3.6.1   Time Recovered

§ 3.36 A gravitational field is said to be *constant* when a reference frame exists in which all the components $g_{\mu\nu}$ are independent of the "time coordinate" x^0. This coordinate, by the way, is usually referred to as "coordinate time", or "world time". The non-relativistic limit given above is an example of constant gravitational field, as the potential V in (3.53) is supposed to be time-independent.

§ 3.37 Let us now examine the relation between proper time and the "coordinate time" $x^0 = ct$. Consider two close events at the same point in space. As $d\vec{x} = 0$, the interval between them will be $ds = cd\tau$, where τ is time as seen at the point. This means that $ds^2 = c^2 d\tau^2 = g_{00}(dx^0)^2$, or that

$$d\tau = \frac{1}{c} \sqrt{g_{00}} \, dx^0 = \sqrt{g_{00}} \, dt \; . \qquad (3.59)$$

As long as the coordinates are defined, the time lapse between two events at the same point in space will be given by

$$\tau = \frac{1}{c} \int \sqrt{g_{00}} \, dx^0 \; . \qquad (3.60)$$

This is the proper time at the point. In a constant gravitational field,

$$\tau = \frac{1}{c} \sqrt{g_{00}} \, x^0 \; . \qquad (3.61)$$

Once a coordinate system is established, it is the coordinate time which is seen from "abroad". Nevertheless, a test particle plunged in the field will see its proper time.

§ 3.38 Consider some periodic phenomenon taking place in a constant gravitational field. Its period will be given by the formula above and, as such, will be different when measured in coordinate time or with a "proper" clock. Its proper frequency will have, for the same reason, the values

$$\omega = \frac{\omega_0}{\sqrt{g_{00}}} \; . \qquad (3.62)$$

In the non-relativistic limit, (3.53) will lead to

$$\omega = \omega_0 \left(1 - \frac{V}{c^2} \right) \; . \qquad (3.63)$$

§ 3.39 Consider a light ray goint from a point "1" to a point "2" in a weak constant gravitational field. If the values of the potential are V_1 and V_2 at points "1" and "2", its proper frequency will be $\omega_1 = \omega_0 \left(1 - \frac{V_1}{c^2} \right)$ at point "1" and $\omega_2 = \omega_0 \left(1 - \frac{V_2}{c^2} \right)$ at point "2". It will consequently change while

moving from one point to the other. As the potential is a negative function, $V = -|V|$, this change will be given by

$$\Delta \omega = \omega_2 - \omega_1 = \omega_0 \left(\frac{|V_2| - |V_1|}{c^2} \right) . \qquad (3.64)$$

If the field is stronger in "1" than in "2", $|V_1| > |V_2|$ and $\omega_2 < \omega_1$. The frequency, if in the visible region, will become redder: this is the phenomenon of gravitational *red-shift*.

<div style="float:right;border:1px solid;">red shift</div>

§ **3.40** This effect provides one of the three "classical" tests of General Relativity. The others will be seen later: they are the precession of the planets perihelia (§ 4.12) and the light ray deviation (§ 4.13).

3.6.2   Space

§ **3.41** In Special Relativity, it is enough to put $dx^0 = 0$ in the interval ds to get the infinitesimal space distance dl. The relationship between the proper time and the coordinate time is the same at every point. This is no more the case in General Relativity. The standard procedure to obtain the space interval runs as follows (see Figure 3.2). Consider two close points in space, $P = (x^\mu)$ and $Q = (x^\mu + dx^\mu)$. Suppose Q sends a light signal to P, at which there is a mirror which sends it back through the same space path. With time τ measured at Q, the distance between Q and P will be $dl = c\tau/2$. The interval, which vanishes for a light signal, is given by

$$ds^2 = g_{ij}dx^i dx^j + 2\ g_{0j}dx^0 dx^j + g_{00}dx^0 dx^0 = 0 . \qquad (3.65)$$

Solving this second-order polynomial for dx^0, we find the two values

$$dx^0_{(1)} = \frac{1}{g_{00}} \left[-g_{0j}dx^j - \sqrt{(g_{0i}g_{0j} - g_{ij}g_{00})dx^i dx^j} \right] ;$$

$$dx^0_{(2)} = \frac{1}{g_{00}} \left[-g_{0j}dx^j + \sqrt{(g_{0i}g_{0j} - g_{ij}g_{00})dx^i dx^j} \right] .$$

The interval in coordinate time from the emission to the reception of the signal at Q will be

$$dx^0_{(2)} - dx^0_{(1)} = \frac{2}{g_{00}} \sqrt{(g_{0i}g_{0j} - g_{ij}g_{00})dx^i dx^j}.$$

The corresponding proper time is obtained by (3.59):

$$d\tau = \frac{2}{c\sqrt{g_{00}}}\sqrt{(g_{0i}g_{0j} - g_{ij}g_{00})dx^i dx^j}.$$

The space interval $dl = c\tau/2$ is then

$$dl = \sqrt{\left(\frac{g_{0i}g_{0j}}{g_{00}} - g_{ij}\right)dx^i dx^j},$$

or

$$dl^2 = \gamma_{ij}dx^i dx^j \ , \text{with} \quad \gamma_{ij} = -\ g_{ij} + \frac{g_{0i}g_{0j}}{g_{00}}\ . \tag{3.66}$$

This is the space metric. Curiously enough, the inverse is simpler: it so happens that

$$\gamma^{ij} = -\ g^{ij}\ . \tag{3.67}$$

This can be seen by contracting g^{ij} with γ_{jk}.
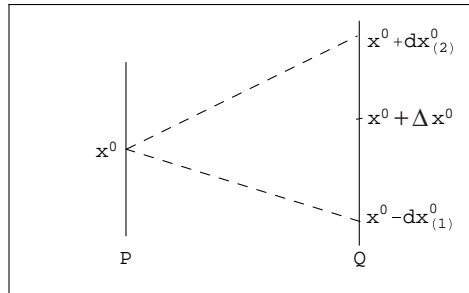
Figure 3.2: *Q sends a light sign towards a mirror at P, which sends it back.*

§ **3.42** Finite distances in space have no meaning in the general case, in which the metric is time-dependent. If we integrate $\int dl$ and take the infimum (as explained in § 2.45), the result will depend on the world-lines. Only constant gravitational fields allow finite space distances to be defined.

§ 3.43 When are two events simultaneous ? In the case of Figure 3.2, the instant x^0 of P should be simultaneous to the instant of Q which is just in the middle between the emission and the reception of the signal, which is

$$x^0 + \Delta x^0 = x^0 + \tfrac{1}{2} \left[dx^0_{(2)} + dx^0_{(1)} \right] .$$

Thus, the difference between the coordinate times of two simultaneous but spatially distinct events is

$$\Delta x^0 = - \; \frac{g_{0i}}{g_{00}} \; dx^i.$$

Time flows differently in different points of space. This relation allows one to synchonize the clocks in a small region. Actually, that can be progressively done along any open line. But not along a closed line: if we go along a closed curve, we arrive back at the starting point with a non-vanishing Δx^0. This is not a property of the field, but of the general frame. For any given field, it is possible to choose a system of coordinates in which the three components g_{0i} vanish. In that system, it is possible to synchonize the clocks.

§ 3.44 Suppose that a reference system exists in which

> synchronous
> system

1. the synchronization condition holds:

$$g_{0i} = 0 \tag{3.68}$$

2. and further

$$g_{00} = 1. \tag{3.69}$$

Then, the coordinate time $x^0 = ct$ in that system will represent proper time in all points covered by the system. That system is called a *synchronous system*. The interval will, in such a system, have the form

$$ds^2 = c^2 dt^2 - \gamma_{ij} dx^i dx^j \tag{3.70}$$

and

$$\gamma_{ij} = - \; g_{ij}. \tag{3.71}$$

89

Direct calculation shows that $\Gamma^\lambda{}_{00} = 0$ in such a coordinate system. This has an important consequence: the four-velocity $u^\lambda = \frac{dx^\lambda}{ds}$, tangent to the world-line $x^i = 0$ (i = 1, 2, 3)) has components $(u^0 = 1, u^i = 0)$ and satisfies the geodesic equation. Such geodesics are orthogonal to the surfaces $ct = constant$. In this way it is possible to conceive spacetime as a direct product of space (the above surfaces) and time. This system is not unique: any transformation preserving the time coordinate, or simple changing its origin, will give another synchronous system.

3.7   Equivalence, Once Again

§ 3.45 Consider a Levi-Civita connection $\overset{\circ}{\Gamma}$ defined on a manifold M and a point $P \in M$. Without loss of generality, we can choose around P a coordinate system $\{x^\mu\}$ such that $x^\mu(P) = 0$. Such a system will cover a neighbourhood N of P (its coordinate neighbourhood), and will provide dual holonomic bases $\{\frac{\partial}{\partial x^\mu}, dx^\mu\}$, with $dx^\lambda(\frac{\partial}{\partial x^\mu}) = \delta^\lambda_\mu$, for vector fields and 1-forms (covector fields) on N. Any other base $\{e_a, e^a$ such that $e^a(e_b) = \delta^a_b\}$, will be given by the components of its members in terms of such initial holonomic bases, as $\{e_a = e_a{}^\mu(x)\frac{\partial}{\partial x^\mu}, e^a = e^a{}_\lambda(x)dx^\lambda\}$. Take another coordinate system $\{y^a\}$ around P, with a coordinate neighborhood N'' such that $P \in N \cap N'' \neq \emptyset$. This second coordinate system will define another holonomic base $\{e_a = \frac{\partial}{\partial y^a} = \frac{\partial x^\mu}{\partial y^a} \frac{\partial}{\partial x^\mu}, e^a = dy^a = \frac{\partial y^a}{\partial x^\mu} dx^\mu\}$. It will be enough for our purposes to consider inside the intersection $N \cap N''$ a non-empty sub–domain N', small enough to ensure that only terms up to first order in the x^μ's can be retained in the calculations.

Let us indicate by $\overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}(x)$ the components of Γ in the first holonomic base, and by $\overset{\circ}{\Gamma}{}^a{}_{b\nu}(x)$ the components of Γ referred to base $\{\frac{\partial}{\partial y^a}, dy^a\}$. These components will be related by

$$\overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}(x) = \frac{\partial x^\lambda}{\partial y^a} \overset{\circ}{\Gamma}{}^a{}_{b\nu}(x) \frac{\partial y^b}{\partial x^\mu} + \frac{\partial x^\lambda}{\partial y^c} \frac{\partial}{\partial x^\nu} \frac{\partial y^c}{\partial x^\mu} \quad , \qquad (3.72)$$

or

$$\overset{\circ}{\Gamma}{}^a{}_{b\nu}(x) = \frac{\partial y^a}{\partial x^\lambda} \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}(x) \frac{\partial x^\mu}{\partial y^b} + \frac{\partial y^a}{\partial x^\rho} \frac{\partial}{\partial x^\nu} \frac{\partial x^\rho}{\partial y^b} \quad . \qquad (3.73)$$

90

Let us indicate by $\gamma^\lambda{}_{\mu\nu} = \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}(P)$ the value of the connection components at the point P in the first holonomic system. On a small enough domain N' the connection components will be approximated by

$$\overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}(x) = \gamma^\lambda{}_{\mu\nu} + x^\rho \, [\partial_\rho \overset{\circ}{\Gamma}{}^\lambda{}_{\mu\nu}]_P$$

to first order in the coordinates x^μ. Choose now the second system of coordinates

$$y^a = \delta^a_\mu x^\mu + \tfrac{1}{2} \, \delta^a_\lambda \gamma^\lambda{}_{\mu\nu} \, x^\nu x^\mu \; . \tag{3.74}$$

Then,

$$\frac{\partial y^a}{\partial x^\mu} = \delta^a_\mu + \delta^a_\lambda \, \gamma^\lambda{}_{\mu\nu} \, x^\nu \; ; \quad \frac{\partial x^\mu}{\partial y^a} = \delta^\mu_a - \delta^\rho_a \, \gamma^\mu{}_{\rho\nu} \, \delta^\nu_c y^c \; .$$

Taken into (3.73), these expressions lead to

$$\overset{\circ}{\Gamma}{}^a{}_{b\nu}(x) = \delta^a_\lambda \delta^\sigma_b \, \left[\partial_\epsilon \overset{\circ}{\Gamma}{}^\lambda{}_{\sigma\nu}(P) - \gamma^\lambda{}_{\rho\nu} \gamma^\rho{}_{\sigma\epsilon} \right] x^\epsilon. \tag{3.75}$$

We see that at the point P, which means $\{x^\mu = 0\}$, the connection components in base $\{\partial_a\}$ vanish: $\overset{\circ}{\Gamma}{}^a{}_{b\nu}(P) = 0$. The curvature tensor at P, however, is not zero:

$$\overset{\circ}{R}{}^\lambda{}_{\sigma\mu\nu}(P) = \partial_\mu \overset{\circ}{\Gamma}{}^\lambda{}_{\sigma\nu}(P) - \partial_\nu \overset{\circ}{\Gamma}{}^\lambda{}_{\sigma\mu}(P) + \gamma^\lambda{}_{\rho\mu}\gamma^\rho{}_{\sigma\nu} - \gamma^\lambda{}_{\rho\nu}\gamma^\rho{}_{\sigma\mu} \; .$$

It is thus possible to make the connection to vanish at each point by a suitable choice of coordinate system. The equation of force, or the geodesic equation, acquire the expressions they have in Special Relativity. The curvature, nevertheless, as a real tensor, cannot be made to vanish by a choice of coordinates.

Furthermore, we have from Eq.(3.74)

$$\frac{d}{ds} \, y^a = \delta^a_\lambda \, \left[U^\lambda + \gamma^\lambda{}_{(\mu\nu)} \, U^\nu x^\mu \right] \; ;$$

$$\frac{d^2}{ds^2} \, y^a = \delta^a_\lambda \left[\frac{d}{ds} U^\lambda + \gamma^\lambda{}_{(\mu\nu)} \, U^\nu U^\mu \right] + \delta^a_\lambda \gamma^\lambda{}_{(\mu\nu)} \, x^\mu \left[\frac{d}{ds} U^\nu \right] \; .$$

Then, at P,

$$\frac{dy^a}{ds} = \delta^a_\mu \, U^\mu(P) \; ;$$

$$\frac{d^2 y^a}{ds^2} = \delta^a_\lambda A^\lambda.$$

Suppose a self–parallel curve goes through P in N' with velocity U^μ. Then, at P, $\frac{dy^a}{ds} = \delta^a_\mu\, U^\mu(P) = \text{constant} = U^a(P)$ and $\frac{d^2}{ds^2}\, y^a = 0$. As by Eq.(3.74) $y^a(P) = 0$, this gives for the geodesic

$$y^a = U^a(P)\ s$$

in some neighborhood around P. The geodesic is, in this system, a straight line.

3.8   More About Curves

3.8.1   Geodesic Deviation

§ 3.46 Curvature can be revealed by the study of two nearby geodesics. Let us take again Eq. (3.1), rewritten in the form

$$U^\beta V^\alpha{}_{;\beta} = V^\beta U^\alpha{}_{;\beta} \tag{3.76}$$

The deviation between two neighboring geodesics in Figure 3.1 is measured by the vector parameter $\eta^\alpha = V^\alpha \delta V$, with δV a constant. It gives the difference between two points (such as a_1 and b_1 in that Figure) with the same value of the parameter u. Or, if we prefer, η relates two geodesics corresponding to V and $V + \delta V$.

Let us examine now the second-order derivative

$$\frac{D^2 V^\alpha}{Du^2} = (U^\beta V^\alpha{}_{;\beta})_{;\gamma} U^\gamma.$$

Use of Eq.(3.76) allows to write

$$\frac{D^2 V^\alpha}{Du^2} = (V^\beta U^\alpha{}_{;\beta})_{;\gamma} U^\gamma = V^\beta{}_{;\gamma} U^\gamma U^\alpha{}_{;\beta} + V^\beta U^\alpha{}_{;\beta;\gamma} U^\gamma$$

$$= U^\beta{}_{;\gamma} V^\gamma U^\alpha{}_{;\beta} + V^\beta U^\alpha{}_{;\beta;\gamma} U^\gamma\ ,$$

where once again use has been made of Eq.(3.76). Now, Eq.(2.55) leads to

$$\frac{D^2 V^\alpha}{Du^2} = (U^\gamma{}_{;\beta} V^\beta U^\alpha{}_{;\gamma} + V^\beta U^\gamma U^\alpha{}_{;\gamma;\beta}) - V^\beta U^\gamma \overset{\circ}{R}{}^\alpha{}_{\epsilon\beta\gamma} U^\epsilon$$

92

$$= V^\beta (U^\alpha{}_{;\gamma} U^\gamma)_{;\beta} + \mathring{R}^\alpha{}_{\epsilon\gamma\beta} U^\epsilon U^\gamma V^\beta \ .$$

The first term on the right-hand side vanishes by the geodesic equation. We have thus, for the parameter η,

$$\frac{D^2 \eta^\alpha}{Du^2} = \mathring{R}^\alpha{}_{\epsilon\gamma\beta} U^\epsilon U^\gamma \eta^\beta \ . \tag{3.77}$$

This the *geodesic deviation equation*.

As a heuristic, qualitative guide: test particles tend to close to each other in regions of positive curvature and to part from each other in regions of negative curvature.

3.8.2 General Observers

§ **3.47** Let us go back to the notion of observer introduced in § 3.11. An ideal observer — a timelike curve — will only feel the connection, and that can be made to vanish along a piece of that curve. Nevertheless, a real observer will have at least two points, each one following its own timelike curve. It will consequently feel curvature — that is, the gravitational field.

§ **3.48** Let us go back to curves, with a mind to observers. Given a curve γ, it is convenient to attach a vector basis at each one of its points. The best bases are those which are (pseudo-)orthogonal. This is to say that, if the members have components $h_a{}^\mu$, then

$$g_{\mu\nu} h_a{}^\mu h_b{}^\nu = \eta_{ab} \ . \tag{3.78}$$

Consider a set of 4 vectors e_0, e_1, e_2, e_3 at a point P on γ, satisfying the following conditions:

$$\frac{De_0}{Ds} = be_3; \ \ \frac{De_3}{Ds} = ce_1 + be_0; \ \ \frac{De_1}{Ds} = de_2 - ce_3; \ \ \frac{De_2}{Ds} = -\, d\, e_1 \ . \tag{3.79}$$

These are called the Frenet–Serret conditions. The choice of non-vanishing parameters is such as to allow the 4–vectors to be orthogonal. Actually, these four vectors are furthermore required to be orthonormal at each point of γ:

$$e_0 \cdot e_0 = 1; \ \ e_1 \cdot e_1 = e_2 \cdot e_2 = e_3 \cdot e_3 = -1 \ \ . \tag{3.80}$$

93

We shall always consider timelike curves, with velocity $U = e_0$. In the Frenet–Serret language, e_1, e_2, e_3 will be the first, second and third normals to γ at P. The parameters b, c, d are real numbers, called the first, second and third curvatures of γ at P.

By what we have said above, $be_3 = A$, the acceleration, which is orthogonal to the velocity. The absolute value of the first curvature of γ at P is, thus, the acceleration modulus: $|b| = |A|$.

For a geodesic, $b = c = d = 0$. The case $b = \text{constant}$, $c = d = 0$ corresponds to an hyperbola of constant curvature and the case $b = \text{constant}$, $c = \text{constant}$, $d = 0$, to a helix.

§ **3.49** Of course, most curves are not geodesics. A geodesic represents an observer in the absence of any external force. An observer may be settled on a linearly accelerated rocket, or turning around Earth, or still going through a mad spiral trajectory. An orthogonal tetrad defined as above remains an orthogonal tetrad under parallel transport, which also preserves the components of a vector in that tetrad. There is, however, a problem with parallel transport: if we take, as above, e_0 as the velocity U at a point P, e_0 will not be the velocity at other points of γ, unless γ is a geodesic.

There are other kinds of "transport" which preserve orthogonal tetrads and components. There is one, in special, which corrects the mentioned problem with the velocity:

The Fermi-Walker derivative of a vector V is defined by

<div style="text-align: right">Fermi–
Walker
transport</div>

$$\frac{D_{FW}V^\lambda}{Ds} = \frac{DV^\lambda}{Ds} - b\, V_\nu(e_0^\nu e_3^\lambda - e_0^\lambda e_3^\nu) = \frac{DV^\lambda}{Ds} - V_\nu(U^\nu A^\lambda - U^\lambda A^\nu). \quad (3.81)$$

A vector V is said to be Fermi–Walker–transported along a curve γ of velocity field U if its the Fermi-Walker derivative vanishes,

$$\frac{D_{FW}V^\lambda}{Ds} = \frac{DV^\lambda}{Ds} - V_\nu(U^\nu A^\lambda - U^\lambda A^\nu) = 0 \ . \qquad (3.82)$$

We see that, applied to U, $\frac{D_{FW}U^\lambda}{Ds} = 0$, and also that $\frac{D_{FW}}{Ds} = \frac{D}{Ds}$ if the curve γ is a geodesic. Take two vector fields X and Y such that $\frac{D_{FW}X^\lambda}{Ds} = 0$ and $\frac{D_{FW}Y^\lambda}{Ds} = 0$. Then, it follows that the component of X along Y is preserved along γ:

$$\frac{D(X_\lambda Y^\lambda)}{Ds} = \frac{d(X_\lambda Y^\lambda)}{ds} = 0.$$

94

In particular, if $e_0 = U$ at P, then e_0 will remain $= U$ along the curve if it is Fermi-Walker transported.

3.8.3 Transversality

§ 3.50 Given a curve γ whose tangent velocity is U, it is interesting to introduce a "transversal metric" by

$$h_{\mu\nu} = g_{\mu\nu} - U_\mu U_\nu \quad . \tag{3.83}$$

Transversality is evident: $h_{\mu\nu}U^\nu = 0$. In particular, $h_{\mu\nu}A^\nu = A_\mu$. A projector is an operator P satisfying $P^2 = P$. Matrices $(h^\sigma{}_\nu) = (g^{\sigma\mu}h_{\mu\nu}) = (\delta^\sigma_\nu - U^\sigma U_\nu)$ are projectors: $h^\sigma{}_\nu h^\nu{}_\mu = h^\sigma{}_\mu$. They satisfy $h^\mu{}_\nu h^\nu{}_\lambda = h^\mu{}_\nu g^{\nu\rho}h_{\rho\lambda} = h^\mu{}_\nu g^{\nu\rho}(g_{\rho\lambda} - U_\rho U_\lambda) = h^\mu{}_\lambda$. Notice that $g^{\mu\nu}h_{\mu\nu} = h^{\mu\nu}h_{\mu\nu} = h^\mu{}_\mu = 3$. The energy–momentum tensor (3.36) of a general fluid can be rewritten as

$$T_{\mu\nu} = (p + \epsilon)U_\mu U_\nu - p\, g_{\mu\nu} = \epsilon\, U_\mu U_\nu - p\, h_{\mu\nu}. \tag{3.84}$$

The transversal metric extracts the pressure:

$$T^{\mu\nu}h_{\mu\nu} = -\,3\,p.$$

The Einstein equations with this energy–momentum tensor give, by contraction,

$$R_{\mu\nu}U^\mu U^\nu = \frac{4\pi G}{c^4}\,(\epsilon + 3p) - \Lambda\ . \tag{3.85}$$

We have seen in § 3.21 that the streamlines of a general fluid — unlike those of a dust cloud — are not geodesics. The equation of force (3.37) has, actually, the form

fluid streamlines

$$(p + \epsilon)\,\frac{DU^\mu}{Ds} = h^{\mu\nu}\partial_\nu p\ . \tag{3.86}$$

The pressure gradient, as usual, engenders a force. In the relativistic case the force is always transversal to the curve. Here, it is the transversal gradient that turns up. The equation above governs the streamlines of a general fluid.

3.8.4 Fundamental Observers

§ 3.51 On a pseudo–Riemannian spacetime, there exists always a family of world–lines which is preferred. They represent the motion of certain preferred observers, the *fundamental observers* and the curves themselves are called the *fundamental world–lines*. Proper time coincides with the line parameter, so that he 4-velocities along these lines are $U^\mu = \frac{dx^\mu}{ds}$ and, consequently $U^2 = U^\mu U_\mu = 1$. The time derivative of a tensor $T^{\rho\sigma\cdots}{}_{\mu\nu\cdots}$ is $\frac{d}{ds}T^{\rho\sigma\cdots}{}_{\mu\nu\cdots} = \frac{d}{dx^\lambda}T^{\rho\sigma\cdots}{}_{\mu\nu\cdots}\frac{dx^\lambda}{ds} = (T^{\rho\sigma\cdots}{}_{\mu\nu\cdots})_{,\lambda} U^\lambda$. This is not covariant. The covariant time derivative, absolute derivative, is

$$\frac{D}{Ds}T^{\rho\sigma\cdots}{}_{\mu\nu\cdots} = (T^{\rho\sigma\cdots}{}_{\mu\nu\cdots})_{;\lambda} u^\lambda.$$

For example, the acceleration is

$$A^\mu = \frac{D}{Ds} U^\mu = U^\mu{}_{;\nu} U^\nu.$$

Using the Christoffel connection, it is easily seen that $U_\mu A^\mu = 0$. This property is analogous to that found in Minkowski space, but here only has an invariant sense if acceleration is covariantly defined, as above.

At each point P, *under a condition given below*, a fundamental observer has a 3-dimensional space which it can consider to be "hers/his own": its rest–space. Such a space is tangent to the pseudo–Riemannian spacetime and, as time runs along the fundamental world–line, orthogonal to that line at P (orthogonal to a line means orthogonal to its tangent vector, here U^μ). At each point of a world–line, that 3-space is determined by the projectors $h_{\mu\nu}$.

It is convenient to introduce the notations

$$U_{(\mu;\nu)} = \tfrac{1}{2} \left(U_{\mu;\nu} + U_{\nu;\mu}\right) \quad ; \quad U_{[\mu;\nu]} = \tfrac{1}{2} \left(U_{\mu;\nu} - U_{\nu;\mu}\right)$$

for the symmetric and antisymmetric parts of $U_{\mu;\nu}$. There are a few important notions to be introduced:

- the vorticity tensor

$$\omega_{\mu\nu} = h^\rho{}_\mu h^\sigma{}_\nu U_{[\rho;\sigma]} = U_{[\mu;\nu]} + U_{[\mu} A_{\nu]} \; ; \qquad (3.87)$$

  it satisfies $\omega_{\mu\nu} = \omega_{[\mu\nu]} = -\omega_{\nu\mu}$ and $\omega_{\mu\nu}U^\nu = 0$; it is frequently indicated by its magnitude $\omega^2 = \tfrac{1}{2} \omega_{\mu\nu} \omega^{\mu\nu} \geq 0$.

96

- the expansion tensor

$$\Theta_{\mu\nu} = h^{\rho}{}_{\mu} h^{\sigma}{}_{\nu} U_{(\rho;\sigma)} = U_{(\mu;\nu)} - U_{(\mu} A_{\nu)} \; ; \qquad (3.88)$$

its transversal trace is called the volume expansion; it is $\Theta = h^{\mu\nu}\Theta_{\mu\nu} = \Theta^{\mu}{}_{\mu} = U^{\mu}{}_{;\mu}$, the covariant divergence of the velocity field; it measures the spread of nearby lines, thereby recovering the original meaning of the word *divergence*; in the Friedmann model, Θ turns up as related to the Hubble expansion function by $\Theta = 3H(t)$.

- $\sigma_{\mu\nu} = \Theta_{\mu\nu} - \frac{1}{3}\,\Theta\, h_{\mu\nu} = \sigma_{(\mu\nu)}$ is the symmetric trace–free shear tensor; it satisfies $\sigma_{\mu\nu}u^{\nu} = 0$ and $\sigma^{\mu}{}_{\mu} = 0$ and its magnitude is defined as $\sigma^2 = \frac{1}{2}\,\sigma_{\mu\nu}\sigma^{\mu\nu} \geq 0$. Notice $\Theta_{\mu\nu}\Theta^{\mu\nu} = 2\,\sigma^2 + \frac{1}{3}\,\Theta^2$.

Decomposing the covariant derivative of the 4-velocity into its symmetric and antisymmetric parts, $U_{\rho;\sigma} = U_{(\rho;\sigma)} + U_{[\rho;\sigma]}$, and using the definitions (3.87) and (3.88), we find

$$U_{\mu;\nu} = \omega_{\mu\nu} + \Theta_{\mu\nu} + A_{\mu}U_{\nu} \; , \qquad (3.89)$$

or

$$U_{\mu;\nu} = \omega_{\mu\nu} + \sigma_{\mu\nu} + \tfrac{1}{3}\,\Theta\, h_{\mu\nu} + A_{\mu}U_{\nu} \; . \qquad (3.90)$$

§ **3.52** With the above characterizations of the energy density and the pressure, the Einstein equations reduce to the Landau–Raychaudhury equation. Let us go back to Eq.(2.55) and take its contracted version

<div style="float:right; border:1px solid black; padding:4px; text-align:center;">

Landau–
Raychaudhury
equation

</div>

$$U^{\alpha}{}_{;\alpha;\gamma} - U^{\alpha}{}_{;\gamma;\alpha} = -\,\mathring{R}_{\alpha\gamma}U^{\alpha} \; . \qquad (3.91)$$

Contracting now with U^{γ},

$$U^{\gamma}U^{\alpha}{}_{;\alpha;\gamma} - U^{\gamma}U^{\alpha}{}_{;\gamma;\alpha} = -\,\mathring{R}_{\alpha\gamma}U^{\alpha}U^{\gamma} \; .$$

$$\therefore \; \frac{D}{Ds}\,U^{\alpha}{}_{;\alpha} - (U^{\gamma}U^{\alpha}{}_{;\gamma})_{;\alpha} + U^{\alpha}{}_{;\gamma}U^{\gamma}{}_{;\alpha} + \mathring{R}_{\alpha\beta}U^{\alpha}U^{\beta} = 0 \; ,$$

$$\therefore \; \frac{d}{ds}\,U^{\alpha}{}_{;\alpha} - A^{\alpha}{}_{;\alpha} + U^{\alpha}{}_{;\gamma}\,U^{\gamma}{}_{;\alpha} + \mathring{R}_{\alpha\beta}U^{\alpha}U^{\beta} = 0 \; .$$

To obtain $U^{\alpha}{}_{;\gamma}\,U^{\gamma}{}_{;\alpha}$, we notice that

$$\Theta_{\mu\nu}\Theta^{\mu\nu} = \tfrac{1}{2}\,\left[U_{\alpha;\beta}\,U^{\alpha;\beta} + U_{\alpha;\beta}\,U^{\beta;\alpha} - A_{\alpha}A^{\alpha}\right] \; ;$$

97

$$\omega_{\mu\nu}\omega^{\mu\nu} = \tfrac{1}{2}\left[U_{\alpha;\beta}\, U^{\alpha;\beta} - U_{\alpha;\beta}\, U^{\beta;\alpha} - A_\alpha A^\alpha\right] \ .$$

It follows that

$$U^{\alpha}_{\ ;\gamma}\, U^{\gamma}_{\ ;\alpha} = \Theta_{\mu\nu}\Theta^{\mu\nu} - \omega_{\mu\nu}\omega^{\mu\nu} = 2\,\sigma^2 - 2\,\omega^2 + \frac{1}{3}\,\Theta^2 \ .$$

The equation acquires the aspect

$$\frac{d}{ds}\,\Theta + 2\,(\sigma^2 - \omega^2) + \frac{1}{3}\,\Theta^2 - A^{\alpha}_{\ ;\alpha} + \overset{\circ}{R}_{\alpha\beta}\, U^\alpha U^\beta = 0 \ . \qquad (3.92)$$

Up to this point, only definitions have been used. Einstein's equations lead, however, to Eq.(3.85), which allows us to put the above expression into the form

$$\frac{d}{ds}\,\Theta = -\,\tfrac{1}{3}\,\Theta^2 -\, 2\,(\sigma^2 - \omega^2) + A^{\mu}_{\ ;\mu} - \frac{4\pi G}{c^4}\,(\epsilon + 3p) + \Lambda \ . \qquad (3.93)$$

The promised condition comes from a detailed examination which shows that, actually, only when $\omega_{\mu\nu} = 0$ there exists a family of 3-spaces everywhere orthogonal to U^μ. In that case, there is a well–defined time which is the same over each 3-space.

From the above equation, we can see the effect of each quantity on expansion: as we proceed along a fundamental world–line, expansion

- decreases (an indication of attraction) with higher values of

    expansion itself

    shear

    energy content

- increases (an indication of repulsion) with higher values of

    vorticity

    second-acceleration

    cosmological constant .

3.9   An Aside: Hamilton-Jacobi

§ 3.53 The action principle we have been using [say, as in Eq.(1.5), or (3.51)] has a "teleological" character which brings forth a causal problem. When we look for the curve γ which minimizes the functional

$$S[\gamma] = \int_{t_0}^{t_1} L dt = \int_{\gamma_{PQ}} L dt \ ,$$

we seem to suppose that the behavior of a particle, starting from a point P at instant t_0, is somehow determined by its future, which forcibly consists in being at a fixed point Q at instant t_1. Another notion of action exists which avoids this difficulty. Instead of as a functional S is conceived, in that version, as a function

$$S(q_1(t), q_2(t), ..., q_n(t), t) =$$

$$\int_{t_0}^{t} dt' L\left[q_1(t'), q_2(t'), ..., q_n(t'), \dot{q}_1(t'), \dot{q}_2(t'), ..., \dot{q}_n(t')\right] \qquad (3.94)$$

of the final time t and the values of the generalized coordinates at that instant *for the real trajectory*. The particle, by satisfying the Lagrange equation at each point of its path, automatically minimizes S. In effect, taking the variation

$$\delta S =$$

$$\int_{t_0}^{t} dt' \left[\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i\right] = \int_{t_0}^{t} dt' \left[\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial}{\partial t'}\left(\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right) - \left(\frac{\partial}{\partial t'}\frac{\partial L}{\partial \dot{q}^i}\right)\delta q^i\right]$$

$$= \left[\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right]_{t_0}^{t} + \int_{t_0}^{t} dt' \left[\frac{\partial L}{\partial q^i} - \left(\frac{\partial}{\partial t'}\frac{\partial L}{\partial \dot{q}^i}\right)\right]\delta q^i \ .$$

The second term vanishes by Lagrange's equation. In the first term, $\delta q^i(t_0) = 0$, so that

$$\delta S = \frac{\partial L}{\partial \dot{q}^i} \ \delta q^i = p_i \ \delta q^i \ , \qquad (3.95)$$

which entails

$$p_i = \frac{\partial S}{\partial q^i} \ . \qquad (3.96)$$

We have been forgetting the time dependence of S. The integral (3.94) says that $L = \frac{dS}{dt}$. On the other hand,

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial q^i}\, \dot{q}^i = \frac{\partial S}{\partial t} + p_i \dot{q}^i.$$

Consequently,

$$\frac{\partial S}{\partial t} = L - p_i \dot{q}^i = -\,H. \tag{3.97}$$

Thus, the total differential of S will be

$$dS = p_i dq^i - H dt\ . \tag{3.98}$$

Variation of the action integral

$$S = \int \left[p_i dq^i - H dt \right] \tag{3.99}$$

leads indeed to Hamilton's equations:

$$\delta S = \int \left[\delta p_i dq^i + p_i d\delta q^i - \frac{\partial H}{\partial q^i}\, \delta q^i dt - \frac{\partial H}{\partial p_i}\, \delta p_i dt \right]$$

$$= \int \left[\delta p_i \frac{dq^i}{dt} + p_i \frac{d\delta q^i}{dt} - \frac{\partial H}{\partial q^i}\, \delta q^i - \frac{\partial H}{\partial p_i}\, \delta p_i \right] dt$$

$$= \int \left[\delta p_i \left(\frac{dq^i}{dt} - \frac{\partial H}{\partial p_i} \right) - \left(\frac{dp_i}{dt} + \frac{\partial H}{\partial q^i} \right) \delta q^i \right] dt.$$

An integration by parts was performed to arrive at the last espression which, to produce $\delta S = 0$ for arbitrary δq^i and δp_i, enforces

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}\ ; \quad \frac{dp_i}{dt} = -\,\frac{\partial H}{\partial q^i}\ . \tag{3.100}$$

§ 3.54 Hamilton's equations are invariant under canonical transformations leading to new variables $Q^i = Q^i(q^i, p_j, t)$, $P_i = P_i(q^i, p_j, t)$, $H'(Q^i, P_j, t)$. This means that if

$$\delta \int \left[p_i dq^i - H dt \right] = 0\ ,$$

then also

$$\delta \int \left[P_i dQ^i - H' dt \right] = 0$$

must hold. In consequence, the two integrals must differ by the total differential of an arbitrary function:

$$p_i dq^i - H dt = P_i dQ^i - H' dt + dF \ .$$

F is the *generating function* of the canonical transformation. It is such that

$$dF = p_i dq^i - P_i dQ^i + (H' - H) dt \ .$$

Therefore,

$$p_i = \frac{\partial F}{\partial q^i} \ ; \quad P_i = - \frac{\partial F}{\partial Q^i} \ ; \quad H' = H + \frac{\partial F}{\partial t} \ . \tag{3.101}$$

In the formulas above, the generating function appears as a function of the old and the new generalized coordinates, $F = F(q, Q, t)$. We obtain another generating function $f = f(q, P, t)$, with the new momenta instead of the new generalized coordinates, by a Legendre transformation: $f = F + Q^i P_i$, for which

$$df = dF + dQ^i P_i + Q^i dP_i = p_i dq^i + Q^i dP_i + (H' - H) dt \ .$$

In this case,

$$p_i = \frac{\partial f}{\partial q^i} \ ; \quad Q^i = \frac{\partial f}{\partial P_i} \ ; \quad H' = H + \frac{\partial f}{\partial t} \ . \tag{3.102}$$

Other generating functions, related to other choices of arguments, can of course be chosen. For instance, the function $g = \sum_i q^i Q^i$ generates a simple interchange of the initial coordinates and momenta.

§ 3.55 Let us go back to Eq. (3.97),

$$\frac{\partial S}{\partial t} + H(q^1, q^2, ..., q^n, p_1, p_2, ..., p_n, t) = 0 \ . \tag{3.103}$$

By Eq. (3.96), the momenta are the gradients of the action function S. Substituting their expressions in the above formula, we find a first order partial differential equation for S,

$$\frac{\partial S}{\partial t} + H\left(q^1, q^2, ..., q^n, \frac{\partial S}{\partial q^1}, \frac{\partial S}{\partial q^2}, ..., \frac{\partial S}{\partial q^n}, t\right) = 0 \ . \tag{3.104}$$

101

This is the *Hamilton-Jacobi equation*. The general solution of such an equation depends on an arbitrary function. The solution which is important for Mechanics is not the general solution, but the so-called complete solution (from which, by the way, the general solution can be recovered). That solution contains one arbitrary constant for each independent variable, (n+1) in the case above. Notice that only derivatives of S appear in the equation. One of the constants (C below) turns up, consequently, isolated. The complete solution has the form

$$S = f\left(q^1, q^2, ..., q^n, a_1, a_2, \ldots, a_n, t\right) + C. \qquad (3.105)$$

We have indicated the arbitrary constants by $a_1, a_2, \ldots, a_n$ and C.

The connection to the mechanical problem is made as follows. Consider a canonical transformation with generating function f, taking the original variables $(q^1, q^2, ..., q^n, p_1, p_2, ..., p_n)$ into $(Q^1, Q^2, ..., Q^n, a_1, a_2, ..., a_n)$. This is a transformation of the type summarized in Eq. (3.102), with $a_1, a_2, ..., a_n$ as the new momenta. From those equations and (3.105),

$$p_i = \frac{\partial S}{\partial q^i} \;;\quad Q^i = \frac{\partial S}{\partial P_i} \;;\quad H' = H + \frac{\partial S}{\partial t} = 0 \;. \qquad (3.106)$$

To get the vanishing of the last expression use has been made of Eq. (3.104). Hamilton equations havee then the solutions $Q^n = $ constant, $a_k = $ constant. From the equations $Q^i = \frac{\partial S}{\partial a_i}$ it is possible to obtain back

$$q^k = q^k\left(Q^1, Q^2, ..., Q^n, a_1, a_2, \ldots, a_n, t\right) \;,$$

that is, the old coordinates written in terms of $2n$ constants and the time. This is the solution of the equation of motion. Summing up, the procedure runs as follows:

- given the Hamiltonian, one looks for the complete solution (3.105) of the Hamilton-Jacobi equation (3.104);

- once the solution S is obtained, one derives with respect to the constants a_k and equate the results to the new constants Q^k; the equations $Q^i = \frac{\partial S}{\partial a_i}$ are algebraic;

- that set of algebraic equations are then solved to give the coordinates $q^k(t)$;

- the momenta are then found by using $p_i = \frac{\partial S}{\partial q^i}$.

§ 3.56 For conservative systems, H is time-independent. The action depends on time in the form

$$S(q,t) = S(q,0) - E\,t.$$

It follows that

$$H\left(q^1, q^2, ..., q^n, \frac{\partial S(q,0)}{\partial q^1},\ \frac{\partial S(q,0)}{\partial q^2},\ ...,\ \frac{\partial S(q,0)}{\partial q^n}\right) = E\ , \qquad (3.107)$$

which is the *time-independent Hamilton-Jacobi equation.*

The same happens whenever some integral of motion is known from the start. Each constant of motion is introduced as one of the constants. For instance, central potentials, for which the angular momentum J is a constant, will lead to a form

$$S(r,t) = S(r,0) - Et + J\phi.$$

These are particular cases, in which time or an angle are cyclic variables. In effect, suppose some coordinate $q_{(c)}$ is cyclic. This means that $q_{(c)}$ does not appear explicitly in the Hamiltonian nor, consequently in the Hamilton-Jacobi equation. The corresponding momentum is therefore constant, $p_{(c)}$ $= \frac{\partial S}{\partial q_{(c)}} = a_{(c)}$. It is an integral of motion, and $S = S(\text{remaining variables})$ $+ a_{(c)} q_{(c)}$.

§ 3.57 Of course, a suitable choice of coordinate system is essential to isolate a cyclic variable. For a particle in a central potential, spherical coordinates are the obvious choice. In the Hamiltonian

$$H = \frac{1}{2m}\left[p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2\theta}\right] + U(r)\ ,$$

the variables θ and ϕ, besides t, are absent. We shall use the knowledge that the angular momentum $J = mr^2\dot{\phi}$ is a constant, and start from the simpler planar Hamiltonian

$$H = \frac{m}{2}\left[\dot{r}^2 + r^2\dot{\phi}^2\right] + U(r) = \frac{p_r^2}{2m} + \frac{J^2}{2mr^2} + U(r)\ .$$

103

The time-independent Hamilton-Jacobi equation is then

$$\left(\frac{\partial S_r}{\partial r}\right)^2 + \frac{J^2}{r^2} = 2m(E - U(r)).$$

Thus,

$$S = -Et + J\phi + \int dr \sqrt{2m(E - U(r)) - \frac{J^2}{r^2}}.$$

Now, $\frac{\partial S}{\partial E} = C$ gives

$$t = \int \frac{m dr}{\sqrt{2m[E - U(r)] - \frac{J^2}{r^2}}} - C.$$

And $\frac{\partial S}{\partial J} = C'$ gives

$$\phi = C' + \int \frac{J\ dr}{r^2 \sqrt{2m[E - U(r)] - \frac{J^2}{r^2}}}.$$

The constants can be chosen $= 0$, fixing simply the origins of time and angle. The first equation,

$$t = \int \frac{m\ dr}{\sqrt{2m[E - U(r)] - \frac{J^2}{r^2}}}\ , \tag{3.108}$$

gives implicitly $r(t)$. The second,

$$\phi = \int \frac{J\ dr}{r^2 \sqrt{2m[E - U(r)] - \frac{J^2}{r^2}}}\ , \tag{3.109}$$

gives the trajectory.

We see that what is actually at work is an effective potential $U_{eff}(r) = U(r) + \frac{J^2}{2mr^2}$, including the angular momentum term. The values of r for which $U_{eff}(r) = E$ represent "turning points". If the function $r(t)$ is at first decreasing, it becomes increasing at that value, and vice versa.

§ 3.58 **Classical planetary motion** There are two general kinds of motion: limited (bound motion) and unlimited (scattering). We shall be concerned here only with the first case, in which $r(t)$ has a finite range $r_{min} \leq r(t) \leq$

r_{max}. In one turn, that is, in the time the variable takes to vary from r_{min} to r_{max} and back to r_{min}, the angle ϕ undergoes a change

$$\Delta\phi = 2 \int_{r_{min}}^{r_{max}} \frac{J}{r^2} \frac{dr}{\sqrt{2m[E - U(r)] - \frac{J^2}{r^2}}} \ . \qquad (3.110)$$

The trajectory will be closed if $\Delta\phi = 2\pi m/n$. A theorem (Bertrand's) says that this can happen only for two potentials, the Kepler potential $U(r) = -K/r$ and the harmonic oscillator potential $U(r) = Kr^2$. We shall here limit ourselves to the first case, which describes the keplerian motion of planets around the Sun. For $U(r) = -K/r$, Eq.(3.109) can be integrated to give

$$\phi = \arccos\left\{ \left[\frac{J}{r} - \frac{mK}{J}\right] \frac{1}{\sqrt{2mE + \frac{m^2 K^2}{J^2}}} \right\} \ . \qquad (3.111)$$

This trajectory corresponds to a closed ellipse. In effect, introduce the "ellipse parameter" $p = \frac{J^2}{mK}$ and the "eccentricity" $e = \sqrt{1 + \frac{2EJ^2}{mK^2}}$. Equation (3.111) can then be put into the form

$$r = \frac{p}{1 + e\ \cos\phi} \ , \qquad (3.112)$$

which is the equation for the ellipse. The above choice of the integration constants corresponds to $\phi = 0$ at $r = r_{min}$, which is the orbit *perihelium*. Suppose we add another potential (for example, $U' = K'/r^3$, with K' small. The orbit, as said above, will be no more closed. Staring from $\phi = 0$ at $r = r_{min}$, the orbit will reach the value $r = r_{min}$ at $\phi \neq 0$ in the first turn, and so on at each turn. The perihelium will change at each turn. This efect (called the *perihelium precession*) could come, for example, from a non-spherical form of the Sun. A turning gas sphere can be expected to be oblate. Observations of the Sun tend to imply that its oblateness is negligible, in any case insufficient to answer for the observed precession of Mercury's perihelium. We shall see later (§4.12) that General Relativity predicts a value in good agreement with observations.

Comment 3.7 We recall that the equation of the ellipse in cartesian coordinates is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $e = \frac{\sqrt{a^2 - b^2}}{a}$ and $p = b^2/a$. For a circle, $a = b$ and $e = 0$.

§ 3.59 We have already seen the main interest of the Hamilton-Jacobi formalism: in the relativistic case, the Hamilton-Jacobi equation (3.17) for a free particle coincides, for vanishing mass, with the eikonal equation (3.18). The formalism allows a unified treatment of test particles and light rays.

Chapter 4

Solutions

Einstein's equations are a nightmare for the searcher of solutions: a system of ten coupled non-linear partial differential equations. Its a tribute to human ingenuity that many (almost thirty to present time) solutions have been found. The non-linear character can be interpreted as saying that the gravitational field is able to engender itself. In consequence, the equations have non-trivial solutions even in the absence of sources. Actually, most known solutions are of this kind. We shall only examine a few examples, divided into two categories: "small scale solutions", which are of "local" interest, idealized models for stars (which gives an idea of what we mean by "small") and objects alike; and "large scale solutions", of cosmological interest.

4.1 Transformations

In tackling the big task of solving so difficult a problem, it is not surprising that the hunters have always supposed a high degree of symmetry. Let us begin with a short comment on symmetries of spacetimes.[*]

§ **4.1** Let us look for the condition for a vector field to generate a symmetry of the metric. Consider an infinitesimal transformation

$$x'^{\mu} = x^{\mu} + \varepsilon^{\mu}(x) \ , \quad \varepsilon^{\mu}(x) << 1 \ , \tag{4.1}$$

arbitrary as longs as ε is an arbitrary, though small, function of x. To the

---

[*] A treatment of the subject is given in the chapter 13 of S. Weinberg, *Gravitation and Cosmology*, J. Wiley, New York, 1972.

first order in ε,

$$\frac{\partial x'^{\mu}}{\partial x^{\lambda}} = \delta^{\mu}_{\lambda} + \frac{\partial \varepsilon^{\mu}}{\partial x^{\lambda}} \ .$$

Under such a tranformation, the metric components will change according to

$$g'^{\mu\nu}(x') = g^{\rho\sigma}(x) \, \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x'^{\nu}}{\partial x^{\sigma}} = g^{\rho\sigma}(x) \, \left(\delta^{\mu}_{\rho} + \frac{\partial \varepsilon^{\mu}}{\partial x^{\rho}}\right) \left(\delta^{\nu}_{\sigma} + \frac{\partial \varepsilon^{\nu}}{\partial x^{\sigma}}\right)$$

$$\approx g^{\mu\nu}(x) + g^{\mu\sigma}(x) \, \frac{\partial \varepsilon^{\nu}}{\partial x^{\sigma}} + g^{\nu\sigma}(x) \, \frac{\partial \varepsilon^{\mu}}{\partial x^{\sigma}} \ .$$

On the other hand, always keeping only the first-order terms,

$$g'^{\mu\nu}(x') = g'^{\mu\nu}(x + \varepsilon) \approx g'^{\mu\nu}(x) + \varepsilon^{\rho}(x)\partial_{\rho}g'^{\mu\nu}(x) \approx g'^{\mu\nu}(x) + \varepsilon^{\rho}(x)\partial_{\rho}g^{\mu\nu}(x) \ .$$

Equating both expressions,

$$g'^{\mu\nu}(x) + \varepsilon^{\rho}(x)\partial_{\rho}g^{\mu\nu}(x) = g^{\mu\nu}(x) + g^{\mu\sigma}(x) \, \frac{\partial \varepsilon^{\nu}}{\partial x^{\sigma}} + g^{\nu\sigma}(x) \, \frac{\partial \varepsilon^{\mu}}{\partial x^{\sigma}} \ ,$$

from which we obtain the variation of the metric components at a fixed point,

$$\bar{\delta}g^{\mu\nu}(x) = g'^{\mu\nu}(x) - g^{\mu\nu}(x) = g^{\mu\sigma}(x) \, \frac{\partial \varepsilon^{\nu}}{\partial x^{\sigma}} + g^{\nu\sigma}(x) \, \frac{\partial \varepsilon^{\mu}}{\partial x^{\sigma}} - \varepsilon^{\rho}(x)\frac{\partial g^{\mu\nu}}{\partial x^{\rho}} \ .$$
$$(4.2)$$

We now calculate the covariant derivative of ε^{μ}, conveniently separating the pieces comming from the Christoffel symbol:

$$\varepsilon^{\mu;\nu} = \partial^{\nu}\varepsilon^{\mu} - \tfrac{1}{2}\varepsilon^{\lambda}\partial_{\lambda}g^{\mu\nu} + \tfrac{1}{2}\varepsilon_{\sigma}(\partial^{\mu}g^{\nu\sigma} - \partial^{\nu}g^{\mu\sigma}) \ .$$

We see then that

$$\varepsilon^{\mu;\nu} + \varepsilon^{\nu;\mu} = \partial^{\mu}\varepsilon^{\nu} + \partial^{\nu}\varepsilon^{\mu} - \varepsilon^{\rho}\partial_{\rho}g^{\mu\nu}. \tag{4.3}$$

Therefore,

$$\bar{\delta}g^{\mu\nu}(x) = \varepsilon^{\mu;\nu} + \varepsilon^{\nu;\mu} \ . \tag{4.4}$$

This gives the change in the functional form of $g^{\mu\nu}$ under the transformation generated by the field $\varepsilon = \varepsilon^{\mu}\partial_{\mu}$. The condition for a symmetry is $\bar{\delta}g^{\mu\nu}(x) = 0$, that is,

$$\boxed{\text{Killing} \atop \text{equation}}$$

$$\varepsilon^{\mu;\nu} + \varepsilon^{\nu;\mu} = 0 \ . \tag{4.5}$$

108

This is the *Killing equation.* Fields satisfying it are called *Killing fields.* They generate transformations preserving the metric, which are called *isometries*, or *motions.* Applied to the Lorentz metric, the ten generators of the Poincaré group are found.[†]

§ **4.2** We shall here quote three theorems:

- the first says that the maximal number of isometries in a d-dimensional space is $d(d+1)/2$.[‡] Consequently, a given spacetime has at most 10 isometries.

- the second theorem says that this maximal number can only be attained if the scalar curvature R is a constant. There are only three kinds of spacetimes with 10 isometries: Minkowski spacetime, for which $R = 0$, and the two families of de Sitter spacetimes, one with $R > 0$ and the other with $R < 0$.

- a third theorem states that the isometries of a given metric constitute a group (group of isometries, or group of motions). The Poincaré group is the group of motions of Minkowski space.

§ **4.3** The converse procedure may be useful in the search of solutions: impose a certain symmetry from the start, and find the metrics satisfying Eq.(4.3) for the case,

$$\varepsilon^\rho \partial_\rho g^{\mu\nu} = \partial^\mu \varepsilon^\nu + \partial^\nu \varepsilon^\mu . \tag{4.6}$$

It should be said, however, that the Killing equation is still more useful in the study of the simmetries of a metric given *a priori.*

§ **4.4** The above procedure is a very particular case of a general and powerful method. How does a transformation acts on a manifold ? We are used to translations and rotations in Euclidean space. The same transformations, plus boosts and time translations, are at work on Minkowski space: they preserve the Lorentz metric. For these we use generators like, for example,

---

[†] See for example W.R. Davis & G.H. Katzin, Am. J. Phys. **30** (1962) 750.

[‡] L. P. Eisenhart, *Riemannian Geometry*, Princeton University Press, 1949.

$L^{\mu\nu} = x^\mu \partial^\nu - x^\nu \partial^\mu$, which is a vector field. This can be extended to general manifolds: given a group of transformations, each generator is represented on the manifold by a vector field X. This vector field presides on the in-finitesimal transformations undergone by every tensor field by the so called *Lie derivative*, an operation denoted L_X. The calculation above must be repeated for each type of tensor T: take a transformation like (4.1), compare $T'(x')$ obtained from the tensor behavior with $T'(x')$ obtained as a Taylor series, etc. The general result is

$$(L_X T)^{ab...r}_{ef...s} = X(T^{ab...r}_{ef...s}) - (\partial_i X^a)T^{ib...r}_{ef...s} - (\partial_i X^b)T^{ai...r}_{ef...s} - ... - (\partial_i X^r)T^{ab...i}_{ef...s}$$

$$+(\partial_e X^i)T^{ab...r}_{if...s} + (\partial_f X^i)T^{ab...r}_{ei...s} + ... + (\partial_s X^i)T^{ab...r}_{ef...i} . \qquad (4.7)$$

The requirement of invariance is $L_X T = 0$. For T a vector field, $L_X T$ is just the commutator:

$$L_X V = [X, V] .$$

The vector V is invariant with respect to the transformations engendered by X if it commutes with X.

Equation (4.2) is just the Lie derivative of $g^{\mu\nu}$ with respect to the field $\varepsilon = \varepsilon^\mu \partial_\mu$:

$$L_\varepsilon g^{\mu\nu}(x) = g'^{\mu\nu}(x) - g^{\mu\nu}(x) = g^{\mu\sigma}(x)\, \frac{\partial \varepsilon^\nu}{\partial x^\sigma} + g^{\nu\sigma}(x)\, \frac{\partial \varepsilon^\mu}{\partial x^\sigma} - \varepsilon^\rho(x)\frac{\partial g^{\mu\nu}}{\partial x^\rho} .$$
$$(4.8)$$

We shall not go into the subject in general.[§] Let us only state a property which holds when the tensor T is a differential form. For that we need a preliminary notion.

§ **4.5** Given a vector field X, the *interior product* of a p-form α by X is that (p-1)-form $i_X \alpha$, which, for any set of fields $\{X^1, X^2, \dots, X^{p-1}\}$, satisfies

$$i_X \alpha(X^1, X^2, \dots, X^{p-1}) = \alpha(X, X^1, X^2, \dots, X^{p-1}). \qquad (4.9)$$

If α is a 1-form, it gives simply its action on X:

$$i_X \alpha = \;< \alpha, X > \;= \alpha(X).$$

[§] A very detailed account can be found in B. Schutz,*Geometrical Methods of Mathematical Physics*, Cambridge University Press, Cambridge, 1985.

The interior product of X by a 2-form Ω is that 1-form satisfying $i_X\Omega = \Omega(X,Y)$ for any field Y. For a form of general degree, it is enough to know that, for a basis element,

$$i_X\left[\alpha^1 \wedge \alpha^2 \wedge \ldots \wedge \alpha^p\right] = \sum_{j=1}^{p}(-)^{j-1}\alpha^1 \wedge \alpha^2 \wedge \ldots [i_X\alpha^j] \wedge \ldots \wedge \alpha^p.$$

§ 4.6 The promised result is as follows: if ω is a differential form, then its Lie derivative has a simple expression in terms of the exterior derivative and the interior product:

$$L_X\omega = d[i_X\omega] + i_X[d\omega] .$$

Notice that L_X preserves the tensor character: it takes an r-covariant, s-contravariant tensor into another tensor of the same type.

4.2 Small Scale Solutions

Life is much simpler when a system of coordinates can be chosen so that invariance means just independence of some of the coordinates. In that case, Eq.(4.8) reduces to the last term and the intuitive property holds: the metric components are independent of those variables.

4.2.1 The Schwarzschild Solution

§ 4.7 Suppose we look for a solution of the Einstein equations which has spherical symmetry in the space section. This would correspond to central potentials in Classical Mechanics. It is better, in that case, to use spherical coordinates $(x^0, x^1, x^2, x^3) = (ct, r, \theta, \phi)$. This is one of the most studied of all solutions, and there is a standard notation for it. The interval is written in the form

$$ds^2 = e^\nu c^2 dt^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) - e^\lambda dr^2 . \tag{4.10}$$

The contravariant metric is consequently

$$g = (g_{\mu\nu}) = \begin{pmatrix} e^\nu & 0 & 0 & 0 \\ 0 & -e^\lambda & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix} \tag{4.11}$$

and its covariant counterpart,

$$
g^{-1} = (g^{\mu\nu}) = \begin{pmatrix} e^{-\nu} & 0 & 0 & 0 \\ 0 & -e^{-\lambda} & 0 & 0 \\ 0 & 0 & -r^{-2} & 0 \\ 0 & 0 & 0 & -r^{-2}\sin^{-2}\theta \end{pmatrix} . \qquad (4.12)
$$

We have now to build Einstein's equations. The first step is to calculate the components of the Levi-Civita connection, given by Eq.(2.40). Those which are non-vanishing are:

$$
\Gamma^0{}_{00} = \tfrac{1}{2}\,\tfrac{d\nu}{cdt} \; ; \quad \Gamma^0{}_{10} = \Gamma^0{}_{01} = \tfrac{1}{2}\,\tfrac{d\nu}{dr} \; ; \quad \Gamma^0{}_{11} = \tfrac{1}{2}\,e^{\lambda-\nu}\,\tfrac{d\lambda}{cdt} \; ;
$$

$$
\Gamma^1{}_{00} = \tfrac{1}{2}\,e^{\nu-\lambda}\,\tfrac{d\nu}{dr} \; ; \quad \Gamma^1{}_{01} = \Gamma^1{}_{10} = \tfrac{1}{2}\,\tfrac{d\lambda}{cdt} \; ; \quad \Gamma^1{}_{11} = \tfrac{1}{2}\,\tfrac{d\lambda}{dr} \; ;
$$

$$
\Gamma^1{}_{22} = -r\,e^{-\lambda} \; ; \quad \Gamma^1{}_{33} = -re^{-\lambda}\sin^2\theta \; ;
$$

$$
\Gamma^2{}_{12} = \Gamma^2{}_{21} = \tfrac{1}{r} \; ; \quad \Gamma^2{}_{33} = -\sin\theta\,\cos\theta \; ;
$$

$$
\Gamma^3{}_{13} = \Gamma^3{}_{31} = \tfrac{1}{r} \; ; \quad \Gamma^3{}_{23} = \Gamma^3{}_{32} = -\cot\theta \; . \qquad (4.13)
$$

As the second step, we must calculate the Ricci tensor of Eq.(2.48) and the Einstein tensor (2.59). We list those which are non-vanishing:

$$
G^0{}_0 = \tfrac{1}{r^2} - e^{-\lambda}\left[\tfrac{1}{r^2} - \tfrac{1}{r}\,\tfrac{d\lambda}{dr}\right] \; ; \quad G^1{}_0 = -\tfrac{1}{r}\,e^{-\lambda}\,\tfrac{d\lambda}{cdt} \; ; \quad G^1{}_1 = \tfrac{1}{r^2} - e^{-\lambda}\left(\tfrac{1}{r}\,\tfrac{d\nu}{dr} + \tfrac{1}{r^2}\right)
$$

$$
G^2{}_2 = G^3{}_3 = -\tfrac{1}{2}\,e^{-\lambda}\left[\tfrac{d^2\nu}{dr^2} + \tfrac{1}{2}\left(\tfrac{d\nu}{dr}\right)^2 + \tfrac{1}{r}\left(\tfrac{d\nu}{dr} - \tfrac{d\lambda}{dr}\right) - \tfrac{1}{2}\left(\tfrac{d\nu}{dr}\,\tfrac{d\lambda}{dr}\right)\right]
$$

$$
+ \tfrac{1}{2}\,e^{-\nu}\left(\tfrac{d^2\lambda}{c^2 dt^2} + \tfrac{1}{2}\left(\tfrac{d\lambda}{cdt}\right)^2 - \tfrac{1}{2}\,\tfrac{d\lambda}{cdt}\,\tfrac{d\nu}{cdt}\right) \; . \qquad (4.14)
$$

Each $G^\mu{}_\nu$ should now be imposed to be equal to $\tfrac{8\pi G}{c^4}\,T^\mu{}_\nu$. The source could be, for example, an electromagnetic field, in which case Eq.(3.35) would be used. Or the fluid inside a star, with the source given by Eq.(3.36) supplemented by an equation of state. Notice that the symmetry requirements made above would be satisfied not only by a static star, but also by a radially

112

pulsating one. We shall here consider the so-called "external", or "vacuum" solution for this case. We shall put $T^\mu{}_\nu = 0$, which is the case outside the star. The four differential equations following from $G^\mu{}_\nu = 0$ in (4.14) reduce in that case to only three (see Comment 4.1 below):

$$G^0{}_0 = 0 \quad \Rightarrow \quad e^{-\lambda} \left[\frac{1}{r^2} - \frac{1}{r} \frac{d\lambda}{dr} \right] = \frac{1}{r^2} \; ;$$

$$G^1{}_1 = 0 \quad \Rightarrow \quad e^{-\lambda} \left[\frac{1}{r^2} + \frac{1}{r} \frac{d\nu}{dr} \right] = \frac{1}{r^2} \; ; \qquad (4.15)$$

$$G^1{}_0 = 0 \quad \Rightarrow \quad \frac{d\lambda}{dt} = 0 \; .$$

A first result from Eqs.(4.15) is that λ is time-independent. Taking the difference between the first two equations shows that

$$\frac{d\lambda}{dr} = - \frac{d\nu}{dr} \; . \qquad (4.16)$$

Substituting this back in those equations lead to

$$e^\lambda = 1 + r \frac{d\nu}{dr} \; . \qquad (4.17)$$

Equation (4.16) says that $\lambda + \nu$ is independent of r, and is consequently a function of time alone: $\lambda + \nu = f(t)$. In the interval (4.10), it is always possible to redefine the time by an arbitrary transformation $t = \phi(t')$, which corresponds to adding an arbitrary function of t to ν. The choice of a new time coordinate $t' = \int_0^t e^{-f(t)/2} dt$ corresponds to changing $\nu \to \nu' = \nu + f(t)$. This means that it is always possible to choose the time coordinate so as to have $\lambda + \nu = 0$.

Comment 4.1 Time-independence of λ entails the vanishing of the last line in (4.14). Using (4.16) in (4.17) and taking the derivative implies that also the one-but-last line vanishes. This shows that the equation $G^2{}_2 = G^3{}_3 = 0$ is indeed redundant.

Integration of the only remaining equation, which is (4.17) rewritten with $\lambda = -\nu$,

$$e^{-\nu} = 1 + r \frac{d\nu}{dr} \; ,$$

leads then to

$$e^{-\lambda} = e^\nu = 1 - \frac{R_S}{r} \; ,$$

where R_S is a constant. Far from the source, when $r \to \infty$, we have $e^{-\lambda} = e^{\nu} \to 1$, so that the metric reduces to that of Minkowski space. Large values of r means a weak gravitational field. To fix the constant R_S, it is enough to impose that, at those values of r the solution reduce to the newtonian approximation, $g_{00} = 1 + 2V/c^2$, with $V = -GM/r$ and M the mass of source body. It follows that

$$R_S = \frac{2GM}{c^2} \ . \tag{4.18}$$

The interval (4.10) is therefore

$$ds^2 = \left(1 - \frac{2GM}{c^2\ r}\right) c^2 dt^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) - \frac{dr^2}{1 - \frac{2GM}{c^2\ r}} \tag{4.19}$$

$$= \left(1 - \frac{R_S}{r}\right) c^2 dt^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) - \frac{dr^2}{1 - \frac{R_S}{r}} \ . \tag{4.20}$$

This is the solution found by K. Schwarzschild in 1916, soon after Eintein had presented his final version of General Relativity. It describes the field caused, outside it, by a symmetrically spherical source. We see that there is a singularity in the metric components at the value $r = R_S$. The parameter R_S, given in Eq.(4.18), is called the *Schwarzschild radius*. Its value for a body with the mass of the Sun would be $R_S \approx 3$ km. For a body with Earth's mass, $R_S \approx 0.9$ cm. For such objects, of course, there exists to real Schwarzschild radius. It would be well inside their matter distribution, where $T_{\mu\nu} \neq 0$ and the solution is not valid.

§ **4.8** The above solution has been obtained in the absence of the cosmological constant. Its presence would change it to[¶]

$$ds^2 = \left(1 - \frac{2GM}{c^2\ r} - \frac{\Lambda}{3}r^2\right) c^2 dt^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) - \frac{dr^2}{1 - \frac{2GM}{c^2\ r} - \frac{\Lambda}{3}r^2}. \tag{4.21}$$

[¶] See the §96 of R.C. Tolman, *Relativity, Thermodynamics and Cosmology*, Dover, New York, 1987.

If we compare with Eq.(3.53), we find the potential

$$V = -\frac{MG}{r} - \frac{\Lambda c^2 r^2}{6} \ . \tag{4.22}$$

Eq.(3.58) would then lead to

$$\frac{d}{dt} v = -\frac{MG}{r^2} + \tfrac{1}{3} \Lambda c^2 r \ . \tag{4.23}$$

We recognize Newton's law in the first term of the righ-hand side. The extra, cosmological term has the aspect of a harmonic oscillator but, for $\Lambda > 0$, produces a repulsive force.

§ 4.9 The field, just as in the newtonian case, depends only on the mass M. At a large distance of any limited source, the field will forget details on its form and tend to have a spherical symmetry. The interval above is approximately given, at larges distances, by

$$ds^2 \approx c^2 dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) - \frac{R_S}{r} \left(dr^2 + c^2 dt^2 \right) \ . \tag{4.24}$$

The last term is a correction to the Lorentz metric and the above interval should be the asymptotic limit, for large values of r, of any field created by any source of limited size. We see that the Schwarzschild coordinate system used in (4.20) is "asymptotically Galilean": Schwarzschild's spacetime tends to Minkowski spacetime when $r \to \infty$.

As $g_{0j} = 0$ in Eq.(3.66), the 3-dimensional space sector induced by (4.20) will have the interval

$$d\sigma^2 = \frac{dr^2}{1 - \frac{R_S}{r}} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ , \tag{4.25}$$

to be compared with the Euclidean interval

$$d\sigma^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{4.26}$$

At fixed θ and ϕ, that is, radially, the distance between two points P and Q standing outside the Schwarzschild radius will be

$$\int_P^Q \frac{dr}{\sqrt{1 - \frac{R_S}{r}}} > r_Q - r_P \ . \tag{4.27}$$

On the other hand, the proper time will be

$$d\tau = \sqrt{g_{00}}\; dt = \sqrt{1 - \frac{R_S}{r}}\; dt < dt \; . \qquad (4.28)$$

We see that $d\tau = dt$ when $r \to \infty$. And we see also that, at finite distances from the source, time "marches slower" than time at infinity. This difference between proper time and coordinate time arrives at an extreme case near the Schwarzschild radius.

§ **4.10** We can make some checking on the results found. Given the metric

$$\begin{pmatrix} 1 - \frac{R_S}{r} & 0 & 0 & 0 \\ 0 & -\frac{1}{1-\frac{R_S}{r}} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix} ,$$

we can proceed to the laborious computation of the Christoffeln and Riemann components. We find, for example,

$$R^1{}_{212} = \frac{R_S}{r^2\;(r-R_S)} \; ; \quad R^1{}_{313} = -\frac{R_S}{2\;r} \; ; \quad R^1{}_{414} = -\frac{R_S\;\sin^2\theta}{2\;r}$$

$$R^2{}_{323} = -\frac{R_S}{2\;r} \; ; \quad R^2{}_{424} = \frac{(R_S - 2\;r)\;\sin^2\theta}{2\;r} \; ; \quad R^3{}_{434} = -\cos^2\theta + \frac{R_S\;\sin^2\theta}{r} \; .$$

All components of the Ricci tensor vanish, as they should for an exterior, $T_{\mu\nu} = 0$ solution. The scalar curvature, of course, vanishes also. This is a good illustration of the statement made in § 3.25, by which $T_{\mu\nu} = 0$ implies $R_{\mu\nu} = 0$ but not necessarily $R^\rho{}_{\sigma\mu\nu} = 0$. It is also a good example of a fundamental point of General Relativity: it is the non-vanishing of $R^\rho{}_{\sigma\mu\nu}$ that indicates the presence of a gravitational field.

§ **4.11** We can, furthermore, examine the space section. With coordinates (r, θ, ϕ), the metric is

$$\begin{pmatrix} \frac{1}{1-\frac{R_S}{r}} & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix} \; .$$

The Christoffeln form the matrices

$$(\Gamma^1{}_{ij}) = \begin{pmatrix} \frac{R_S}{2r(R_S-r)} & 0 & 0 \\ 0 & R_S - r & 0 \\ 0 & 0 & (R_S - r)\sin^2\theta \end{pmatrix}$$

$$(\Gamma^2{}_{ij}) = \begin{pmatrix} 0 & \frac{1}{r} & 0 \\ \frac{1}{r} & 0 & 0 \\ 0 & 0 & -\sin\theta\cos\theta \end{pmatrix}$$

$$(\Gamma^3{}_{ij}) = \begin{pmatrix} 0 & 0 & \frac{1}{r} \\ 0 & 0 & \cot\theta \\ \frac{1}{r} & \cot\theta & 0 \end{pmatrix} .$$

The Ricci tensor is given by

$$(R_{ij}) = \begin{pmatrix} \frac{R_S}{r^2(R_S-r)} & 0 & 0 \\ 0 & \frac{R_S}{2r} & 0 \\ 0 & 0 & \frac{R_S \sin^2\theta}{2r} \end{pmatrix} .$$

Thus, the Ricci tensor of the space sector is non-trivial. The scalar curvature, however, is zero.

§ 4.12 Perihelium precession The Hamilton-Jacobi equation (3.17) and the eikonal equation (3.18) provide, as we have seen, a unified approach to the trajectories of massive particles and light rays. Let us first examine the motion of a particle of mass m in the above gravitational field. As angular momentum is conserved, it will be a plane motion, with constant θ. For reasons of simplicity, we shall choose the value $\theta = \pi/2$. With the metric (4.19), the Hamilton-Jacobi equation $g^{\mu\nu}(\partial_\mu S)(\partial_\nu S) = m^2 c^2$ acquires the form

$$\left(1 - \frac{R_S}{r}\right)^{-1}\left(\frac{\partial S}{\partial ct}\right)^2 - \left(1 - \frac{R_S}{r}\right)\left(\frac{\partial S}{\partial r}\right)^2 - \frac{1}{r^2}\left(\frac{\partial S}{\partial \phi}\right)^2 - m^2 c^2 = 0 .$$

(4.29)

The solution is looked for by the Hamilton-Jacobi method described in Section 3.9. With some constant energy E and constant angular momentum J, we write

$$S = -Et + J\phi + S_r(r).$$

(4.30)

117

This, once inserted in (4.29), gives

$$S_r = \int dr \left\{ \frac{E^2}{c^2} \left(1 - \frac{R_S}{r} \right)^{-2} - \left(m^2 c^2 + \frac{J^2}{c^2} \right) \left(1 - \frac{R_S}{r} \right)^{-1} \right\}^{1/2} . \quad (4.31)$$

By the method, $r = r(t)$ is obtained from the equation $\frac{\partial S}{\partial E} = $ constant, from which comes

$$ct = \frac{E}{mc^2} \int \frac{dr}{\left(1 - \frac{R_S}{r} \right) \sqrt{ \left(\frac{E}{mc^2} \right)^2 - \left(1 + \frac{J^2}{m^2 c^2 r^2} \right) \left(1 - \frac{R_S}{r} \right) }} . \quad (4.32)$$

The trajectory is found from $\frac{\partial S}{\partial J} = $ constant, which gives

$$\phi = \int \frac{J}{r^2} \frac{dr}{\sqrt{ \frac{E^2}{c^2} - \left(m^2 c^2 + \frac{J^2}{r^2} \right) \left(1 - \frac{R_S}{r} \right) }} . \quad (4.33)$$

This leads to an elliptic integral. We are putting the additive integration constants, which merely fix the origins of the coordinates ct and ϕ, equal to zero.

We should compare the above results with their non-relativistic counterparts given in Eqs.(3.108), (3.109). However, in order to calculate the small corrections given by the theory to the trajectories of the planets turning around the Sun, it is wiser to make approximations in (4.31) before taking the derivative $\frac{\partial S}{\partial J}$. We shall suppose radial distances very large with respect to the Schwarzschild radius: $R_S << r$. We also change the integration variable to $r' = \sqrt{r^2 - r R_S}$ (and drop the pirmes afterwards). Writing E' for the non-relativistic energy, we find

$$S_r =$$

$$\int dr \left\{ \frac{E'^2}{c^2} + 2mE' + \frac{1}{r} \left(2m^2 MG + 4E' M R_S \right) - \frac{1}{r^2} \left(J^2 - \tfrac{3}{2} m^2 c^2 R_S^2 \right) \right\}^{1/2}$$
$$(4.34)$$

The term in $1/r^2$ will produce a secular displacement of the orbit perihelium. The remaining terms cause changes in the relationships between the four-momentum of the particle and the newtonian ellipse. We shall be interested only in the perihelium precession. The trajectory is determined by $\phi + \frac{\partial S_r}{\partial J}$

118

= constant. The variation of S_r in one revolution is, in the approximation considered,

$$\Delta S_r = \Delta S_r^{(0)} - \frac{3m^2c^2R_S^2}{4J} \frac{\partial \Delta S_r}{\partial J} \ .$$

$\Delta S_r^{(0)}$ is the closed ellipse case. The variation of the angle ϕ in one revolution will be

$$\Delta \phi = -\frac{\partial \Delta S_r}{\partial J} \ .$$

Taking into account that

$$-\frac{\partial \Delta S_r^{(0)}}{\partial J} = \Delta \phi^{(0)} = 2\pi \ ,$$

we find

$$\Delta \phi = 2\pi + \frac{3\pi m^2 c^2 R_S^2}{2J^2} = 2\pi + \frac{6\pi G^2 m^2 M^2}{c^2 J^2} \ .$$

The last piece gives the precession $\delta\phi$. It is usual to express it in terms of the ellipse parameters. If the great axis is a and the eccentricity is e, we have

$$\frac{J^2}{GMm^2} = a(1-e^2) \ .$$

The perihelium precession is then

$$\delta\phi = \frac{6\pi GM}{a(1-e^2)c^2} \ .$$

For the Earth, this is a very small variation: in seconds of arc, $3.8''$ per century. For Mercury, it is $43.0''$ per century. This is in good agreement with measurements.

§ **4.13 Light-ray deviation** The eikonal equation (3.18) is just the Hamilton-Jacobi equation (3.17) with $m = 0$. The trajectory will still be given by Eq.(4.33). The interpretation is, of course, quite another. S is now the eikonal, the energy E must be replaced by the light frequency ω_0, and it is convenient to introduce a new constant, the impact parameter $\rho = Jc/\omega_0$. Thus,

$$\phi = \int \frac{dr}{r^2 \sqrt{\frac{1}{\rho^2} - \frac{1}{r^2}\left(1 - \frac{R_S}{r}\right)}} \ . \tag{4.35}$$

119

This gives $r = \rho/\cos\phi$ — a straight line passing at a distance ρ of the coordinate origin — in the non-relativistic case $R_S = 0$. The procedure to analyse the small corrections due to $R_S \neq 0$ is analogous to that used for $m \neq 0$. We go back to Eq. (4.31),

$$S_r(r) = \frac{\omega_0}{c} \int dr \left\{ r^2 (r - R_S)^{-2} - \rho^2 \left(r^2 - rR_S \right)^{-1} \right\}^{1/2} . \qquad (4.36)$$

With the same transformations used previously, this becomes

$$S_r(r) = \frac{\omega_0}{c} \int dr \sqrt{1 - \rho^2 r^{-2} + 2R_S r^{-1}} . \qquad (4.37)$$

Expanding in powers of $R_S r^{-1}$,

$$S_r \approx S_r^{R_S=0} + \frac{R_S \omega_0}{c} \int \frac{dr}{\sqrt{r^2 - \rho^2}} = S_r^{R_S=0} + \frac{R_S \omega_0}{c} \ \text{arccosh} \ \frac{r}{\rho} . \qquad (4.38)$$

The deviation undergone by a ray coming from a large distance R down to a distance ρ and then again to the same distance R will be

$$\Delta S_r = \Delta S_r^{R_S=0} + 2 \ \frac{R_S \omega_0}{c} \ \text{arccosh} \ \frac{r}{\rho} . \qquad (4.39)$$

To get the variation in the angle ϕ, it is enough to take the derivative with respect to $J = \rho \omega_0 / c$:

$$\Delta\phi = -\frac{\partial \Delta S_r}{\partial J} = -\frac{\partial \Delta S_r^{R_S=0}}{\partial J} + 2 \ \frac{R_S R}{\rho \sqrt{R^2 - \rho^2}} . \qquad (4.40)$$

The term corresponding to the straight line has $\Delta\phi = \pi$. Taking the asymptotic limit $R \to \infty$,

$$\Delta\phi = \pi + 2 \ \frac{R_S}{\rho} . \qquad (4.41)$$

This gives a deviation towards the centre of an angle

$$\delta\phi = 2 \ \frac{R_S}{\rho} = \frac{4GM}{\rho c^2}. \qquad (4.42)$$

For a light ray grazing the Sun, this gives $\delta\phi = 1.75''$. This effect has been observed by a team under the leadership of Eddington during the 1919 solar eclipse, at Sobral. It has been considered the first positive experimental test of General Relativity.

§ 4.14 The event horizon We can compare the energy of the particle as seen by an observer using the Schwarzschild coordinate system with the its energy as seen in the proper frame. The first is, as previously seen, $E = -\frac{\partial S}{\partial t}$, and the second is $E_0 = -\frac{\partial S}{\partial \tau} = mc^2$. But $\frac{\partial S}{\partial t} = \frac{\partial \tau}{\partial t}\frac{\partial S}{\partial \tau}$, so that $E = \sqrt{g_{00}}\, mc^2$, or

$$E = mc^2 \sqrt{1 - \frac{R_S}{r}}\ . \tag{4.43}$$

Let us examine the case of a particle falling in purely radial motion ($\theta = 0$, $\phi = 0$, $\therefore$ $J = 0$) towards the center $r = 0$. If it starts from a point r_0 at the instant t_0, its energy will be $E = mc^2 \sqrt{1 - \frac{R_S}{r_0}}$. At a moment t and a distance r, Eq.(4.32) gives

$$c(t - t_0) = \sqrt{1 - \frac{R_S}{r_0}} \int_r^{r_0} \frac{dr'}{\left(1 - \frac{R_S}{r'}\right)\sqrt{\frac{R_S}{r'} - \frac{R_S}{r_0}}}\ . \tag{4.44}$$

This coordinate time diverges when $r \to R_S$. Seen from an external observer, the particle will take an infinite time to arrive at the Schwarzschild radius. On the other hand, we can calculate the proper interval of time for the same thing to happen: it will be

$$c(\tau - \tau_0) = \int_r^{r_0} ds = \int_r^{r_0} dr' \sqrt{g_{00}c^2 \left(\frac{dt}{dr'}\right)^2 + g_{11}}\ .$$

From (4.44),

$$\frac{dt}{dr} = - \sqrt{1 - \frac{R_S}{r_0}}\ \frac{dr}{\left(1 - \frac{R_S}{r}\right)\sqrt{\frac{R_S}{r} - \frac{R_S}{r_0}}}$$

$$\therefore\ g_{00}c^2 \left(\frac{dt}{dr}\right)^2 + g_{11} = \frac{1}{\frac{R_S}{r} - \frac{R_S}{r_0}}\ .$$

Thus,

$$c(\tau - \tau_0) = \int_r^{r_0} \frac{dr'}{\sqrt{\frac{R_S}{r'} - \frac{R_S}{r_0}}}\ . \tag{4.45}$$

This is a convergent integral, giving a finite value when $r \to R_S$. The particle, looking at things from its own proper frame and measuring time in

its own clock, arrives at the Schwarzschild radius in a finite interval of time. It will even arrive at the center $r = 0$ in finite time.

Suppose now a gas of particles, each one in the conditions above. They will all fall towards the center. Each will do it in a finite interval of time in its own frame. The gas will eventually colapse. A quite distinct picture will be seen from the coordinate, asymptotically flat frame. Seen from a distant observer, the particles will never actually traverse the Schwarzschild radius. All that happens inside the Schwarzschild sphere lies "beyond the infinity of time". The sphere is a *horizon*, technically called an *event horizon*.

We have seen (in § 3.37 and ensuing paragraphs) how coordinate time and proper time are related. Here we have an extreme difference, given by Eq.(4.28). Consider a light source near the Schwarzschild radius. Seen from a distant observer, it will be strongly red-shifted. It will actually be more and more red-shifted as the source is closer and closer to the radius. The red-shift will tend to infinity at the radius itself: the external observer cannot receive any signal from the sphere surface.

§ **4.15** In consequence of all that has been said, a massive object can eventually fall inside its own Schwarzschild sphere. In a real star, such a gravitational colapse is kept at bay by the centrifugal effects of pressure caused by the energy production through nuclear fusion. A massive enough star can, however, collapse when its energy sources are exhausted. Once inside the radius, no emission will be able to scape and reach an external observer. Particles and radiation will go on falling down to the sphere, but nothing will get out. Such a collapsed object, such a *black hole*, will only be observable by indirect means, as the emission produced by those external particles which are falling down the gravitational field.

§ **4.16** It should be said, however, that the singularity in the *components* of the metric does not imply that the metric itself, an invariant tensor, be singular. A real singularity must be independent of coordinates and should manifest itself in the invariants obtained from the metric. For example, the determinant is an invariant. It is $g = -r^4 \sin^2 \theta$. This shows that the point $r = 0$ is a real singularity, in which the metric is no more invertible. The Schwarzschild sphere, however, is not. It is, as seen, an event horizon, but

122

not a singularity. This seems to have been first noticed by Lemaître in 1938, and can be verified by transforming to other coordinate systems. Take for instance the family of transformations

$$ct' = \pm ct \pm \int \frac{f(r)dr}{1 - \frac{R_S}{r}} \; ; \quad r' = ct + \int \frac{dr}{(1 - \frac{R_S}{r})f(r)} \; , \qquad (4.46)$$

involving an arbitrary function $f(r)$ and which lead to

$$ds^2 = \frac{1 - \frac{R_S}{r}}{1 - f(r)^2} \, (c^2 dt'^2 - f^2 dr'^2) - r^2(d\theta^2 + \sin^2\theta d\phi^2) \; .$$

The Schwarzschild singularity will disappear for an f such that $f(\frac{R_S}{r}) = 1$. The better choice is $f(r) = \sqrt{\frac{R_S}{r}}$, which gives a synchomous system. Notice that there are two possible choices of the signs in (4.46). One leads to an expanding reference system, the other to a contracting frame. In effect, the upper signs in (4.46) give

$$r' - ct' = \int dr \, \frac{(1 - f(r)^2)}{(1 - \frac{R_S}{r})f(r)} \; = \int \frac{dr}{f(r)} = \int dr \frac{r}{R_S} = \frac{2r^{3/2}}{R_S^{1/2}} \; .$$

This shows that the system contracts in the old system:

$$r = R_S^{1/3} \left[\frac{3}{2}(r' - ct') \right]^{2/3} \; . \qquad (4.47)$$

As $r \geq 0$, $r' \geq ct'$, the equality corresponding to the center real singularity. The singularity would correspond to the value

$$\frac{3}{2} \, (r' - ct') = R_S \; .$$

The interval takes the form given by Lemaître

$$ds^2 = c^2 dt'^2 - \frac{dr'^2}{\left[\frac{3}{2R_S}(r' - ct') \right]^{2/3}} - \left[\frac{3}{2}(r' - ct') \right]^{4/3} R_S^{2/3}(d\theta^2 + \sin^2\theta d\phi^2) \; .$$

$$\qquad (4.48)$$

The Schwarzschild singularity does not turn up. To get a feeling, we can use Eq.(4.47) to rewrite the interval in mixed coordinates:

$$ds^2 = c^2 dt'^2 - \frac{R_S}{r} \, dr'^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \; . \qquad (4.49)$$
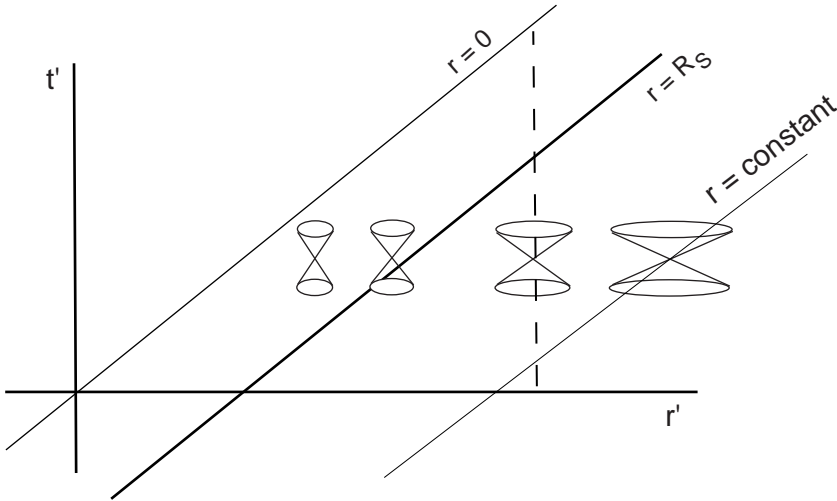
Figure 4.1: *In Lemaître coordinates, a particle can go through the Schwarzschild radius in a finite amount of time. The cones become narrower as r′ decreases.*

By Eq.(4.47), to each value $r = $ constant in the old coordinates corresponds a straight line $r' = a + ct'$ in the new system (see Figure 4.1). These lines indicate, consequently, immobile particles in the old reference frame.

Lemaître's system is synchonous (§ 3.44). Geodesics are represented by vertical lines, as the dashed line in the diagram. A free particle can fall through the Schwarzschild radius and attain the center at $r = 0$.

Light cones have an interesting behavior. Consider a radial ($d\theta = 0, d\phi = 0$) light ray. Equation (4.49) gives, for $ds = 0$, the two (future and past) cones, one for each sign in

$$c \, \frac{dt'}{dr'} = \pm \sqrt{\frac{R_S}{r}} \,. \tag{4.50}$$

We see that the cone solid angle becomes smaller for smaller values of r. Consider again the lines $r = $ constant, indicating immobility in the original system of coordinates. Their inclination is $c \, \frac{dt'}{dr'} = 1$, and there are two possibilities:

for $r > R_S$**,** then $\left| c \, \frac{dt'}{dr'} \right| < 1$: a line $r = $ constant through the vertex lies inside the cone; immobility is consequently causally possible;

124

for $r < R_S$**,** then $\left| c \frac{dt'}{dr'} \right| > 1$: a line $r =$ constant through the vertex lies outside the cone; immobility is causally impossible; any particle will fall to the center.

The cones defined in (4.50) have one more curious behavior: they will deform themselves progressively as they approach the center. The derivative becomes infinite at $r = 0$. The lines representing the cones in the Figure will meet the line $r = 0$ as vertical lines.

If we choose the lower signs in (4.46), the line element will be (4.48), but with $t' \rightarrow - t'$:

$$ds^2 = c^2 dt'^2 - \frac{dr'^2}{\left[\frac{3}{2R_S}(r' + ct') \right]^{2/3}} - \left[\frac{3}{2}(r' + ct') \right]^{4/3} R_S^{2/3}(d\theta^2 + \sin^2 \theta d\phi^2) \; .$$

(4.51)

An analogous discussion can be done concerning the behavior of cones and the issue of immobility. Immobility is still forbidden inside the radius. The difference is that, instead of falling fatally towards the center, a particle will inexorably draw away from it. Contrary to the previous case, the system expands in the old system:

$$r = R_S^{1/3} \left[\frac{3}{2}(r' + ct') \right]^{2/3} \; .$$

(4.52)

§ 4.17 A reference frame is *complete* when the world line of every particle either go to infinity or stop at a true singularity. In this sense, neither of the above coordinate systems is complete. The Schwarzschild coordinates do not apply to the interior of the sphere. An outside particle in the contracting Lemaître system can only fall down towards the centre: initial conditions in the opposite sense are not allowed. Just the contrary happens in the expanding Lemaître systems. Both leave some piece of space unattainable. That a complete system of coordinates does exist was first shown by Kruskal and Fronsdal. We shall here only mention a few aspects of this question.[||] In such kind of system no singularity at all appears in the Schwarzschild radius. The coordinates are given as implicit functions of those used above.

In a form given by Novikov, the metric is looked for in a form generalizing the mixed-coordinate interval (4.49). New time and radial variables τ, R are defined so as to put the line element in the form

$$ds^2 = c^2 d\tau^2 - \left[1 + \frac{R_S^2}{R^2}\right] dr^2 - r^2(\tau, R)(d\theta^2 + \sin^2\theta d\phi^2) \qquad (4.53)$$

and supposing a dust gas as source. The coordinates τ, R are given implicitly and in parametric form by

$$r = \frac{R_S}{2}\left[1 + \frac{R_S^2}{R^2}\right](1 - \cos\eta); \qquad (4.54)$$

$$\tau = \frac{R_S}{2}\left[1 + \frac{R_S^2}{R^2}\right]^{3/2}(\pi - \eta + \sin\eta). \qquad (4.55)$$

The parameter η take values in the interval $2\pi, 0$. When it runs from 2π to 0 the time variable increases monotonically, while r increases from zero up to a maximum value

$$r = R_S\left[1 + \frac{R_S^2}{R^2}\right] \qquad (4.56)$$

and then decreases back to zero. The *Kruskal diagram* of Figure 4.2 sumarizes the whole thing. Coordinates τ, R are complete, so that all situations are described. The Schwarzschild coordinates describe only situations external to the line $r = R_S$. Contracting Lemaître coordinates cover the shaded area, expanding coordinates cover the domain which appears shaded after specular vertical reflection. The small arrows indicate the forcible sense particles follow inside the Schwarzschild sphere: contracting in the upper side, expanding in the lower one.

We have seen in § 3.17 that dust particles follow geodesics. The system is synchronous (§ 3.44), so that such geodesics are the vertical lines $R =$ constant in the diagram. An example is shown as a dashed line. Starting at $\tau = 0$

$^{\parallel}$ Details are given in an elegant form in L.D. Landau & E.M. Lifshitz, *Théorie des Champs*, 4th french edition, § 103.

126
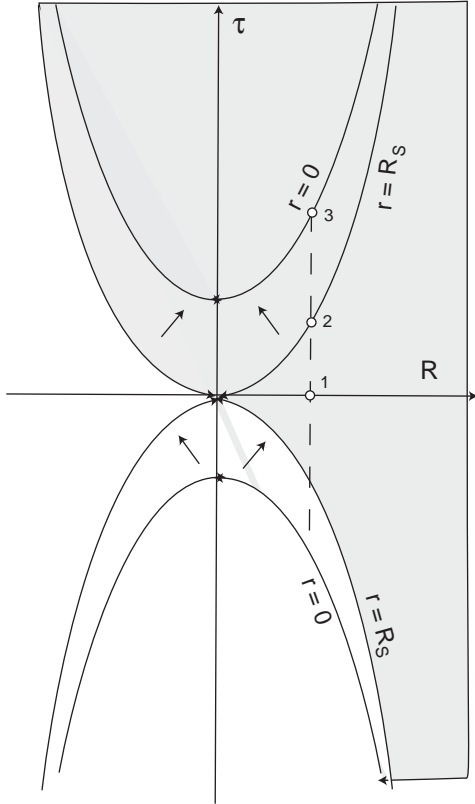
Figure 4.2: *Kruskal diagram.*

(point 1 in the Figure), a particle attains the center $r = 0$ in a finite time

$$\tau = \frac{\pi R_S}{2}\left[1 + \frac{R_S^2}{R^2}\right]^{3/2}. \tag{4.57}$$

Starting with outward initial conditions at the center, a situation corresponding to the lower part of the diagram, a particle will follow a trajectory like the dashed line. It will cross the Schwarzschild radius in the outward direction, attain a maximal coordinate distance given by (4.56) at $\tau = 0$ (point 1 again), and then fall back, crossing the Schwarzschild radius in an inward progression at point 2 and reaching the center at point 3.

4.3 Large Scale Solutions

§ 4.18 Two of the four fundamental interactions of Nature — the weak and the strong — are of very short range. Electromagnetism has a long range but, as opposite charges attract each other and tend thereby to neutralize, the field is, so to speak, "compensated" at large distances. Gravitation remains as the only uncompensated field and, at large scales, dominates. This is why Cosmology is deeply involved with it. We shall now examine two of the main solutions of Einstein's equations which are of cosmological import. The Friedmann solution provides the background for the so-called Standard Cosmological Model, which — despite some difficulties — gives a good description of the large-scale Universe during most of its evolution. It represents a spacetime where time, besides being separated from space, is position–independent, and space is homogeneous and isotropic at each point. The Universe "begins" with a high-density singularity (the "big-bang") and evolves through two main periods, one radiation-dominated and another matter-dominated, which lasts up to present time. There are difficulties both at the beginning and at present time. The first problem is mainly related to causality, and could be solved by a dominant cosmological term. The second comes from the recent observation that a cosmological term is, even at present time, dominant. It is consequently of interest to examine models in which the cosmological term gives the only contribution. These are the de Sitter universes, which have an additional theoretical advantage: all calculation are easily done. A more detailed account of both the Friedmann and the de Sitter solutions, mainly concerned with cosmological aspects, can be found in the companion notes *Physical Cosmology.*

> **Standard model**

4.3.1 The Friedmann Solutions

§ 4.19 We have seen under which conditions the second term in the interval $ds^2 = g_{00}dx^0dx^0 + g_{ij}dx^idx^j$ represents the 3-dimensional space. We shall suppose g_{00} to be space–independent. In that case, the coordinate x^0 can be chosen so that the time piece is simply c^2dt^2, where t will be the coordinate time. It will be a "universal time", the same at every point of space.

The "Universe", that is, the space part, is supposed to respect the Cos-

mological Principle, or Copernican Principle: it is homogeneous as a whole. Homogeneous means looking the same at each point. Once this is accepted, imposing isotropy around one point (for instance, that point where we are) is enough to imply isotropy around every point. This means in particular that space has the same curvature around each point. There are only 3 kinds of 3–dimensional spaces with constant curvature:

- the sphere S^3, a closed space with constant positive curvature;

- the open hyperbolic space $S^{2,1}$, or (a pseudo–sphere, or sphere with imaginary radius), whose curvature is negative; and

- the open euclidean space $\mathbb{E}^3$ of zero curvature (that is, flat).

These three types of space are put together with the help of a parameter k: $k = +1$ for S^3, $k = -1$ for $S^{2,1}$ and $k = 0$ for $\mathbb{E}^3$. The 3–dimensional line element is then, in convenient coordinates,

$$dl^2 = \frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \ . \tag{4.58}$$

The last two terms are simply the line element on a 2-dimensional sphere S^2 of radius r — a clear manifestation of isotropy. Notice that these symmetries refer to space alone. The "radius" can be time-dependent.

§ **4.20** The energy content is given by the energy-momentum of an ideal fluid, which in the Standard Model is supposed to be homogeneous and isotropic:

$$T_{\mu\nu} = (p + \rho c^2) \ U_\mu U_\nu - p \ g_{\mu\nu} \ . \tag{4.59}$$

Here U_μ is the four-velocity and $\rho = \epsilon/c^2$ is the mass equivalent of the energy density. The pressure p and the energy density are those of the matter (visible plus "dark") and radiation. When $p = 0$, the fluid reduces to "dust".

§ **4.21** We can now put together all we have said. Instead of a time–dependent radius, it is more convenient to use fixed coordinates as above,

and introduce an overall scale parameter $a(t)$ for 3–space, so as to have the spacetime line element in the form

$$ds^2 = c^2 dt^2 - a^2(t) dl^2. \qquad (4.60)$$

Thus, with the high degree of symmetry imposed, the metric is entirely fixed by the sole function $a(t)$. The spacetime line element will be

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right] . \qquad (4.61)$$

This is the Friedmann–Robertson–Walker line element.

§ **4.22** The above line element is a pure consequence of symmetry considerations. We have now to impose the dynamical equations. Recent data, as said, point to a non-vanishing cosmological constant. It is, consequently, wiser to use the Einstein equations in the form (3.42). The extreme simplicity of the model is reflected in the fact that those 10 partial differential equations reduce to 2 ordinary differential equations (in the variable t) for the scale parameter.

In effect, once (4.59) is used, the field equations (3.42) reduce to the two Friedmann equations for $a(t)$:

$$\dot{a}^2 = \left[2 \left(\frac{4\pi G}{3} \right) \rho + \frac{\Lambda c^2}{3} \right] a^2 - kc^2 \ ; \qquad (4.62)$$

$$\ddot{a} = \left[\frac{\Lambda c^2}{3} - \frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2} \right) \right] a(t) . \qquad (4.63)$$

It will be convenient to absorb the length dimensionality in $a(t)$, so that the variable r can be seen as dimensionless. The cosmological constant has the dimension $(length)^{-2}$. The second equation determines the concavity of the function $a(t)$. This has a very important qualitative consequence when $\Lambda = 0$. In that case, for normal sources with $\rho > 0$ and $p \geq 0$, $\ddot{a}$ is forcibly negative for all t and the general aspect of $a(t)$ is that of Figure 4.3. It will consequently vanish for some time $t_{initial}$. Distances and volumes vanish at that time and densities become infinite. This moment $t_{initial}$ is taken as *the*

beginning, the "Big Bang" itself. It is usual to take $t_{initial}$ as the origin of the time coordinate: $t_{initial} = 0$. If $\Lambda > 0$, there is a competition between the two terms. It may even happen that the scale parameter be $\neq 0$ for all finite values of t.
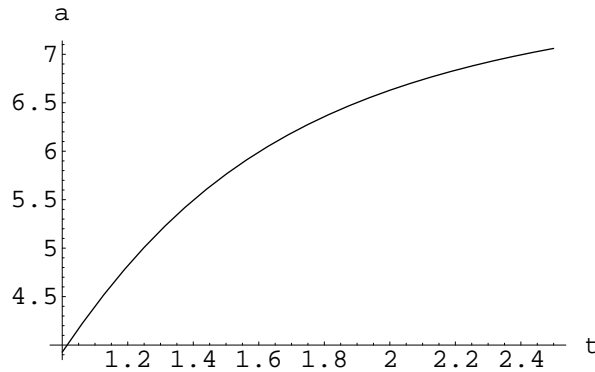
Figure 4.3: *Concavity of $a(t)$ for $\Lambda = 0$.*

Combining both Friedmann equations, we find

$$\frac{d\rho}{dt} = -3 \frac{\dot{a}}{a} \left(\rho + \frac{p}{c^2} \right) , \qquad (4.64)$$

which is equivalent to

$$\frac{d}{da}(\epsilon a^3) + 3 \, p \, a^2 = 0. \qquad (4.65)$$

This equation reflects the energy–momentum conservation: it can be alternatively obtained from $T^{\mu\nu}{}_{;\nu} = 0$.

§ **4.23** It is convenient to introduce the Hubble function

$$H(t) = \frac{\dot{a}(t)}{a(t)} = \frac{d}{dt} \ln a(t) , \qquad (4.66)$$

whose present-day value is the *Hubble constant*

$$H_0 = 100 \, h \, km \, s^{-1} \, Mpc^{-1} = 3.24 \times 10^{-18} \, h \, s^{-1} .$$

<div style="text-align:right">

Hubble function & constant

</div>

131

The parameter h, of the order of unity, encapsulates the uncertainty in present-day measurements, which is large ($0.45 \leq h \leq 1$). Another function of interest is the deceleration

$$q(t) = -\frac{\ddot{a}a}{\dot{a}^2} = -\frac{\ddot{a}}{\dot{a}H(t)} = -\frac{1}{H^2(t)}\frac{\ddot{a}}{a} . \qquad (4.67)$$

Equivalent expressions are

$$\dot{H}(t) = -H^2(t)(1 + q(t)) ; \qquad \frac{d}{dt}\frac{1}{H(t)} = 1 + q(t) . \qquad (4.68)$$

Uncertainty is very large for the *deceleration parameter*, which is the present–day value $q_0 = q(t_0)$. Data seem consistent with $q_0 \approx 0$. Notice that we are using what has become a standard notation, the index "0" for present-day values: H_0 for the Hubble constant, t_0 for present time, etc.

The Hubble constant and the deceleration parameter are basically integration constants, and should be fixed by initial conditions. As previously said, the present–day values are used.

§ 4.24 The flat Universe Let us, as an exercise, examine in some detail the particular case $k = 0$. The Friedmann–Robertson–Walker line element is simply

$$ds^2 = c^2dt^2 - a^2(t)dl^2 , \qquad (4.69)$$

where dl^2 is the Euclidean 3-space interval. In this case, calculations are much simpler in cartesian coordinates. The metric and its inverse are

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -a^2(t) & 0 & 0 \\ 0 & 0 & -a^2(t) & 0 \\ 0 & 0 & 0 & -a^2(t) \end{pmatrix} ;$$

$$(g^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -a^{-2}(t) & 0 & 0 \\ 0 & 0 & -a^{-2}(t) & 0 \\ 0 & 0 & 0 & -a^{-2}(t) \end{pmatrix} .$$

132

In the Christoffel symbols $\overset{\circ}{\Gamma}{}^\alpha{}_{\beta\nu}$, the only nonvanishing derivatives are those with respect to x^0. Consequently, only Christoffel symbols with at least one index equal to 0 will be nonvanishing. For example, $\overset{\circ}{\Gamma}{}^k{}_{ij} = 0$. Actually, the only Christoffels $\neq 0$ are:

$$\overset{\circ}{\Gamma}{}^k{}_{0j} = \delta^k_j \frac{1}{c} \frac{\dot{a}}{a} \; ; \; \overset{\circ}{\Gamma}{}^0{}_{ij} = \delta_{ij} \frac{1}{c} a \dot{a}.$$

The nonvanishing components of the Ricci tensor are

$$R_{00} = -\frac{3}{c^2} \frac{\ddot{a}}{a} = 3 \frac{H^2(t)}{c^2} q(t) \; ;$$

$$R_{ij} = \frac{\delta_{ij}}{c^2} [a\ddot{a} + 2\dot{a}^2] = \frac{\delta_{ij}}{c^2} a^2 H^2(t)[2 - q(t)] \; .$$

In consequence, the scalar curvature is

$$R = g^{00} R_{00} + g^{ij} R_{ij} = -6 \left[\frac{\ddot{a}}{c^2 a} + \left(\frac{\dot{a}}{ca}\right)^2 \right] = 6 \frac{H^2(t)}{c^2} [q(t) - 1] \; .$$

The nonvanishing components of the Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$ are

$$G_{00} = 3 \left(\frac{\dot{a}}{ca}\right)^2 \; ; \quad G_{ij} = -\frac{\delta_{ij}}{c^2} [\dot{a}^2 + 2a\ddot{a}] \; .$$

Let us consider the sourceless case with cosmological constant. The Einstein equations are then

$$G_{00} - \Lambda g_{00} = 3 \left(\frac{\dot{a}}{ca}\right)^2 - \Lambda = 0 \; ; \; G_{ij} - \Lambda g_{ij} = -\frac{\delta_{ij}}{c^2} [\dot{a}^2 + 2a\ddot{a} - \Lambda a^2] = 0 \; .$$

Subtracting 3 times one equation from the other, we arrive at the equivalent set

$$\dot{a}^2 - \frac{\Lambda c^2}{3} a^2 = 0 \; ; \; \ddot{a} - \frac{\Lambda c^2}{3} a = 0 \; . \tag{4.70}$$

These equations are (4.62) and (4.63) for the case under consideration. Of the two solutions, $a(t) = a_0 e^{\pm H_0(t-t_0)}$, only

$$a(t) = a_0 \, e^{H_0(t-t_0)} = a_0 \, e^{\sqrt{\frac{\Lambda c^2}{3}}(t-t_0)} \tag{4.71}$$

133

would be consistent with expansion. Expansion is a fact well established by observation. This is enough to fix the sign, and the model implies an everlasting exponential expansion. Notice that the scalar curvature is $R = -4\Lambda$, as is always the case in the absence of sources. Equation (4.71) is actually a de Sitter solution. The quick growth has been called "inflation" and is supposed to have taken place in the very early history of the Universe.

§ **4.25 Thermal History** The present–day content of the Universe consists of matter (visible or not) and radiation, the last constituting the cosmic microwave background. The energy density of the latter is very small, much smaller than that of visible matter alone. Nevertheless, it comes from the equations of state that radiation energy increases faster than matter energy with the temperature. Thus, though matter dominates the energy content of the Universe at present time, this dominance ceases at a "turning point" time in the past. At that point radiation takes over. At about the same time, hydrogen — the most common form of matter — ionizes. The photons of the background radiation establish contact with the electrons and the whole system is thermalized. Above that point, there exists a single temperature. And, above the turning point, the dominating photons increase progressively in number while their concentration grows by contraction. The opportunity for interactions between them becomes larger and larger. When they approach the mass of an electrons, pair creation sets up as a stable process. Radiation is now more than a gas of photons: it contains more and more electrons and positrons. Concomitantly, nucleosynthesis stops. As we insist in going up the temperature ladder, the photons, which are more and more energetic, break the composite nuclei. The nucleosynthesis period is the most remote time from which we have reasonably sure information nowadays.

The Standard model starts from present-day data and moves to the past, taking into account the changes in the equations of state. It goes consequently from a matter-dominated era through the time of hydrogen recombination, then to the changeover period in the which radiation establishes its dominance. These successive "eras", the so-called "thermal history" of the Universe, is analysed in the text *Physical Cosmology*. We shall here only examine the de Sitter solutions — also of fundamental cosmological interest

— because, by their simplicity, they give a beautiful example in which all calculation can be done without much ado.

4.3.2 de Sitter Solutions

§ **4.26** de Sitter spacetimes are hyperbolic spaces of constant curvature. They are solutions of vacuum Einstein's equation with a cosmological term. There are two different kinds of them: one with positive scalar Ricci curvature, and another one with negative scalar Ricci curvature. As the calculations are remarkably simple, we shall give a fairly detailed account.

We shall denote by $\mathcal{R}$ the de Sitter pseudo-radius, by $\eta_{\alpha\beta}$ $(\alpha, \beta, \cdots = 0, 1, 2, 3)$ the Lorentz metric of the Minkowski spacetime, and ξ^A $(A, B, \ldots = 0, \ldots, 4)$ will be the Cartesian coordinates of the pseudo-Euclidean 5–spaces. There are two types of spacetime named after de Sitter:

1. **de Sitter spacetime** $dS(4, 1)$: hyperbolic 4-surface whose inclusion in the pseudo–Euclidean space $\mathbb{E}^{4,1}$ satisfy

$$\eta_{AB}\,\xi^A\xi^B = \eta_{\alpha\beta}\,\xi^\alpha\xi^\beta - \left(\xi^4\right)^2 = -\mathcal{R}^2\ . \tag{4.72}$$

It is a one-sheeted hyperboloid (a 4-dimensional version of the surface seen in § 2.65) with topology $R^1 \times S^3$, and — within our conventions — negative scalar curvature. Its group of motions is the pseudo–orthogonal group $SO(4, 1)$

2. **anti–de Sitter spacetime** $dS(3, 2)$: hyperbolic 4-surface whose inclusion in the pseudo–Euclidean space $\mathbb{E}^{3,2}$ satisfy

$$\eta_{AB}\,\xi^A\xi^B = \eta_{\alpha\beta}\,\xi^\alpha\xi^\beta + \left(\xi^4\right)^2 = \mathcal{R}^2\ . \tag{4.73}$$

It is a two-sheeted hyperboloid (a 4-dimensional version of the space seen in § 2.64) with topology $S^1 \times R^3$, and positive scalar curvature. Its group of motions is $SO(3, 2)$

With the notation $\eta_{44} = s$, both de Sitter spacetimes can be put together in

$$\eta_{AB}\,\xi^A\xi^B = \eta_{\alpha\beta}\,\xi^\alpha\xi^\beta + s\left(\xi^4\right)^2 = s\mathcal{R}^2\ , \tag{4.74}$$

where we have the following relation between s and the de Sitter spaces:

$$s = -1 \text{ for } dS(4,1)$$
$$s = +1 \text{ for } dS(3,2).$$

§ 4.27 The metric Let us find now the line element of the de Sitter spaces. The most convenient coordinates are the stereographic conformal. The passage from the Euclidean ξ^A to the stereographic conformal coordinates x^α ($\alpha, \beta, \cdots = 0, 1, 2, 3$) is done by the transformation:

$$\xi^\alpha = \Omega x^\alpha \; ; \; \xi^4 = \mathcal{R}(1 - 2\Omega), \tag{4.75}$$

(a sign in the last expression would have no consequence for what follows) with $\Omega(x)$ a function of x^α which we shall determine. Two expressions preparatory to the calculation of the line element can be immediately obtained by taking differentials:

$$d\xi^\alpha = x^\alpha d\Omega + \Omega dx^\alpha , \tag{4.76}$$

and

$$\left(d\xi^4\right)^2 = 4\mathcal{R}^2 d\Omega^2. \tag{4.77}$$

Let us introduce $\rho^2 = \eta_{\alpha\beta}\, x^\alpha x^\beta$ and rewrite the defining relation (4.74) as

$$\Omega^2 \rho^2 + s \left(\xi^4\right)^2 = s\mathcal{R}^2 . \tag{4.78}$$

Equating $\left(\xi^4\right)^2$ got from (4.75) and (4.78), we find

$$\Omega = \frac{1}{1 + s \, \frac{\rho^2}{4\mathcal{R}^2}} \quad . \tag{4.79}$$

Notice that from this expression it follows that

$$d\Omega = -\, s \, \frac{\Omega^2}{2\mathcal{R}^2} \rho d\rho = -s \, \frac{\Omega^2}{4\mathcal{R}^2} 2\eta_{\alpha\beta}\, x^\alpha dx^\beta , \tag{4.80}$$

from which another preparatory result is obtained:

$$2\eta_{\alpha\beta}\, x^\alpha dx^\beta = -\, s \, \frac{4\mathcal{R}^2}{\Omega^2} d\Omega . \tag{4.81}$$

Now, the de Sitter line element is

$$d\Sigma^2 = \eta_{AB}\, d\xi^A d\xi^B = \eta_{\alpha\beta}\, d\xi^\alpha d\xi^\beta + s \, \left(d\xi^4\right)^2 ,$$

136

or, by using (4.76),

$$d\Sigma^2 = \eta_{\alpha\beta} \left(x^\alpha d\Omega + \Omega dx^\alpha \right) \left(x^\beta d\Omega + \Omega dx^\beta \right) + s \ \left(d\xi^4 \right)^2 \ .$$

Expanding and using (4.77),

$$d\Sigma^2 = \eta_{\alpha\beta} x^\alpha x^\beta d\Omega^2 + 2\eta_{\alpha\beta} x^\alpha dx^\beta \Omega d\Omega + \Omega^2 \eta_{\alpha\beta} dx^\alpha dx^\beta + s \ 4\mathcal{R}^2 d\Omega^2 \ .$$

Now, using (4.81),

$$d\Sigma^2 = \Omega^2 \eta_{\alpha\beta} dx^\alpha dx^\beta + \left[\rho^2 - s \ \frac{4\mathcal{R}^2}{\Omega} \right] d\Omega^2 + s \ 4\mathcal{R}^2 d\Omega^2 \ ,$$

and then (4.79),

$$d\Sigma^2 = \Omega^2 \eta_{\alpha\beta} dx^\alpha dx^\beta + \left[- s \ 4\mathcal{R}^2 \right] d\Omega^2 + s \ 4\mathcal{R}^2 d\Omega^2 \ ,$$

so that finally

$$d\Sigma^2 = g_{\alpha\beta} dx^\alpha dx^\beta \ , \tag{4.82}$$

where the metric $g_{\alpha\beta}$ is

$$g_{\alpha\beta} = \Omega^2 \ \eta_{\alpha\beta} \ = \ \frac{1}{\left[1 + s \ \frac{\rho^2}{4\mathcal{R}^2} \right]^2} \ \eta_{\alpha\beta} \ . \tag{4.83}$$

The de Sitter spaces are, therefore, conformally flat (see § 2.46), with the conformal factor given by $\Omega^2(x)$.

§ **4.28** The Christoffel symbol corresponding to a conformally flat metric $g_{\mu\nu}$ with conformal factor $\Omega^2(x)$ has the form

$$\overset{\circ}{\Gamma}{}^\alpha{}_{\beta\nu} = \left[\delta^\alpha_\beta \delta^\sigma_\nu + \delta^\alpha_\nu \delta^\sigma_\beta - \eta_{\beta\nu} \eta^{\alpha\sigma} \right] \partial_\sigma \ln \Omega(x) \ . \tag{4.84}$$

Taking derivatives in (4.79), we find for the de Sitter spaces

$$\overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma} = - s \ \frac{\Omega}{2\mathcal{R}^2} \left[\delta^\alpha_\beta \eta_{\gamma\sigma} + \delta^\alpha_\sigma \eta_{\gamma\beta} - \eta_{\beta\sigma} \delta^\alpha_\gamma \right] x^\gamma \ . \tag{4.85}$$

The Riemann tensor components can be found by taking the following steps. First, take the derivative of the de Sitter connection $\overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma}$:

$$\partial_\rho \ \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma} = - s \ \frac{\Omega}{2\mathcal{R}^2} \left[\delta^\alpha_\beta \eta_{\gamma\sigma} + \delta^\alpha_\sigma \eta_{\gamma\beta} - \eta_{\beta\sigma} \delta^\alpha_\gamma \right] \delta^\gamma_\rho + \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma} \ \partial_\rho \ln \Omega$$

137

$$= -s \frac{\Omega}{2\mathcal{R}^2} \left[\delta_\beta^\alpha \eta_{\rho\sigma} + \delta_\sigma^\alpha \eta_{\rho\beta} - \eta_{\beta\sigma}\delta_\rho^\alpha\right] - \frac{sx_\rho}{2\mathcal{R}^2\Omega} \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma}$$

$$= -s \frac{\Omega}{2\mathcal{R}^2} \left[\delta_\beta^\alpha \eta_{\rho\sigma} + \delta_\sigma^\alpha \eta_{\rho\beta} - \eta_{\beta\sigma}\delta_\rho^\alpha\right] + \frac{x_\rho x^\gamma}{4\mathcal{R}^4}\left[\delta_\beta^\alpha \eta_{\gamma\sigma} + \delta_\sigma^\alpha \eta_{\gamma\beta} - \eta_{\beta\sigma}\delta_\gamma^\alpha\right]$$

$$= -s \frac{\Omega}{2\mathcal{R}^2} \left[\delta_\beta^\alpha \eta_{\rho\sigma} + \delta_\sigma^\alpha \eta_{\rho\beta} - \eta_{\beta\sigma}\delta_\rho^\alpha\right] + \frac{x_\rho x^\gamma}{4\mathcal{R}^4\Omega^2}\left[\delta_\beta^\alpha g_{\gamma\sigma} + \delta_\sigma^\alpha g_{\gamma\beta} - g_{\beta\sigma}\delta_\gamma^\alpha\right]$$

$$= -s \frac{\Omega}{2\mathcal{R}^2} \left[\delta_\beta^\alpha \eta_{\rho\sigma} + \delta_\sigma^\alpha \eta_{\rho\beta} - \eta_{\beta\sigma}\delta_\rho^\alpha\right] + \frac{1}{4\mathcal{R}^4\Omega^2}\left[\delta_\beta^\alpha x_\sigma x_\rho + \delta_\sigma^\alpha x_\beta x_\rho - g_{\beta\sigma}x^\alpha x_\rho\right] \ .$$

Indicating by $[\rho\sigma]$ the antisymmetrization (without any factor) of the included indices, we get

$$\partial_\rho \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma} - \partial_\sigma \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\rho} = -s \frac{\Omega}{2\mathcal{R}^2} \left(\delta_{[\sigma}^\alpha \eta_{\rho]\beta} - \eta_{\beta[\sigma}\delta_{\rho]}^\alpha\right) + \frac{1}{4\mathcal{R}^4\Omega^2}\left(x_\beta \delta_{[\sigma}^\alpha x_{\rho]} - x^\alpha g_{\beta[\sigma}x_{\rho]}\right)$$

$$= -s \frac{\Omega}{\mathcal{R}^2}\delta_{[\sigma}^\alpha \eta_{\rho]\beta} + \frac{1}{4\mathcal{R}^4\Omega^2}\left(x_\beta \delta_{[\sigma}^\alpha x_{\rho]} - x^\alpha g_{\beta[\sigma}x_{\rho]}\right) \ .$$

This is the contribution of the derivative terms.

The product terms are

$$\overset{\circ}{\Gamma}{}^\alpha{}_{\lambda\rho}\overset{\circ}{\Gamma}{}^\lambda{}_{\beta\sigma} = \frac{\Omega^2}{4\mathcal{R}^4}\left[\delta_\lambda^\alpha \eta_{\gamma\rho} + \delta_\rho^\alpha \eta_{\gamma\lambda} - \eta_{\lambda\rho}\delta_\gamma^\alpha\right]\left[\delta_\beta^\lambda \eta_{s\sigma} + \delta_\sigma^\lambda \eta_{s\beta} - \eta_{\beta\sigma}\delta_s^\lambda\right]x^\gamma x^s;$$

$$\overset{\circ}{\Gamma}{}^\alpha{}_{\lambda\rho}\overset{\circ}{\Gamma}{}^\lambda{}_{\beta\sigma} - \overset{\circ}{\Gamma}{}^\alpha{}_{\lambda\sigma}\overset{\circ}{\Gamma}{}^\lambda{}_{\beta\rho} = \frac{1}{4\mathcal{R}^4\Omega^2}\left[x_\beta \delta_{[\rho}^\alpha x_{\sigma]} + x^\alpha x_{[\rho}g_{\sigma]\beta} + \Omega^2\rho^2 g_{\beta[\rho}\delta_{\sigma]}^\alpha\right] \ .$$

A provisional expression for the curvature is, therefore,

$$\overset{\circ}{R}{}^\alpha{}_{\beta\rho\sigma} = \partial_\rho \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\sigma} - \partial_\sigma \overset{\circ}{\Gamma}{}^\alpha{}_{\beta\rho} + \overset{\circ}{\Gamma}{}^\alpha{}_{\lambda\rho}\overset{\circ}{\Gamma}{}^\lambda{}_{\beta\sigma} - \overset{\circ}{\Gamma}{}^\alpha{}_{\lambda\sigma}\overset{\circ}{\Gamma}{}^\lambda{}_{\beta\rho}$$

$$= -s \frac{\Omega}{\mathcal{R}^2}\delta_{[\sigma}^\alpha \eta_{\rho]\beta} + \frac{1}{4\mathcal{R}^4\Omega^2}\left(x_\beta \delta_{[\sigma}^\alpha x_{\rho]} - x^\alpha g_{\beta[\sigma}x_{\rho]}\right)$$

$$+ \frac{1}{4\mathcal{R}^4\Omega^2}\left[x_\beta \delta_{[\rho}^\alpha x_{\sigma]} + x^\alpha x_{[\rho}g_{\sigma]\beta} + \Omega^2\rho^2 g_{\beta[\rho}\delta_{\sigma]}^\alpha\right] \ .$$

The first two terms in the last line just cancel the last two in the line above them. Therefore,

$$\overset{\circ}{R}{}^\alpha{}_{\beta\rho\sigma} = -s \frac{\Omega}{\mathcal{R}^2}\delta_{[\sigma}^\alpha \eta_{\rho]\beta} + \frac{1}{4\mathcal{R}^4\Omega^2}\Omega^2\rho^2 g_{\beta[\rho}\delta_{\sigma]}^\alpha$$

$$= -s \frac{\Omega}{\mathcal{R}^2}\eta_{\beta[\rho}\delta_{\sigma]}^\alpha + \frac{1}{4\mathcal{R}^4}\Omega^2\rho^2 \eta_{\beta[\rho}\delta_{\sigma]}^\alpha$$

$$= \left[-s \frac{\Omega}{\mathcal{R}^2} + \frac{1}{4\mathcal{R}^4}\Omega^2\rho^2\right]\eta_{\beta[\rho}\delta_{\sigma]}^\alpha$$

$$= s \frac{\Omega}{\mathcal{R}^2}\left[\frac{s}{4\mathcal{R}^2}\Omega\rho^2 - 1\right]\eta_{\beta[\rho}\delta_{\sigma]}^\alpha$$

138

Using (4.79), we find that the bracketed term is $= -\Omega$. Therefore, we get

$$\overset{\circ}{R}{}^{\alpha}{}_{\beta\rho\sigma} = -s \frac{\Omega^2}{\mathcal{R}^2}\eta_{\beta[\rho}\delta^{\alpha}_{\sigma]} = -\frac{s}{\mathcal{R}^2}\left[\delta^{\alpha}_{\sigma}g_{\beta\rho} - \delta^{\alpha}_{\rho}g_{\beta\sigma}\right] . \tag{4.86}$$

The Ricci tensor will be

$$\overset{\circ}{R}_{\mu\nu} = \frac{3\,s}{\mathcal{R}^2}\,g_{\mu\nu} \tag{4.87}$$

and the scalar curvature,

$$\overset{\circ}{R} = \frac{12\,s}{\mathcal{R}^2} . \tag{4.88}$$

The de Sitter spacetimes are spaces of constant curvature. We can now make contact with the cosmological term. From the expressions above, we find that

$$\overset{\circ}{R}_{\mu\nu} - \frac{1}{2}\,g_{\mu\nu}\,\overset{\circ}{R} + \frac{3\,s}{\mathcal{R}^2}\,g_{\mu\nu} = 0. \tag{4.89}$$

Thus, the de Sitter spaces are solutions of the sourceless Einstein's equations with a cosmological constant

$$\Lambda = -\frac{3\,s}{\mathcal{R}^2} = -\overset{\circ}{R}/4. \tag{4.90}$$

Notice the relationships to the de Sitter and the anti-de Sitter spaces:

$$s = -1 \text{ for the de Sitter space } dS(4,1) \longrightarrow \Lambda > 0$$
$$s = +1 \text{ for the anti-de Sitter space } dS(3,2) \longrightarrow \Lambda < 0 .$$

§ **4.29** We have said (§ 3.46) that positive scalar curvature tends to make curves to close to each other, and just the contrary for negative curvature. The relative sign in (4.90) shows that the cosmological constant has the opposite effect: $\Lambda > 0$ leads to diverging curves, $\Lambda < 0$ to converging ones. This actually depends on the initial conditions. Let us look at the geodetic deviation equation,

$$\frac{D^2X^{\alpha}}{Du^2} = \overset{\circ}{R}{}^{\alpha}{}_{\beta\rho\sigma}U^{\beta}U^{\rho}X^{\sigma} . \tag{4.91}$$

Using Eq.(4.86),

$$\frac{D^2 X^\alpha}{Ds^2} = - \tfrac{s}{\mathcal{R}^2} \left[\delta^\alpha_\sigma g_{\beta\rho} - \delta^\alpha_\rho g_{\beta\sigma} \right] U^\beta U^\rho X^\sigma$$

$$= - \tfrac{s}{\mathcal{R}^2} \left[\delta^\alpha_\sigma - U^\alpha U_\sigma \right] X^\sigma = - \tfrac{s}{\mathcal{R}^2} h^\alpha{}_\sigma X^\sigma \ . \quad (4.92)$$

We have recognized the transversal projector $h^\alpha{}_\sigma$ (see § 3.50). By the geodesic equation, the component of X along U will have vanishing contributions to the left-hand side. Consequently, only the transversal part $X_\perp$ will appear in the equation, which is now

$$\frac{D^2 X^\alpha_\perp}{Ds^2} + \tfrac{s}{\mathcal{R}^2} X^\alpha_\perp = \frac{D^2 X^\alpha_\perp}{Ds^2} + \tfrac{\mathring{R}}{12} X^\alpha_\perp = \frac{D^2 X^\alpha_\perp}{Ds^2} - \tfrac{\Lambda}{3} X^\alpha_\perp = 0 \ . \quad (4.93)$$

Negative Λ leads to oscillatory solutions. Positive Λ can lead both to contracting and expanding congruences. If two lines are initially separating, they will separate indefinitely more and more.

§ **4.30** We have been using carefully two coordinate systems. The most convenient system for cosmological considerations is the so–called comoving system, in which the Friedmann equations, in particular, have been written. In that system the scale parameter appears in its utmost simplicity. The stereographic coordinates are of interest for de Sitter spaces. We could perform a transformation between the two systems, but that is not really necessary: we have only taken scalar parameters from one system into the other. The only exception, Eq. (4.89), is a tensor which vanishes in a system and, consequently, vanishes also in the other.

The expression (4.83) for the metric is very different from the original one. De Sitter has found it in another coordinate system, in the form

$$ds^2 = \left(1 - \frac{\Lambda}{3} r^2 \right) c^2 dt^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) - \frac{dr^2}{1 - \frac{\Lambda}{3} r^2} \ . \quad (4.94)$$

This expression is just the Schwarzschild solution in the presence of a Λ term, Eq.(4.21), when the source mass tends to zero. There are many other coordinates and metric expressions of interest for different aims.** One of them exhibits clearly the inflationary property above discussed:

$$ds^2 = c^2 dt^2 - e^{ct/\mathcal{R}} \left(dx^2 + dy^2 + dz^2 \right) \ . \quad (4.95)$$

** A few, included that given below, are given by R.C. Tolman, *Relativity, Thermodynamics and Cosmology*, Dover, New York, 1987, § 142.

Chapter 5

Tetrad Fields

5.1 Tetrads

For each source, set of symmetries and boundary conditions, there will be a different solution of Einstein's equations. Each solution will be a spacetime. It will be interesting to go back and review the general characteristics of spacetimes. While in the process, we shall revisit some previously given notions and reintroduce them in a more formal language.

§ **5.1** A spacetime S is a four–dimensional differentiable manifold whose tangent space (§ 2.24) at each point is a Minkowski space. Loosely speaking, we implant a Lorentz metric η_{ab} on each tangent space. Bundle language is more specific: it considers the tangent bundle TS on spacetime, an 8-dimensional space which is locally the direct product of S and a typical fibre representing the tangent space. For a spacetime, the typical fiber is the Minkowski space M. The fiber is "typical" because it is an "ideal" (in the platonic sense) Minkowski space. The relationship between the typical Minkowski fiber and the spaces tangent to spacetime is established by tetrad fields (see Figure 5.1). A tetrad field will determine a copy of M on each tangent space. M is considered not only as a flat pseudo-Riemannian space, but as a vector space as well. We are going to use letters of the latin alphabet, $a, b, c, \ldots = 0, 1, 2, 3$, to label components on M, and those of the greek alphabeth, $\mu, \nu, \rho, \ldots = 0, 1, 2, 3$ for spacetime components. The first will be called "Minkowski indices", the latter "Riemann indices".

§ **5.2** We shall need an initial vector basis in M. We take the simplest one, the standard "canonical" basis

$$K_0 = (1, 0, 0, 0) \qquad K_1 = (0, 1, 0, 0)$$
$$K_2 = (0, 0, 1, 0) \qquad K_3 = (0, 0, 0, 1) \;.$$

Each K_a is given by the entries $(K_a)_b = \delta_{ab}$. Each tetrad field h will be a mapping

$$h : M \rightarrow TS, \quad h(K_a) = h_a \;.$$

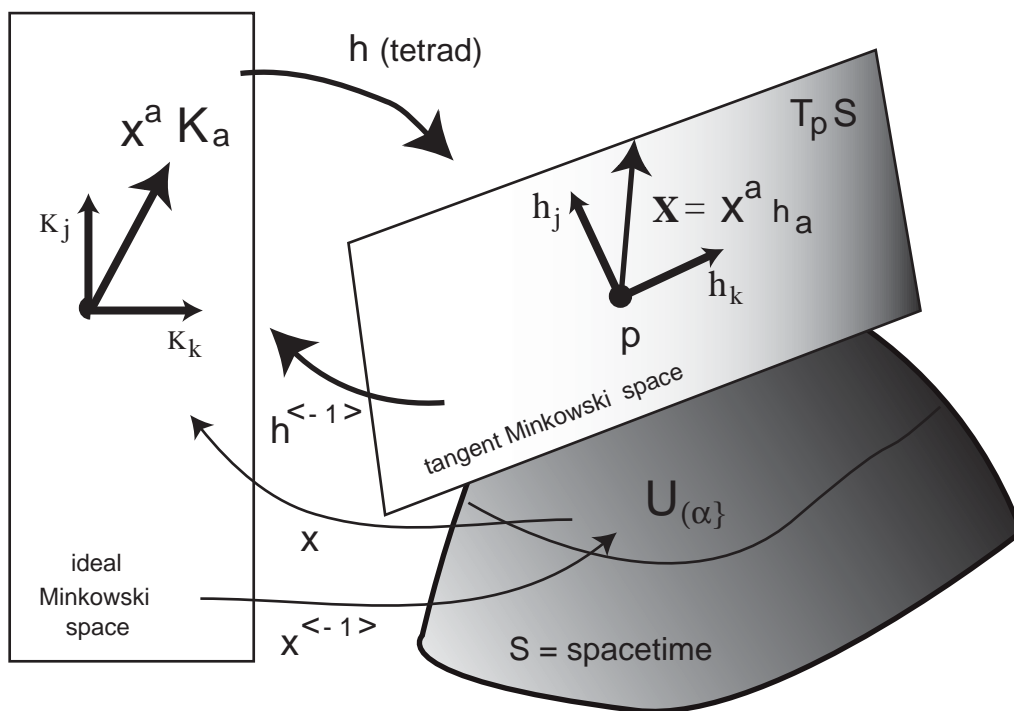The four vectors h_a will constitute a vector basis on S. Actually, this is a

Figure 5.1: *The role of a tetrad field.*

local affair: given a point $p \in S$, and around it an euclidean open set U, the h_a will constitute a vector basis not only for the tangent space to S at p, denoted T_pS, but also for all the T_qS with $q \in U$. The extension from p

to U is warranted by the differentiable structure. The dual forms h^b, such that $h^b(h_a) = \delta^b{}_a$, will constitute a vector basis on the cotangent space at p, denoted T_pS^*. The dual base $\{h^a\}$ can be equally extended to all the points of U. For example, each coordinate system $\{x^\mu\}$ on U will define a "natural" vector basis $\{\partial_\mu = \partial/\partial x^\mu\}$, with its concomitant covector dual basis, $\{dx^\mu\}$. This is a very particular and convenient tetrad field, frequently called a "trivial" tetrad, given by $e_a = e(K_a) = \delta_a{}^\mu \partial_\mu$. It is usual to fix a coordinate system around each point p from the start, and in this sense this basis is indeed "natural".

Another tetrad field, as the above generic $\{h_a\}$ and its dual $\{h^b\}$, can be written in terms of $\{\partial_\mu\}$ as

$$h_a = h_a{}^\mu \partial_\mu \quad \text{and} \quad h^b = h^b{}_\nu \, dx^\nu \, , \tag{5.1}$$

with

$$h^b{}_\mu \, h_a{}^\mu = \delta^b{}_a \quad \text{and} \quad h^a{}_\mu \, h_a{}^\nu = \delta_\mu{}^\nu \, . \tag{5.2}$$

The components $h_a{}^\mu(x)$ have one label in Minkowski space and one in spacetime, constituting a matrix with the inverse given by $h^a{}_\mu(x)$. It is usual to designate the tetrad by the sets $\{h_a{}^\mu(x)\}$ or $\{h^a{}_\mu\}$, but their meaning should be clear: they represent the components of a general tetrad field in a natural basis previously chosen.

§ **5.3** The tetrad Minkowski labels are vector indices, changing under Lorentz transformations according to

$$h^{a'} = \Lambda^{a'}{}_b \, h^b \, , \tag{5.3}$$

or, in terms of components,

$$h^{a'}{}_\mu(x) = \Lambda^{a'}{}_b(x) \, h^b{}_\mu(x) \, . \tag{5.4}$$

For each tangent Minkowski space, $\Lambda^{a'}{}_b$ is constant. It will, however, depend on the point of spacetime, which we indicate by its coordinates $x = \{x^\mu\}$. This is better seen if we contract the last expression above with

$h_c{}^\mu(x)$, to obtain the Lorentz transformation in terms of the initial and final tetrad basis:

$$\Lambda^{a'}{}_b(x) = h^{a'}{}_\mu(x)\, h_b{}^\mu(x) \ . \tag{5.5}$$

Equation (5.4) says that each tetrad component behaves, on each Minkowski fiber, as a Lorentz vector. For each fixed Riemann index μ, $h^a{}_\mu$ transforms according to the vector representation of the Lorentz group. A Lorentz (actually, co-)vector on a Minkowski space has components transforming, under a Lorentz transformation with parameters α^{cd}, as

$$\phi^{a'} = \Lambda^{a'}{}_b(x)\, \phi^b = \left[\exp\left(\tfrac{i}{2}\,\alpha^{cd}J_{cd}\right)\right]^{a'}{}_b \phi^b \ . \tag{5.6}$$

Here, each J_{cd} is a 4×4 matrix representing one of the Lorentz group generators:

$$[J_{cd}]^{a'}{}_b = i\left(\eta_{cb}\delta^{a'}{}_d - \eta_{db}\delta^{a'}{}_c\right) \ . \tag{5.7}$$

This means also that, for each fixed Riemann index μ, the $h^a{}_\mu$'s constitute a Lorentz basis (or frame) for M.

§ **5.4** A tetrad field converts tensors on M into tensors on spacetime, transliterating one index at a time. A general Lorentz tensor T, transforming according to

$$T^{a'b'c'\cdots} = \Lambda^{a'}{}_a\Lambda^{b'}{}_b\Lambda^{c'}{}_c\ldots T^{abc\cdots} \ ,$$

will satisfy automatically $T^{a'b'c'\cdots} = h^{a'}{}_\mu h^{b'}{}_\nu h^{c'}{}_\rho\ldots T^{\mu\nu\rho\cdots}$, which shows how tetrad fields can "mediate" Lorentz transformations. As an example of that transmutation, a tetrad will produce a field on spacetime out of a vector in Minkowski space by

$$\phi^\mu(x) = h_a{}^\mu(x)\, \phi^a(x) \ . \tag{5.8}$$

As the tetrad, in its Minkowski label, transforms under Lorentz transformation as a vector should do, $\phi^\mu(x)$ is Lorentz–invariant. Another case of "tensor transliteration" is

$$\Lambda^\mu{}_\nu(x) = h_{a'}{}^\mu(x)\, \Lambda^{a'}{}_b(x)\, h^b{}_\nu(x) = \delta^\mu{}_\nu \tag{5.9}$$

144

[using (5.5)]. Thus, there is no Lorentz transformation on spacetime itself.

The Minkowski indices — also called "tetrad indices" — are lowered by the Lorentz metric η_{ab}:

$$h_{a\sigma} = \eta_{ab} h^b{}_\sigma \ .$$

An important consequence is that the Lorentz metric η_{ab} is transmuted into the Riemannian metric

$$g_{\mu\nu} = \eta_{ab}\, h^a{}_\mu\, h^b{}_\nu \ . \tag{5.10}$$

Of course, also $g_{\mu\nu}$ (as any component of a spacetime tensor) is Lorentz–invariant. Here, given η_{ab}, the tetrad field defines the metric $g_{\mu\nu}$. Different tetrad fields transmute the same η_{ab} into different spacetime pseudo-Riemannian metrics.

§ 5.5 The members of a general tetrad field $\{h_a\}$, as vector fields, will satisfy a Lie algebra (see § 2.34) with a commutation table

$$[h_a, h_b] = c^c{}_{ab}\, h_c \ . \tag{5.11}$$

The structure coefficients $c^c{}_{ab}$ measure its anholonomicity — they are sometimes called "anoholonomicity coefficients" — and are given by

$$c^c{}_{ab} = [h_a(h_b{}^\mu) - h_b(h_a{}^\mu)]\ h^c{}_\mu \ . \tag{5.12}$$

If $\{h_a\}$ is holonomous, $c^c{}_{ab} = 0$, then $h^a = dx^a$ for some coordinate system $\{x^a\}$, and

$$dx^{a'} = \Lambda^{a'}{}_b\, dx^b \ .$$

Expression (5.10) would, in that holonomic case, give just the Lorentz metric written in another system of coordinates. This is the usual choice when we are interested only in Minkowski space transformations, because then $\Lambda^{a'}{}_b = \partial x^{a'}/\partial x^b$. In this case the tetrad components can be identified with the Lamé coefficients of coordinate transformations, and the metric $g_{\mu\nu}$ will be the Lorentz metric written in a general coordinate system.

We have up to now left quite indefinite the choice of the tetrad field itself. In fact its choice depends on the physics under consideration. Trivial tetrads are relevant when only coordinate transformations are considered. A non–trivial tetrad reveals the presence of a gravitational field, and is the fundamental tool in the description of such a field.

145

5.2 Linear Connections

Let us examine, in a purely descroptive way, the transformation properties of a linear connection. A linear connection is a 1-form with values in the linear algebra, that is, the Lie algebra of the linear group $GL(4, \mathbb{R})$ of all invertible real 4×4 matrices. This means a matrix of 1-forms. A Lorentz connection is a 1-form with values in the Lie algebra of the Lorentz group, which is a subgroup of $GL(4, \mathbb{R})$. All connections of interest to gravitation (the Levi-Civita in particular) are ultimately linear connections.

5.2.1 Linear Transformations

§ **5.6** A linear transformation of N variables x^r is an invertible transformation of the type

$$x^{r'} = M^{r'}{}_s \, x^s \, . \tag{5.13}$$

In this case $(M^{r'}{}_s)$ is an invertible matrix. Linear transformations form groups, which include all the groups of matrices. The set of invertible $N \times N$ matrices with real entries constitutes a group, called the real linear group $GL(N, \mathbb{R})$.

§ **5.7** Consider the set of $N \times N$ matrices. This set is, among other things, a vector space. The simplest of such matrices will be those $\Delta_\alpha{}^\beta$ whose entries are all zero except for that of the α-th row and β-th column, which is 1:

$$(\Delta_\alpha{}^\beta)^\delta{}_\gamma = \delta^\delta_\alpha \, \delta^\beta_\gamma \, . \tag{5.14}$$

An arbitrary $N \times N$ matrix K can be written $K = K^\alpha{}_\beta \, \Delta_\alpha{}^\beta$.

The $\Delta_\alpha{}^\beta$'s have one great quality: they are linearly independent (none can be written as linear combinations of the other). Thus, the set $\{\Delta_\alpha{}^\beta\}$ constitutes a basis (the "canonical basis") for the vector space of the $N \times N$ matrices. An advantage of this basis is that the components of a matrix K as a vector written in basis $\{\Delta_\alpha{}^\beta\}$ are the very matrix elements: $(K)^\alpha{}_\beta = K^\alpha{}_\beta$.

§ **5.8** Consider now the product of matrices: it takes each pair (A, B) of matrices into another matrix AB. In our notation, matrix product is performed

146

coupling lower–right indices to higher–left indices, as in

$$\left(\Delta_\alpha{}^\beta \Delta_\phi{}^\xi\right)^\delta{}_\epsilon = \left(\Delta_\alpha{}^\beta\right)^\delta{}_\gamma \left(\Delta_\phi{}^\xi\right)^\gamma{}_\epsilon = \delta^\beta{}_\phi \left(\Delta_\alpha{}^\xi\right)^\delta{}_\epsilon , \qquad (5.15)$$

where in $(\Delta_\alpha{}^\beta)^\delta{}_\gamma$, γ is the column index.

The structure of vector space (§ 2.20) includes an addition operation and its inverse, subtraction. Once the product is given, another operation can be defined by the commutator $[A, B] = AB - BA$, the subtraction of two products. The Lie algebra of the $N \times N$ real matrices with the operation defined by the commutator is called the real N-linear algebra, denoted $gl(N, \mathbb{R})$. The set $\{\Delta_\alpha{}^\beta\}$ is called the *canonical base* for $gl(N, \mathbb{R})$. A Lie algebra is summarized by its commutation table. For $gl(N, \mathbb{R})$, the commutation relations are

$$\left[\Delta_\alpha{}^\beta, \Delta_\phi{}^\zeta\right] = f_{\binom{\alpha}{\beta}\binom{\phi}{\zeta}}{}^{\binom{\delta}{\gamma}} \Delta_\gamma{}^\delta . \qquad (5.16)$$

The constants appearing in the right-hand side are the *structure coefficients*, whose values in the present case are

$$f_{\binom{\alpha}{\beta}\binom{\phi}{\zeta}}{}^{\binom{\delta}{\gamma}} = \delta_\phi{}^\beta \delta_\alpha{}^\gamma \delta_\delta{}^\zeta - \delta_\alpha{}^\zeta \delta_\phi{}^\gamma \delta_\delta{}^\beta . \qquad (5.17)$$

The choice of index positions may seem a bit awkward, but will be convenient for use in General Relativity. There, linear connections $\Gamma^\alpha{}_{\beta\mu}$ and Riemann curvatures $R^\alpha{}_{\beta\mu\nu}$ will play fundamental roles. These notations are quite well–established. Now, a linear connection is actually a matrix of 1-forms $\Gamma = \Delta_\alpha{}^\beta \Gamma^\alpha{}_\beta$, with the components $\Gamma^\alpha{}_\beta$ being usual 1-forms, just $\Gamma^\alpha{}_{\beta\mu} dx^\mu$. The first two indices refer to the linear algebra, the last to the covector character of Γ. A Riemann curvature is a matrix of 2-forms, $R = \Delta_\alpha{}^\beta R^\alpha{}_\beta$, where each $R^\alpha{}_\beta$ is an usual 2-form $R^\alpha{}_\beta = \frac{1}{2} R^\alpha{}_{\beta\mu\nu} dx^\mu \wedge dx^\nu$. In both cases, we talk of "algebra–valued forms". In consequence, the notation we use here for the matrices seem the best possible choice: in order to have a good notation for the components one must sacrifice somewhat the notation for the base.

The invertible $N \times N$ matrices constitute, as said above, the real linear group $GL(N, \mathbb{R})$. Each member of this group can be obtained as the exponential of some $K \in gl(N, \mathbb{R})$. $GL(N, \mathbb{R})$ is thus a Lie group, of which $gl(N, \mathbb{R})$ is the Lie algebra. The generators of the Lie algebra are also called, by extension, generators of the Lie group.

147

5.2.2 Orthogonal Transformations

§ 5.9 A group of continuous transformations preserving an invertible real symmetric bilinear form η (see § 2.44) on a vector space is an *orthogonal group* (or, if the form is not positive–definite, a pseudo–orthogonal group).

A symmetric bilinear form is a mapping taking two vectors into a real number: $\eta(\mathbf{u}, \mathbf{v}) = \eta_{\alpha\beta}\, u^\alpha v^\beta$, with $\eta_{\alpha\beta} = \eta_{\beta\alpha}$. It is represented by a symmetric matrix, which can always be diagonalized. Consequently, it is usually presented in its simplest, diagonal form in terms of some coordinates: $\eta(x, x) = \eta_{\alpha\beta} x^\alpha x^\beta$. Thus, the usual orthogonal group in $\mathbb{E}^3$ is the set $SO(3)$ of rotations preserving $\eta(\mathbf{x}, \mathbf{x}) = x^2 + y^2 + z^2$; the Lorentz group is the pseudo–orthogonal group preserving the Lorentz metric of Minkowski spacetime, $\eta(x, x) = c^2 t^2$ - x^2 - y^2 - z^2. These groups are usually indicated by $SO(\eta) = SO(r, s)$, with (r, s) fixed by the signs in the diagonalized form of η. The group of rotations in n-dimensional Euclidean space will be $SO(n)$, the Lorentz group will be $SO(3, 1)$, etc.

§ 5.10 Given the transformation $x^{\alpha'} = \Lambda^{\alpha'}{}_\alpha\, x^\alpha$, to say that "$\eta$ is preserved" is to say that the distance calculated in the primed frame and the distance calculated in the unprimed frame are the same. Take the squared distance in the primed frame, $\eta_{\alpha'\beta'}\, x^{\alpha'} x^{\beta'}$, and replace $x^{\alpha'}$ and $x^{\beta'}$ by their transformation expressions. We must have

$$\eta_{\alpha'\beta'}\, x^{\alpha'} x^{\beta'} = \eta_{\alpha'\beta'}\, \Lambda^{\alpha'}{}_\alpha \Lambda^{\beta'}{}_\beta\, x^\alpha x^\beta = \eta_{\alpha\beta}\, x^\alpha x^\beta \ , \quad \forall x \ . \tag{5.18}$$

This is the group–defining property, a condition on the $\Lambda^{\alpha'}{}_\alpha$'s. When η is an Euclidean metric, the matrices are orthogonal — that is, their columns are vectors orthogonal to each other. If η is the Lorentz metric, the $\Lambda^{\alpha'}{}_\alpha$'s belong to the Lorentz group. We see that it is necessary that

$$\eta_{\alpha\beta} = \eta_{\alpha'\beta'}\, \Lambda^{\alpha'}{}_\alpha \Lambda^{\beta'}{}_\beta = \Lambda^{\alpha'}{}_\alpha\, \eta_{\alpha'\beta'}\, \Lambda^{\beta'}{}_\beta.$$

The matrix form of this condition is, for each group element Λ,

$$\Lambda^T\, \eta\, \Lambda = \eta \ , \tag{5.19}$$

where Λ^T is the transpose of Λ.

§ 5.11 There is a corresponding condition on the members of the group Lie algebra. For each member A of the algebra, there will exist a group member Λ such that $\Lambda = e^A$. Taking $\Lambda = I + A + \frac{1}{2}A^2 + \dots$ and $\Lambda^T = I + A^T + \frac{1}{2}(A^T)^2 + \dots$ in the above condition and comparing order by order, we find that A must satisfy

$$A^T = -\ \eta^{-1}\ A\ \eta \tag{5.20}$$

and will consequently have vanishing trace: tr A = tr A^T = - tr $(\eta^{-1}\ A\ \eta)$ = - tr $(\eta\eta^{-1}\ A)$ = - tr A $\therefore$ tr $A = 0$.

§ 5.12 If η is defined on an N-dimensional space, the Lie algebras $so(\eta)$ of the orthogonal or pseudo–orthogonal groups will be subalgebras of $gl(N, \mathbb{R})$. Given an algebra $so(\eta)$, both basis and entry indices can be lowered and raised with the help of η. We define new matrices $\Delta_{\alpha\beta}$ by lowering labels with $\eta : (\Delta_{\alpha\beta})^\delta{}_\gamma = \delta^\delta{}_\alpha \eta_{\beta\gamma}$. Their commutation relations become

$$[\Delta_{\alpha\beta}, \Delta_{\gamma\delta}] = \eta_{\beta\gamma}\,\Delta_{\alpha\delta} - \eta_{\alpha\delta}\,\Delta_{\gamma\beta}\ . \tag{5.21}$$

The generators of $so(\eta)$ will then be $J_{\alpha\beta} = \Delta_{\alpha\beta}$ - $\Delta_{\beta\alpha}$, with commutation relations

$$[J_{\alpha\beta}, J_{\gamma\delta}] = \eta_{\alpha\delta}J_{\beta\gamma} + \eta_{\beta\gamma}J_{\alpha\delta} - \eta_{\beta\delta}J_{\alpha\gamma} - \eta_{\alpha\gamma}J_{\beta\delta}\ . \tag{5.22}$$

These are the general commutation relations for the generators of the orthogonal or pseudo–orthogonal group related to η. We shall meet many cases in what follows. Given η, the algebra is fixed up to conventions. The usual group of rotations in the 3-dimensional Euclidean space is the special orthogonal group, denoted by $SO(3)$. Being "special" means connected to the identity, that is, represented by 3×3 matrices of determinant = +1.

The group $O(N)$ is formed by the orthogonal $N \times N$ real matrices. $SO(N)$ is formed by all the matrices of $O(N)$ which have determinant = +1. In particular, the group $O(3)$ is formed by the orthogonal 3×3 real matrices. $SO(3)$ is formed by all the matrices of $O(3)$ which have determinant = +1. The Lorentz group, as already said, is $SO(3, 1)$. Its generators have just the algebra (5.22), with η the Lorentz metric.

5.2.3 Connections, Revisited

§ 5.13 Suppose we are given the connection by components $\Gamma^a{}_{b\nu}$, the first two indices being "algebraic" and the last a Riemann index. This supposes a basis in the linear algebra and a basis of vector fields on the manifold. Taking the canonical basis $\{\Delta_a{}^b\}$ for the algebra, and a holonomic vector basis $\{dx^\nu\}$ on the manifold, for example, the connection is given in invariant form by

$$\Gamma = \tfrac{1}{2}\, \Delta_a{}^b\, \Gamma^a{}_{b\nu}\, dx^\nu \; . \tag{5.23}$$

For reasons which will become clear later, the set of components $\{\Gamma^a{}_{b\nu}\}$ will be called *spin connection.*

Connections have been introduced in in Section 2.3 through their behavior under coordinate tranformations, that is, under change of holonomic tetrads. We proceed now to a series of steps extending that presentation to general, holonomic or not, tetrads. First, we (i) change from Minkowski indices to Riemann indices by

$$\Gamma^a{}_{b\nu} \rightarrow \Gamma^\lambda{}_{\mu\nu} = h^\lambda{}_a\, \Gamma^a{}_{b\nu}\, h^b{}_\mu + h^\lambda{}_a \partial_\nu h^a{}_\mu \; . \tag{5.24}$$

This generalizes Eq.(2.33). Then, we (ii) change again through a Lorentz–transformed tetrad, $\Gamma^\lambda{}_{\mu\nu} \rightarrow \Gamma^{a'}{}_{b'\nu} = h^{a'}{}_\lambda\, \Gamma^\lambda{}_{\mu\nu}\, h^\mu{}_{b'} + h^{a'}{}_\lambda\, \partial_\nu h^\rho{}_{b'}$, which means that

$$\Gamma^{a'}{}_{b'\nu} = \Lambda^{a'}{}_a\, \Gamma^a{}_{b\nu}\, \left(\Lambda^{-1}\right)^b{}_{b'} + \Lambda^{a'}{}_c \partial_\nu \left(\Lambda^{-1}\right)^c{}_{b'} \; . \tag{5.25}$$

This gives the effect on $\Gamma^a{}_{b\nu}$ of a Lorentz transformation $\Lambda^{a'}{}_a = h^{a'}{}_\lambda\, h^\lambda{}_a$. In the notation adopted, $\Lambda^{a'}{}_a$ changes V^a into $V^{a'}$; we can write simply $\Lambda^c{}_{b'}$ for the inverse $(\Lambda^{-1})^c{}_{b'}$, understanding that

$$V^c = \Lambda^c{}_{b'}\, V^{b'} = \left(\Lambda^{-1}\right)^c{}_{b'}\, V^{b'} \; .$$

Now, start instead with $\Gamma^\lambda{}_{\mu\nu}$, and (iii) change from Riemann indices to Minkowski indices by

$$\Gamma^\lambda{}_{\mu\nu} \rightarrow \Gamma^a{}_{b\nu} = h^a{}_\lambda\, \Gamma^\lambda{}_{\mu\nu}\, h^\mu{}_b + h^a{}_\lambda\, \partial_\nu h^\lambda{}_b \; , \tag{5.26}$$

and (iv) go back to modified Riemann indices by

$$\Gamma^{\lambda'}{}_{\mu'\nu'} = h^{\lambda'}{}_a\, \Gamma^a{}_{b\nu}\, h^b{}_{\mu'} + h^{\lambda'}{}_c\, \partial_\nu h^c{}_{\mu'} \; .$$

Consequently,

$$\Gamma^{\lambda'}{}_{\mu'\nu'} = h^{\lambda'}{}_a\, h^a{}_\lambda\, \Gamma^\lambda{}_{\mu\nu}\, h^\mu{}_b + h^{\lambda'}{}_a\, h^a{}_\lambda \left(\partial_\nu h^\lambda{}_b\right) h^b{}_{\mu'} + h^{\lambda'}{}_c\, \partial_\nu h^c{}_{\mu'}\ ,$$

or

$$\Gamma^{\lambda'}{}_{\mu'\nu'} = B^{\lambda'}{}_\lambda\, \Gamma^\lambda{}_{\mu\nu}\, B^{\mu'}{}_{\mu'} + B^{\lambda'}{}_\lambda\, \partial_\nu B^\lambda{}_{\mu'}\ . \qquad (5.27)$$

This is the effect of a change of basis given by $B^{\lambda'}{}_\lambda = h^{\lambda'}{}_a\, h^a{}_\lambda$.

§ **5.14** Vector fields transform according to (5.6):

$$\phi^{e'}(x) = h^{e'}{}_\mu(x)\phi^\mu(x) = \Lambda^{e'}{}_b\, h^b{}_\mu(x)\, \phi^\mu(x) = \Lambda^{e'}{}_b\, \phi^b(x)\ . \qquad (5.28)$$

What happens to their derivatives? Clearly, they transform in another way:

$$\partial_\lambda\phi^{e'} = \partial_\lambda \left[\Lambda^{e'}{}_b\right]\phi^b + \Lambda^{e'}{}_b\, \partial_\lambda\phi^b\ .$$

As the name indicates, the covariant derivative of a given object is a derivative modified in such a way as to keep, under transformations, just the same behavior of the object. Here, it will have to obey

$$D'_\lambda\phi^{e'} = \Lambda^{e'}{}_b\, D_\lambda\phi^b\ . \qquad (5.29)$$

§ **5.15** The way physicists introduce a connection is as a "compensating field", an object $\Gamma^a{}_{b\nu}$ with a very special behavior whose action on the field, once added to the usual derivative, produces a covariant result. In the present case we look for a connection such that

$$\partial_\lambda\phi^{e'} + \Gamma^{e'}{}_{d'\lambda}\,\phi^{d'} = \Lambda^{e'}{}_b\left[\partial_\lambda\phi^b + \Gamma^b{}_{d\lambda}\,\phi^d\right]\ .$$

A direct calculation shows then that the required behavior is just (5.25), which can be written also as

$$\Gamma^{a'}{}_{b'\nu} = \Lambda^{a'}{}_d\left[\delta^d{}_c\partial_\mu + \Gamma^d{}_{c\nu}\right]\Lambda^c{}_{b'}\ . \qquad (5.30)$$

§ **5.16** As a rule, all indexed objects are tensor components and transform accordingly. A connection, written as $\Gamma^a{}_{b\nu}$, is an exception: it is tensorial in the last index, but not in the first two, which change in the peculiar

151

way shown above. Any connection transforming in this way will lead to a covariant derivative of the form

$$\nabla_\mu \phi^a = \partial_\mu \phi^a + \Gamma^a{}_{b\nu}\, \phi^b = h^e{}_\mu \left[e_e \phi^a + \Gamma^a{}_{be}\, \phi^b \right] \; . \tag{5.31}$$

Applied in particular to $h^a{}_\sigma$, it gives

$$\nabla_\mu h^a{}_\sigma = \partial_\mu h^a{}_\sigma + \Gamma^a{}_{b\mu}\, h^b{}_\sigma \; ;$$

applied to its inverse $h_a{}^\sigma$,

$$\nabla_\mu h_a{}^\sigma = \partial_\mu h_a{}^\sigma - \Gamma^b{}_{a\mu}\, h_b{}^\sigma \; .$$

It is easily checked that $\Gamma^a{}_{b\mu} = \frac{i}{2}\, \Gamma^{cd}{}_\mu\, (J_{cd})^a{}_b$, with J_{cd} given in (5.7). It has its values in the Lie algebra of the Lorentz group. The spin connetion $\Gamma^a{}_{b\mu}$ is, for this reason, said to be a *Lorentz connection*. We can define the matrix $\Gamma_\mu = \frac{i}{2}\, \Gamma^a{}_{b\mu}\, J_a{}^b$ whose entries are $\Gamma^a{}_{b\mu}$. Then,

$$\nabla_\mu \phi^a = \partial_\mu \phi^a + \frac{i}{2}\, \Gamma^{cd}{}_\mu\, (J_{cd})^a{}_b\, \phi^b = [\partial_\mu \phi + \Gamma_\mu\, \phi]^a \; .$$

We find also that

$$\Gamma^\lambda{}_{\nu\mu} = h_a{}^\lambda\, \Gamma^a{}_{b\mu}\, h^b{}_\nu + h_b{}^\lambda\, \partial_\mu h^b{}_\nu \tag{5.32}$$

is Lorentz invariant, that is,

$$h_{a'}{}^\lambda \left[\delta^{a'}{}_{b'}\, \partial_\mu + \Gamma^{a'}{}_{b'\mu} \right] h^{b'}{}_\nu = h_a{}^\lambda \left[\delta^a{}_b\, \partial_\mu + \Gamma^a{}_{b\mu} \right] h^b{}_\nu \; .$$

Notice that the components $\Gamma^a{}_{b\mu}$ and $\Gamma^\lambda{}_{\nu\mu}$ refer to different spaces. Equation (5.32) describes how Γ changes when the algebra indices are changed into Riemann indices. It should not be confused with (5.25), which relates the connection components in two Lorentz–related frames.

Comment 5.1 Equation (5.32) is frequently written as the vanishing of a "total covariant derivative" of the tetrad:

$$\partial_\lambda h^a{}_\mu - \Gamma^\sigma{}_{\mu\lambda}\, h^a{}_\mu + \Gamma^a{}_{c\mu}\, h^c{}_\mu = 0 \; .$$

§ **5.17** As the two first indices in $\Gamma^a{}_{b\mu}$ are not "tensorial", the behavior of Γ is very special. On the other hand, the indices in J_{ab} are tensorial. The consequence is that the contraction $\Gamma_\mu = \frac{i}{2}\, \Gamma^{ab}{}_\mu J_{ab}$ is not Lorentz invariant. Actuallly,

$$\Gamma^{a'b'}{}_\mu \, J_{a'b'} = \Lambda^{a'}{}_d \, \partial_\mu \Lambda^{db'} \, J_{a'b'} + \Gamma^{ab}{}_\mu \, J_{ab} \ .$$

Again, decomposing Λ in terms of the tetrad, we find

$$\left[\Gamma^{a'b'}{}_\mu + h^{b'\lambda} \, \partial_\mu h^{a'}{}_\lambda \right] J_{a'b'} = \left[\Gamma^{ab}{}_\mu + h^{b\lambda} \, \partial_\mu h^a{}_\lambda \right] J_{ab} \ .$$

Thus, what is really invariant is

$$\Gamma_\mu = \tfrac{1}{2} \left[\Gamma^{ab}{}_\mu + h^{b\lambda} \, \partial_\mu h^{a\lambda} \right] J_{ab} \ . \tag{5.33}$$

§ **5.18** The "archaic" approach to connections is more intuitive and very suggestive. It is worth recalling, as it complements the above one. It starts with the assumption that, under an infinitesimal displacement dx^λ, a field suffers a change which is proportional to its own value and to dx^λ. The proportionality coefficient is an "affine coefficient" $\Gamma^\mu{}_{\nu\lambda}$ (an oldish name !), so that

$$\delta\phi^\mu(x) = -\, \Gamma^\mu{}_{\nu\lambda}(x)\, \phi^\nu(x)\, dx^\lambda \ .$$

If we introduce the entries of the matrix J_{ab}, we verify that this is the same as

$$\delta\phi^\mu(x) = -\, \tfrac{1}{2}\, \Gamma^{ab}{}_\lambda(x)\, (J_{ab})^\mu{}_\nu\, \phi^\nu(x)\, dx^\lambda \ .$$

Consequently, the variation in the functional form of ϕ will be

$$\bar\delta\phi^\mu(x) = \delta\phi^\mu(x) - \partial_\lambda\phi^\mu(x)\, dx^\lambda = -\, \nabla_\lambda\phi^\mu(x)\, dx^\lambda \ ,$$

which defines the covariant derivative

$$\nabla_\lambda\phi^\mu(x) = \partial_\lambda\phi^\mu(x) + \Gamma^\mu{}_{\nu\lambda}\, \phi^\nu(x) \ .$$

§ **5.19** When $\nabla_\lambda\phi^\mu = 0$, we say that the field is "parallel transported". In that case, we have $\bar\delta\phi^\mu(x) = 0$, or $\phi^{\mu\prime}(x) = \phi^\mu(x)$. In parallel transport,

the functional form of the field does not change. To learn more on the meaning of parallel-transport and its expression as $\nabla_\lambda \phi^\mu = 0$, let us look at the functional variation of the vector field along a curve γ of tangent vector (velocity) U. It will be the 1-form $\bar{\delta}\phi^\mu(x)$ applied to the field $U = d/ds$:

$$\bar{\delta}\phi^\mu(x)[U] = [\partial_\lambda \phi^\mu + \Gamma^\mu{}_{\nu\lambda}\, \phi^\nu]\, dx^\lambda[U] = U^\lambda\, \nabla_\lambda \phi^\mu = \frac{D\phi^\mu}{ds}\ .$$

The purely-functional variation along the curve will vanish if $\nabla_\lambda \phi^\mu = 0$. This is what is meant when we say that ϕ^μ is parallel-transported along a curve γ: the field is transported in such a way that it suffers no change in its functional form. The change coming from the argument,

$$\frac{d\phi^\mu}{ds} = \frac{dx^\lambda}{ds}\, \partial_\lambda \phi^\mu = U^\lambda\, \partial_\lambda \phi^\mu\ ,$$

is exactly compensated by the term $U^\lambda \Gamma^\mu{}_{\nu\lambda}\, \phi^\nu$.

When $\phi^\mu = U^\mu$ itself, we have the acceleration DU^μ/ds. The condition of no variation of the velocity-field along the curve will lead to the geodesic equation

$$\frac{DU^\mu}{ds} = U^\lambda\, [\partial_\lambda U^\mu + \Gamma^\mu{}_{\nu\lambda}\, U^\nu] = \frac{dU^\mu}{ds} + \Gamma^\mu{}_{\nu\lambda}\, U^\nu\, U^\lambda = 0\ ,$$

which is an equation for γ.

5.2.4 Back to Equivalence

§ 5.20 We have seen in Section 3.7 how a convenient choice of coordinates leads to the vanishing of the Levi-Civita connection at a point. Nevertheless, we had in § 3.11 introduced an observer as a timelike curve, and qualified that notion in the ensuing paragraphs. A timelike curve is, actually, an *ideal* observer, which is point-like in any local space section. Real observers are extended in space and can always detect gravitation by comparing the neighboring curves followed by hers/his parts (§ 3.46). Ideal observers can be arbitrary curves (subsection 3.8.2) and, eventually, can have well-defined space-sections all along (subsection 3.8.4).

We shall now see how a tetrad $\{H_a\}$ can be chosen so that, seen from the frame it represents, the connection can be made to vanish all along a curve.

Looking from that frame, an ideal observer will not feel the gravitational field.

§ 5.21 Take a differentiable curve γ which is an integral curve of a field U, with $U^\mu = \frac{dx^\mu}{ds} = \frac{d\gamma^\mu(s)}{ds}$. The condition for the connection to vanish along γ, $\Gamma^a{}_{b\nu}(\gamma(s)) = 0$, will be

$$U^\nu \partial_\nu\ H_a{}^\lambda(\gamma(s)) + \Gamma^\lambda{}_{\mu\nu}(\gamma(s))\ U^\nu\ H_a{}^\mu(\gamma(s)) = 0,$$

that is,

$$\frac{d}{ds}\ H_a{}^\lambda(\gamma(s)) + \Gamma^\lambda{}_{\mu\nu}(\gamma(s))\ U^\nu\ H_a{}^\mu(\gamma(s)) = 0. \qquad (5.34)$$

This is simply the requirement that the tetrad (each member H_a of it) be parallel-transported along γ. Given a curve and a linear connecion, any vector field can be parallel-transported along γ. The procedure is then very simple: take, in the way previously discussed, a point P on the curve and find the corresponding trivial tetrad $\{H_a(P)\}$; then, parallel-transport it along the curve.

For the dual base $\{H^a\}$, the above formula reads

$$\frac{d}{ds}\ H^a{}_\mu(\gamma(s)) - \Gamma^\rho{}_{\mu\nu}(\gamma(s))\ U^\nu\ H^a{}_\rho(\gamma(s)) = 0. \qquad (5.35)$$

§ 5.22 Take (5.34) in the form

$$H^b{}_\mu(x)\frac{d}{ds}\ H_b{}^\lambda(x) = -\ \Gamma^\lambda{}_{\mu\nu}(x)U^\nu$$

and contract it with U^μ:

$$U^\mu H^b{}_\mu(x)\frac{d}{ds}\ H_b{}^\lambda(x) = -\ \Gamma^\lambda{}_{\mu\nu}(x)U^\mu U^\nu. \qquad (5.36)$$

Seen from the tetrad, the tangent field will be $U^b = H^b{}_\mu U^\mu$, and the above formula is

$$U^b\frac{d}{ds}\ H_b{}^\lambda(x) + \Gamma^\lambda{}_{\mu\nu}(x)U^\mu U^\nu = 0, \qquad (5.37)$$

or

$$H_b{}^\lambda(x)\ \frac{d}{ds}\ U^b = \frac{d}{ds}\ U^\lambda(x) + \Gamma^\lambda{}_{\mu\nu}(x)U^\mu U^\nu = \frac{D}{Ds}\ U^\lambda(x). \qquad (5.38)$$

155

This shows that, if γ is a self-parallel curve, then

$$\frac{d}{ds}\, U^a = 0. \tag{5.39}$$

This is the equation for a geodesic, as seen from the frame $\{H_a = H_a{}^\lambda \frac{\partial}{\partial x^\lambda}\}$. If an external force is present, then $m\frac{D}{Ds}\, U^\lambda(x) = F^\lambda$ and

$$m\frac{d}{ds}\, U^a = F^a. \tag{5.40}$$

This means that, looking from that tetrad, the observer will see the laws of Physics as given by Special Relativity.

§ **5.23** Consider now a Levi-Civita geodesic. In that case there exists a preferred tetrad h_a, which is not parallel-transported along the curve. Its deviation from parallelism is measured by the spin connection. Indeed, from (5.32),

$$\frac{d}{ds}\, h_b{}^\lambda + \Gamma^\lambda{}_{\mu\nu}\, U^\nu\, h_b{}^\mu = h_a{}^\lambda\, \Gamma^a{}_{b\nu}\, U^\nu . \tag{5.41}$$

This is the same as

$$\frac{d}{ds}\, h^b{}_\lambda = h^b{}_\sigma\, \Gamma^\sigma{}_{\lambda\nu}\, U^\nu - h^d{}_\lambda \Gamma^b{}_{dc}\, \mathring{U}^c . \tag{5.42}$$

We call $\mathring{U}$ the velocity as seen from the frame $\{h_a\}$. Equation (5.32) is actually a representation of the Equivalence Principle. Let us write it into still another form,

$$\frac{D}{Ds}\, h_b = \frac{d}{ds}\, h_b + \Gamma\left(\frac{d}{ds}\right)\, h_b = \Gamma^a{}_b\left(\frac{d}{ds}\right)\, h_a . \tag{5.43}$$

This holds for *any* curve with tangent vector $(\frac{d}{ds})$. It means that the frame $\{h_a\}$ can be parallel-transported along *no* curve. The spin connection forbids it, and gives the rate of change with respect to parallel transport.

§ **5.24** One of the versions of the Principle — the etymological one — says that a gravitational field is equivalent to an accellerated frame. Which frame ? We see now the answer: the frame equivalent to the field represented by the metric $g_{\mu\nu} = \eta_{ab}h^a{}_\mu h^b{}_\nu$ is just the anholonomous frame $\{h_a\}$. Another

156

piece of the Principle says that it is possible to choose a frame in which the connection vanishes. Let us see now how to change from the "equivalent" frame $\{h_a\}$ to the free-falling frame $\{H_a\}$. Contracting (5.41) with $H^a{}_\lambda$,

$$ H^a{}_\lambda \frac{d}{ds}\, h_b{}^\lambda + \Gamma^\lambda{}_{\mu\nu}\, U^\nu\, H^a{}_\lambda h_b{}^\mu = H^a{}_\lambda h_c{}^\lambda\, \Gamma^c{}_{b\nu}\, U^\nu\ , $$

which is the same as

$$ \frac{d}{ds}\, \left(H^a{}_\lambda h_b{}^\lambda \right) - h_b{}^\mu \left(\frac{d}{ds}\, H^a{}_\mu - \Gamma^\lambda{}_{\mu\nu}\, U^\nu\, H^a{}_\lambda \right) = H^a{}_\lambda h_c{}^\lambda\, \Gamma^c{}_{b\nu}\, U^\nu\ . $$

The second term in the left-hand side vanishes by Eq.(5.35), so that we remain with

$$ \frac{d}{ds}\, \left(H^a{}_\lambda h_b{}^\lambda \right) - H^a{}_\lambda h_c{}^\lambda\, \Gamma^c{}_{b\nu}\, U^\nu\ = 0\ . \tag{5.44} $$

What appears here is a point-dependent Lorentz transformation relating the metric tetrad h_a to the frame H_a in which Γ vanishes:

$$ H^a{}_\lambda = \Lambda^a{}_b\, h^b{}_\lambda\ , \quad H^a = \Lambda^a{}_b\, h^b\ . \tag{5.45} $$

This means

$$ \Lambda^a{}_b = H^a{}_\lambda\, h_b{}^\lambda\ . \tag{5.46} $$

This relation holds on the common domain of definition of both tetrad fields. Taking this into (5.44), we arrive at a relationship which holds on the intersection of that domain with a geodesics of the Levi-Civita connection $\overset{\circ}{\Gamma}$:

$$ \frac{d}{ds}\, \Lambda^a{}_b - \Lambda^a{}_c\, \overset{\circ}{\Gamma}{}^c{}_{bd}\, \overset{\circ}{U}{}^d = 0\ . \tag{5.47} $$

This equation gives the change, along the metric geodesic, of the Lorentz transformation taking the metric tetrad h_a into the frame H_a. The vector formed by each row of the Lorentz matrix is parallel-transported along the line. Contracting with the inverse Lorentz transformation

$$ (\Lambda^{-1})^a{}_b = h^a{}_\rho H_b{}^\rho\ , \tag{5.48} $$

the expression above gives

$$ \overset{\circ}{\Gamma}{}^a{}_{bd}\, \overset{\circ}{U}{}^d = (\Lambda^{-1})^a{}_c\, \frac{d}{ds}\, \Lambda^c{}_b\ = (\Lambda^{-1} \frac{d}{ds}\, \Lambda)^a{}_b\ . \tag{5.49} $$

157

This is, in the language of differential forms,

$$\overset{\circ}{\Gamma}{}^a{}_{bd}\, h^d\left(\frac{d}{ds}\right) = (\Lambda^{-1}d\Lambda)^a{}_b\left(\frac{d}{ds}\right).\tag{5.50}$$

Thus *on the points of the curve*, the connection has the form of a gauge Lorentz vacuum:

$$\overset{\circ}{\Gamma}{}^a{}_b = (\Lambda^{-1}d\Lambda)^a{}_b\tag{5.51}$$

It is important to stress that this is only true along a curve — a one-dimensional domain — so that curvature is not affected. Curvature, the real gravitational field, only manisfests itself on two-dimensional domains. Seen from the frame h_a, the geodesic equation has the form

$$\frac{d}{ds}\,\overset{\circ}{U}{}^a + \overset{\circ}{\Gamma}{}^a{}_{bc}\,\overset{\circ}{U}{}^b\overset{\circ}{U}{}^c = 0,\tag{5.52}$$

which is the same as

$$\frac{d}{ds}\,\overset{\circ}{U}{}^a + (\Lambda^{-1}\frac{d}{ds}\,\Lambda)^a{}_b\,\overset{\circ}{U}{}^b = 0.$$

This expression, once multiplied on the left by Λ, gives

$$\Lambda^c{}_a\,\frac{d}{ds}\,\overset{\circ}{U}{}^a + \frac{d}{ds}\,(\Lambda^c{}_a)\,\overset{\circ}{U}{}^a = \frac{d}{ds}\,(\Lambda^c{}_a\,\overset{\circ}{U}{}^a) = 0.\tag{5.53}$$

This is Eq. (5.39) for the present case. Summing up: at each point of the curve/observer there is a Lorentz transformation taking the accelerated frame $\{h_a\}$, equivalent to the gravitational field, into the inertial frame $\{H_a\}$, in which the force equation acquires the form it would have in Special Relativity.

These considerations can be enlarged to general Lorentz tensors. Take, for instance, a second order tensor:

$$T^{ab} = \Lambda^a{}_c\,\Lambda^b{}_d\,\overset{\circ}{T}{}^{cd}.$$

Taking $\frac{d}{ds}$ of this expression leads to

$$\frac{d}{ds}\,\overset{\circ}{T}{}^{ab} + \overset{\circ}{\Gamma}{}^a{}_c\,\overset{\circ}{T}{}^{cb} + \overset{\circ}{\Gamma}{}^b{}_c\,\overset{\circ}{T}{}^{ac} = (\Lambda^{-1})^a{}_c\,(\Lambda^{-1})^b{}_d\,\frac{d}{ds}\,T^{cd}.$$

The covariant derivative according to the connection $\overset{\circ}{\Gamma}$, which is Γ seen from the frame h_a, is the Lorentz transform of the simple derivative as seen from the inertial frame H_a, in which Γ vanishes.

158

Summing up again: at each point of the curve/observer there is a Lorentz transformation taking the accelerated frame $\{h_a\}$, equivalent to the gravitational field, into the free falling frame $\{H_a\}$, in which all tensorial (that is, covariant) equations acquire the form they would have in Special Relativity. This is the content of the Equivalence Principle.

§ **5.25** In the gauge theories describing the other fundamental interactions, an analogous property turns up: at each point of the curve/observer a gauge can be found in which the potential vanishes. However, it is not the potential but the field strength which appears in the force equation — the Lorentz force equation (1.18) is typical. The field strength is the curvature in these theories, and cannot be made to vanish. This is a crucial difference between gravitation and the other interactions.*

5.2.5 Two Gates into Gravitation

§ **5.26** Starting with a given nontrivial tetrad field $\{h_a\}$, two different but equivalent ways of describing gravitation are possible. In the first, according to (5.10), the nontrivial tetrad field is used to define a Riemannian metric $g_{\mu\nu}$, from which we can contruct the Levi–Civita connection and the corresponding curvature tensor. As the starting point was a nontrivial tetrad field, we can say that such a tetrad is able to induce a metric structure in spacetime, which is the structure underlying the General Relativity description of the gravitational field. We have seen that $\{h_a\}$ is not parallel-transported by the Levi-Civita connection.

On the other hand, a nontrivial tetrad field can be used to define a very special linear connection, called Weitzenböck connection, with respect to which the tetrad $\{h_a\}$ is parallel. For this reason, this kind of structure has received the name of teleparallelism, or absolute parallelism, and is the stage-set of the so called teleparallel description of gravitation. The important point to be kept from these considerations is that a nontrivial tetrad field is able to induce in spacetime both a teleparallel and a Riemannian structure. In what follows we will explore these structures in more detail.

---

* Details can be found in R. Aldrovandi, P. B. Barros e J. G. Pereira, *The equivalence principle revisited*, *Foundations of Physics* **33** (2003) 545-575 - ArXiv: gr-qc/0212034.

§ 5.27 Let us consider now the covariant derivative of the metric tensor $g_{\mu\nu}$. It is

$$\nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\lambda{}_{\mu\rho} g_{\lambda\nu} - \Gamma^\lambda{}_{\nu\rho} g^{\mu\lambda} \ ,$$

or, by using (5.24)

$$\nabla_\rho g_{\mu\nu} = h^a{}_\mu \, h^b{}_\nu \left(\Gamma_{ab\rho} + \Gamma_{ba\rho} \right) \ . \tag{5.54}$$

Therefore, the metricity condition

$$\nabla_\rho g_{\mu\nu} = 0 \tag{5.55}$$

will only hold when the connection is either purely antisymmetric [(pseudo) orthogonal],

$$\Gamma_{ab\rho} = - \, \Gamma_{ba\rho} \ ,$$

or when it vanishes identically:

$$\Gamma_{ab\rho} = 0 \ .$$

The Levi–Civita connection falls into the first case and, as we are going to see, the Weitzenböck connection into the second case. This means that both connections preserve the metric.

Chapter 6

Gravitational Interaction of the Fundamental Fields

6.1 Minimal Coupling Prescription

The interaction of a general field Ψ with gravitation can be obtained through the application of the so called minimal coupling prescription, according to which the Minkowski metric must be replaced by the riemannian metric

$$\eta^{ab} \to g^{\mu\nu} = \eta^{ab}\, h_a{}^\mu\, h_b{}^\nu \;, \tag{6.1}$$

and all ordinary derivatives must be replaced by Fock-Ivanenko covariant derivatives [1],

$$\partial_\mu \Psi \to \mathcal{D}_\mu \Psi = \partial_\mu \Psi - \frac{i}{2}\, \omega^{ab}{}_\mu\, S_{ab}\, \Psi \tag{6.2}$$

where $\omega^{ab}{}_\mu$ is a connection assuming values in the Lie algebra of the Lorentz group, usually called *spin connection*, and S_{ab} is a Lorentz generator written in a representation appropriate to the field Ψ. This "double" coupling prescription is a characteristic property of the gravitational interaction as for all other interactions of Nature only the derivative replacement (6.2) is necessary.

Let us explore deeper this point. The metric replacement (6.1) is a consequence of the local invariance of the lagrangian under *translations* of the tangent–space coordinates [2]. This part of the prescription, therefore, is related to the coupling of the field *energy–momentum* to gravitation, and is *universal* in the sense that it is the same for all fields. On the other hand,

the derivative replacement (6.2) is a consequence of the invariance of the lagrangian under local *Lorentz* transformations of the tangent–space coordinates. This part of the prescription, therefore, is related to the coupling of the field *spin* to gravitation, and is not *universal* because it depends on the spin contents of the field.

Another important point of the above coupling prescription is that the metric change (6.1) is appropriate only for integer–spin fields, whose lagrangian is quadratic in the field derivative, and consequently a metric tensor is always present to contract these two derivatives. For half-integer spin fields, however, the lagrangian is linear in the field derivative, and consequently no metric will be present to be changed. In this case, the metric change must be replaced by the equivalent rule in terms of the tetrad field,

$$e^a{}_\mu \longrightarrow h^a{}_\mu \, , \tag{6.3}$$

where $e^a{}_\mu$ is the trivial tetrad

$$e^a{}_\mu = \frac{\partial x^a}{\partial x^\mu} \, , \tag{6.4}$$

and $h^a{}_\mu$ is a nontrivial tetrad representing a true gravitational field.

Besides being more fundamental than the metric change (6.1), the tetrad change (6.3) allows the introduction of a *full* coupling prescription which encompasses both the metric — or equivalently the tetrad — and the derivative changes. It is given by

$$\partial_a \to \mathcal{D}_a \equiv h_a{}^\mu \, \mathcal{D}_\mu = h_a{}^\mu \left(\partial_\mu - \frac{i}{2} \, \omega^{ab}{}_\mu \, S_{ab} \right) \, , \tag{6.5}$$

This coupling prescription is general in the sense that it holds for both integer and half–integer spin fields. For the case of integer spin fields, it yields automatically the metric replacement (6.1). For the case of half–integer spin fields, it yields automatically the tetrad replacement (6.3).

6.2 General Relativity Spin Connection

As is well known, a tetrad field can be used to transform Lorentz into space-time indices, and *vice–versa*. For example, a Lorentz vector field V^a is related to the corresponding spacetime vector V^μ through

$$V^a = h^a{}_\mu V^\mu \, . \tag{6.6}$$

It is important to notice that this applies to tensors only. Connections, for example, acquire an extra *vacuum* term under such change [3]:

$$\omega^a{}_{b\nu} = h^a{}_\rho \omega^\rho{}_{\mu\nu} h_b{}^\mu + h^a{}_\rho \partial_\nu h_b{}^\rho .$$ (6.7)

On the other hand, because they are used in the construction of covariant derivatives, connections (or potentials, in physical parlance) are the most important personages in the description of an interaction. Concerning the specific case of the general relativity description of gravitation, the spin connection, denoted here by $\omega^a{}_{b\nu} = \mathring{A}^a{}_{b\nu}$, is given by [4]

$$\mathring{A}^a{}_{b\nu} = h^a{}_\rho \mathring{\Gamma}^\rho{}_{\mu\nu} h_b{}^\mu + h^a{}_\rho \partial_\nu h_b{}^\rho \equiv h^a{}_\rho \mathring{\nabla}_\nu h_b{}^\rho .$$ (6.8)

We see in this way that the spin connection $\mathring{A}^a{}_{b\nu}$ is nothing but the Levi–Civita connection

$$\mathring{\Gamma}^\rho{}_{\mu\nu} = \tfrac{1}{2} g^{\rho\lambda} \left[\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\lambda\mu} - \partial_\lambda g_{\mu\nu} \right]$$ (6.9)

rewritten in the tetrad basis. Therefore, the *full* coupling prescription of general relativity is

$$\partial_a \to \mathring{\mathcal{D}}_a \equiv h_a{}^\mu \mathring{\mathcal{D}}_\mu$$ (6.10)

with

$$\mathring{\mathcal{D}}_\mu = \partial_\mu - \frac{i}{2} \mathring{A}^{ab}{}_\mu S_{ab}$$ (6.11)

the general relativity Fock–Ivanenko [1] covariant derivative operator.

Now, comes an important point. The covariant derivative (6.11) applied to a general *Lorentz* tensor field reduces to the usual Levi-Civita covariant derivative of the corresponding *spacetime* tensor. For example, take again a vector field V^a for which the appropriate Lorentz generator is [5]

$$(S_{ab})^c{}_d = i \left(\delta^c{}_a \eta_{bd} - \delta^c{}_b \eta_{ad} \right) .$$ (6.12)

It is then an easy task to verify that [6]

$$\mathring{\mathcal{D}}_\mu V^a = h^a{}_\rho \mathring{\nabla}_\mu V^\rho .$$ (6.13)

On the other hand, no Levi–Civita covariant derivative can be defined for half-integer spin fields [7]. For these fields, the only possible form of the

163

covariant derivative is that given in terms of the spin connection. For a Dirac spinor ψ, for example, the covariant derivative is

$$\overset{\circ}{\mathcal{D}}_{\mu}\psi = \partial_{\mu}\psi - \frac{i}{2}\,\overset{\circ}{A}{}^{ab}{}_{\mu}\,S_{ab}\psi\ , \tag{6.14}$$

where

$$S_{ab} = \frac{1}{2}\sigma_{ab} = \frac{i}{4}[\gamma_a,\gamma_b] \tag{6.15}$$

is the Lorentz spin-1/2 generator, with γ_a the Dirac matrices. Therefore, we may say that the covariant derivative (6.11), which take into account the spin contents of the fields as defined in the tangent space, is more fundamental than the Levi–Civita covariant derivative in the sense that it is able to describe the gravitational coupling of both tensor and spinor fields. For tensor fields it reduces to the Levi–Civita covariant derivative, but for spinor fields it remains as a Fock–Ivanenko derivative.

6.3 Application to the Fundamental Fields

6.3.1 Scalar Field

Let us consider first a scalar field ϕ in a Minkowski spacetime, whose lagrangian is

$$\mathcal{L}_{\phi} = \frac{1}{2}\left[\eta^{ab}\,\partial_a\phi\,\partial_b\phi - \mu^2\phi^2\right]\ , \tag{6.16}$$

with

$$\mu = \frac{mc}{\hbar}\ . \tag{6.17}$$

The corresponding field equation is the so called Klein–Gordon equation:

$$\partial_a\partial^a\phi + \mu^2\phi = 0\ . \tag{6.18}$$

In order to get the coupling of the scalar field with gravitation, we use the *full* coupling prescription

$$\partial_a \to \overset{\circ}{\mathcal{D}}_a \equiv h_a{}^{\mu}\,\overset{\circ}{\mathcal{D}}_{\mu} = h_a{}^{\mu}\left(\partial_{\mu} - \frac{i}{2}\,\overset{\circ}{A}{}^{ab}{}_{\mu}\,S_{ab}\right)\ . \tag{6.19}$$

For a scalar field, however,

$$S_{ab} \, \phi = 0 \, , \tag{6.20}$$

and the coupling prescription in this case becomes

$$\partial_a \longrightarrow h_a{}^\mu \partial_\mu \, . \tag{6.21}$$

Applying this prescription to the lagrangian (6.16), we get

$$\mathcal{L}_\phi = \frac{\sqrt{-g}}{2} \left[g^{\mu\nu} \, \partial_\mu \phi \, \partial_\nu \phi - \mu^2 \phi^2 \right] \, . \tag{6.22}$$

Then, by using the identity

$$\partial_\mu \sqrt{-g} = \frac{\sqrt{-g}}{2} \, g^{\rho\lambda} \, \partial_\mu g_{\rho\lambda} \equiv \sqrt{-g} \, \overset{\circ}{\Gamma}{}^\rho{}_{\mu\rho} \, , \tag{6.23}$$

it is easy to see that the corresponding field equation is

$$\overset{\circ}{\Box} \phi + \mu^2 \phi = 0 \, , \tag{6.24}$$

where

$$\overset{\circ}{\Box} = \overset{\circ}{\nabla}_\mu \, \partial^\mu \equiv \frac{1}{\sqrt{-g}} \, \partial_\mu \left(\sqrt{-g} \, g^{\rho\mu} \, \partial_\rho \right) \tag{6.25}$$

is the Laplace–Beltrami operator, with $\overset{\circ}{\nabla}_\mu$ the Levi–Civita covariant derivative. We notice in passing that it is completely equivalent to apply the minimal coupling prescription to the lagrangian or to the field equations. Furthermore, we notice that, in a locally inertial coordinate system, the first derivative of the metric tensor vanishes, the Levi–Civita connection vanishes as well, and the Laplace–Beltrami becomes the free–field d'Alambertian operator. This is the usual version of the (weak) equivalence principle.

6.3.2 Dirac Spinor Field

In Minkowski spacetime, the spinor field lagrangian is

$$\mathcal{L}_\psi = \frac{ic\hbar}{2} \left(\bar{\psi} \, \gamma^a \partial_a \psi - \partial_a \bar{\psi} \gamma^a \, \psi \right) - mc^2 \, \bar{\psi}\psi \, . \tag{6.26}$$

The corresponding field equation is the Dirac equation

$$i\hbar\gamma^a \, \partial_a \psi - mc \, \psi = 0 \, . \tag{6.27}$$

165

In the context of general relativity, the coupling of a Dirac spinor with gravitation is obtained through the application of the *full* coupling prescription

$$\partial_a \to \mathring{\mathcal{D}}_a \equiv h_a{}^\mu \mathring{\mathcal{D}}_\mu = h_a{}^\mu \left(\partial_\mu - \frac{i}{2} \mathring{A}^{ab}{}_\mu S_{ab} \right) , \tag{6.28}$$

where now S_{ab} stands for the spin-1/2 generators of the Lorentz group, given by

$$S_{ab} = \frac{\sigma_{ab}}{2} = \frac{i}{4} [\gamma_a, \gamma_b] . \tag{6.29}$$

The spin connection, according to Eq.(6.8), is written in terms of the tetrad field as

$$\mathring{A}^a{}_{b\nu} = h^a{}_\rho \mathring{\Gamma}^\rho{}_{\mu\nu} h_b{}^\mu + h^a{}_\rho \partial_\nu h_b{}^\rho . \tag{6.30}$$

Applying the above coupling prescription to the free lagrangian (6.26), we get

$$\mathcal{L}_\psi = \sqrt{-g}\, c \left[\frac{ic\hbar}{2} \left(\bar\psi \gamma^\mu \mathring{\mathcal{D}}_\mu \psi - \mathring{\mathcal{D}}_\mu \bar\psi \gamma^\mu \psi \right) - m\,c^2\, \bar\psi\psi \right] , \tag{6.31}$$

where $\gamma^\mu = e_a{}^\mu \gamma^a$ is the local Dirac matrix, which satisfy

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{ab} h_a{}^\mu h_b{}^\nu = 2g^{\mu\nu} . \tag{6.32}$$

The corresponding Dirac equation can be obtained through the use of the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}_\psi}{\partial \bar\psi} - \mathring{\mathcal{D}}_\mu \frac{\partial \mathcal{L}_\psi}{\partial (\mathring{\mathcal{D}}_\mu \bar\psi)} = 0 . \tag{6.33}$$

The result is the Dirac equation in a Riemann spacetime

$$i\hbar \gamma^\mu \mathring{\mathcal{D}}_\mu \psi - mc\psi = 0 . \tag{6.34}$$

6.3.3 Electromagnetic Field

In Minkowski spacetime, the electromagnetic field is described by the lagrangian density

$$\mathcal{L}_{em} = -\frac{1}{4} F_{ab} F^{ab} , \tag{6.35}$$

166

where

$$F_{ab} = \partial_a A_b - \partial_b A_a \tag{6.36}$$

is the Maxwell field strength. The corresponding field equation is

$$\partial_a F^{ab} = 0 \ , \tag{6.37}$$

which along with the Bianchi identity

$$\partial_a F_{bc} + \partial_c F_{ab} + \partial_b F_{ca} = 0 \ , \tag{6.38}$$

constitute Maxwell's equations. In the Lorentz gauge $\partial_a A^a = 0$, the field equation (6.37) acquires the form

$$\partial_c \partial^c A^a = 0 \ . \tag{6.39}$$

In the framework of general relativity, the form of Maxwell's equations can be obtained through the application of the *full* minimal coupling prescription (6.5), which amounts to replace

$$\partial_a \to h_a{}^\mu \, \mathring{\mathcal{D}}_\mu = h_a{}^\mu \left(\partial_\mu - \frac{i}{2} \, \mathring{A}^{ab}{}_\mu S_{ab} \right) \ , \tag{6.40}$$

with

$$(S_{ab})^c{}_d = i \left(\delta_a{}^c \eta_{bd} - \delta_b{}^c \eta_{ad} \right) \tag{6.41}$$

the vector representation of the Lorentz generators. For the specific case of the electromagnetic vector field A^a, the Fock-Ivanenko derivative acquires the form

$$\mathring{\mathcal{D}}_\mu A^a = \partial_\mu A^a + \mathring{A}^a{}_{b\mu} \, A^b \ . \tag{6.42}$$

It is important to remark once more that the Fock–Ivanenko derivative is concerned *only* to the local Lorentz indices. In other words, it ignores the spacetime tensor character of the fields. For example, the Fock–Ivanenko derivative of the tetrad field is

$$\mathring{\mathcal{D}}_\mu h^a{}_\nu = \partial_\mu h^a{}_\nu + \mathring{A}^a{}_{b\mu} \, h^b{}_\nu \ . \tag{6.43}$$

Substituting

$$\mathring{A}^a{}_{b\mu} = h^a{}_\rho \mathring{\nabla}_\mu h_b{}^\rho \ , \tag{6.44}$$

167

we get

$$\overset{\circ}{\mathcal{D}}_\mu h^a{}_\nu = \overset{\circ}{\Gamma}{}^\rho{}_{\nu\mu}\, h^a{}_\rho \ . \tag{6.45}$$

As a consequence, the *total* covariant derivative of the tetrad $h^a{}_\nu$, that is, a covariant derivative which takes into account both indices of $h^a{}_\nu$, vanishes identically:

$$\partial_\mu h^a{}_\nu + \overset{\circ}{A}{}^a{}_{b\mu}\, h^b{}_\nu - \overset{\circ}{\Gamma}{}^\rho{}_{\nu\mu}\, h^a{}_\rho = 0 \ . \tag{6.46}$$

Now, any Lorentz vector field A^a can be transformed into a spacetime vector field A^μ through

$$A^\mu = h_a{}^\mu\, A^a \ , \tag{6.47}$$

where A^μ transforms as a vector under a general spacetime coordinate transformation. Substituting into equation (6.42), and making use of (6.45), we get

$$\overset{\circ}{\mathcal{D}}_\mu A^a = h^a{}_\rho\, \overset{\circ}{\nabla}_\mu A^\rho \ . \tag{6.48}$$

We see in this way that the Fock–Ivanenko derivative of a Lorentz vector field A^c reduces to the usual Levi–Civita covariant derivative of general relativity. This means that, for a vector field, the minimal coupling prescription (6.40) can be restated as

$$\partial_a A^c \to h_a{}^\mu\, h^c{}_\rho\, \overset{\circ}{\nabla}_\mu A^\rho \ . \tag{6.49}$$

Therefore, in the presence of gravitation, the electromagnetic field lagrangian acquires the form

$$\mathcal{L}_{em} = -\frac{1}{4}\,\sqrt{-g}\,F_{\mu\nu}F^{\mu\nu} \ , \tag{6.50}$$

where

$$F_{\mu\nu} = \overset{\circ}{\nabla}_\mu A_\nu - \overset{\circ}{\nabla}_\nu A_\mu \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \ , \tag{6.51}$$

the connection terms canceling due to the symmetry of the Levi–Civita connection in the last two indices. The corresponding field equation is

$$\overset{\circ}{\nabla}_\mu F^{\mu\nu} = 0 \ , \tag{6.52}$$

168

or equivalently, assuming the covariant Lorentz gauge $\overset{\circ}{\nabla}_\mu A^\mu = 0$,

$$\overset{\circ}{\nabla}_\mu \overset{\circ}{\nabla}^\mu A_\nu - \overset{\circ}{R}^\mu{}_\nu A_\mu = 0 \ . \tag{6.53}$$

Analogously, the Bianchi identity (6.38) can be shown to assume the form

$$\partial_\mu F_{\nu\sigma} + \partial_\sigma F_{\mu\nu} + \partial_\nu F_{\sigma\mu} = 0 \ . \tag{6.54}$$

We notice in passing that the presence of gravitation does not spoil the U(1) gauge invariance of Maxwell theory. Furthermore, like in the case of the scalar field, it results the same to apply the coupling prescription in the lagrangian or in the field equations.

Chapter 7

General Relativity with Matter Fields

7.1 Global Noether Theorem

Let us start by briefly reviewing the results of the global — or first — Noether's theorem [8]. As is well known, the global Noether theorem is concerned with the invariance of the action functional under global transformations. For each of such invariances, Noether's theorem determines a conservation law. In the specific case of the invariance under a global *translation* of the spacetime coordinates, the corresponding Noether conserved current is the *canonical* energy–momentum tensor

$$\theta^a{}_b = \frac{\partial \mathcal{L}_\Psi}{\partial \partial_a \Psi} \, \partial_b \Psi - \delta^a{}_b \, \mathcal{L}_\Psi \, , \tag{7.1}$$

with $\mathcal{L}_\Psi$ the lagrangian of the field Ψ.

On the other hand, in the case of the invariance under a global *rotation* of the spacetime coordinates — that is, under a Lorentz transformation — the corresponding Noether conserved current is the *canonical* angular–momentum tensor

$$\mathcal{J}^a{}_{bc} = \mathcal{M}^a{}_{bc} + \mathcal{S}^a{}_{bc} \, , \tag{7.2}$$

where

$$\mathcal{M}^a{}_{bc} = x_b \, \theta^a{}_c - x_c \, \theta^a{}_b \, , \tag{7.3}$$

is the *orbital* angular–momentum, and

$$\mathcal{S}^a{}_{bc} = i \, \frac{\partial \mathcal{L}_\Psi}{\partial \partial_a \Psi} \, S_{bc} \Psi \tag{7.4}$$

is the *spin* angular–momentum, with S_{bc} the generators of Lorentz transformations written in the representation appropriate for the field Ψ. Notice that $\mathcal{M}^a{}_{bc}$ is the same for all fields, whereas $\mathcal{S}^a{}_{bc}$ depends on the spin contents of the field Ψ.

Notice that the canonical energy–momentum tensor $\theta^a{}_b$ is not symmetric in general. However, using the Belinfante procedure [9] it is possible to define a symmetric energy–momentum tensor for the spinor field,

$$\Theta^{ab} = \theta^{ab} - \frac{1}{2} \partial_c \varphi^{cab}, \tag{7.5}$$

where

$$\varphi^{cab} = -\varphi^{acb} = \mathcal{S}^{cab} + \mathcal{S}^{abc} - \mathcal{S}^{bca} \, . \tag{7.6}$$

It can be easily verified that

$$\partial_c \varphi^{cab} = \theta^{ab} - \theta^{ba} \, , \tag{7.7}$$

which together with (7.5) show that Θ^{ab} is in fact symmetric.

7.2 Energy–Momentum as Source of Curvature

An old and controversial problem of gravitation is the conservation of energy–momentum density for both gravitational and matter fields. Concerning the energy–momentum tensor of matter fields, it becomes problematic mainly when spinor fields are present [10]. In order to explore deeper these problems, we are going to study the definition as well as the conservation law of the *gravitational* energy–momentum density of a general matter field. By *gravitational* energy–momentum tensor we mean the source of gravitation, that is, the tensor appearing in the right hand–side of the gravitational field equations. For the specific case of a spinor field, this energy–momentum tensor is sometimes believed to acquire a genuine non–symmetric part. As the left hand–side of the gravitational field equations are always symmetric, this would call for a generalization of general relativity. However, the

gravitational energy–momentum tensor is actually always symmetric, even for a spinor field, which shows the consistency and completeness of general relativity.

Let us consider the lagrangian

$$\mathcal{L} = \mathcal{L}_G + \mathcal{L}_\Psi \,, \tag{7.8}$$

where

$$\mathcal{L}_G = -\frac{c^4}{16\pi G} \sqrt{-g} \ \overset{\circ}{R} \tag{7.9}$$

is the Einstein–Hilbert lagrangian of general relativity, and $\mathcal{L}_\Psi$ is the lagrangian of a general matter field Ψ. The functional variation of $\mathcal{L}$ in relation to the metric tensor $g^{\mu\nu}$ yields the field equation

$$\overset{\circ}{R}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \ \overset{\circ}{R} = \frac{4\pi G}{c^4} \, \mathcal{T}_{\mu\nu} \,, \tag{7.10}$$

where

$$\mathcal{T}_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta \mathcal{L}_\Psi}{\delta g^{\mu\nu}} \tag{7.11}$$

is the *gravitational* energy–momentum tensor of the field Ψ. The contravariant components of the *gravitational* energy–momentum tensor is

$$\mathcal{T}^{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta \mathcal{L}_\Psi}{\delta g_{\mu\nu}} \,. \tag{7.12}$$

In these expressions,

$$\frac{\delta \mathcal{L}_\Psi}{\delta g^{\mu\nu}} = \frac{\partial \mathcal{L}_\Psi}{\partial g^{\mu\nu}} - \partial_\rho \frac{\partial \mathcal{L}_\Psi}{\partial_\rho \partial g^{\mu\nu}} \tag{7.13}$$

is the Lagrange functional derivative. In general relativity, therefore, energy and momentum are the source of gravitation, or equivalently, are the source of curvature. As the metric tensor is symmetric, the *gravitational* energy–momentum tensor obtained from either expression (7.11) or (7.12) is always symmetric. These expressions yield the energy–momentum tensor not only in the case of the presence of a gravitational field, but also in the absence. In the absence of a gravitational field, a transition to curvilinear coordinates must be done before the calculation of $\mathcal{T}^{\mu\nu}$. Of course, the metric tensor in this case will not represent a true gravitational field, but only effects of coordinates.

7.3 Energy–Momentum Conservation

Let us obtain now the conservation law of the *gravitational* energy–momentum tensor of a general source field Ψ. Denoting by $\mathcal{L}_\Psi$ the lagrangian of the field Ψ, the corresponding action integral is written in the form

$$S = \frac{1}{c} \int \mathcal{L}_\Psi \, d^4 x \; . \tag{7.14}$$

As a spacetime scalar, it does not change under a general transformation of coordinates. Of course, under a coordinate transformation, the field Ψ will change by an amount $\delta\Psi$. Due to the equation of motion satisfied by this field, the coefficient of $\delta\Psi$ vanishes, and for this reason we are not going to take these variations into account. For our purposes, it will be enough to consider only the variations in the metric tensor $g^{\mu\nu}$. Accordingly, by using Gauss theorem, and by considering that $\delta g^{\mu\nu} = 0$ at the integration limits, the variation of the action integral (7.14) can be written in the form [11]

$$\delta S = \frac{1}{c} \int \frac{\delta \mathcal{L}_\Psi}{\delta g^{\mu\nu}} \delta g^{\mu\nu} \, d^4 x = -\frac{1}{c} \int \frac{\delta \mathcal{L}_\Psi}{\delta g_{\mu\nu}} \delta g_{\mu\nu} \, d^4 x \; , \tag{7.15}$$

with $\delta \mathcal{L}_\Psi / \delta g^{\mu\nu}$ the Lagrange functional derivative (7.13). But, we have already seen in section 7.2 that

$$\frac{\delta \mathcal{L}_\Psi}{\delta g^{\mu\nu}} = \frac{\sqrt{-g}}{2} \, \mathcal{T}_{\mu\nu} \; , \tag{7.16}$$

where $\mathcal{T}_{\mu\nu}$ is the *gravitational* energy–momentum tensor of the field Ψ. Therefore, we have

$$\delta S = \frac{1}{2c} \int \mathcal{T}_{\mu\nu} \, \delta g^{\mu\nu} \, \sqrt{-g} \, d^4 x = -\frac{1}{2c} \int \mathcal{T}^{\mu\nu} \, \delta g_{\mu\nu} \, \sqrt{-g} \, d^4 x \; . \tag{7.17}$$

On the other hand, under a spacetime general coordinate transformation

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu \; , \tag{7.18}$$

with ϵ^μ small quantities, the components of the metric tensor change according to

$$\delta g_{\mu\nu} \equiv g'_{\mu\nu}(x^\rho) - g_{\mu\nu}(x^\rho) = -g_{\mu\lambda} \partial_\nu \epsilon^\lambda - g_{\lambda\nu} \partial_\mu \epsilon^\lambda - \partial_\lambda g_{\mu\nu} \epsilon^\lambda \; , \tag{7.19}$$

where only terms linear in the transformation parameter ϵ^μ have been kept. Substituting into (7.17), we get

$$\delta S = -\frac{1}{2c} \int \mathcal{T}^{\mu\nu} \left[-g_{\mu\lambda} \partial_\nu \epsilon^\lambda - g_{\lambda\nu} \partial_\mu \epsilon^\lambda - \partial_\lambda g_{\mu\nu} \epsilon^\lambda \right] \sqrt{-g} \, d^4 x \; . \tag{7.20}$$

173

Integrating by parts the second and third terms, neglecting integrals over hypersurfaces, and making use of the symmetry of $\mathcal{T}^{\mu\nu}$, we get

$$\delta S = -\frac{1}{c} \int \left[\partial_\nu (\sqrt{-g} \mathcal{T}^\nu{}_\lambda) - \frac{1}{2} \partial_\lambda g_{\mu\nu} \sqrt{-g} \, \mathcal{T}^{\mu\nu} \right] \epsilon^\lambda \, d^4 x \, , \qquad (7.21)$$

or equivalently

$$\delta S = -\frac{1}{c} \int \left[\partial_\nu (\sqrt{-g} \mathcal{T}^\nu{}_\lambda) - \overset{\circ}{\Gamma}{}^\rho{}_{\lambda\nu} \, \mathcal{T}^\nu{}_\rho \sqrt{-g} \right] \epsilon^\lambda \, d^4 x \, . \qquad (7.22)$$

Then, by using the identity

$$\partial_\nu \sqrt{-g} = \sqrt{-g} \, \overset{\circ}{\Gamma}{}^\mu{}_{\mu\nu} \, , \qquad (7.23)$$

we get

$$\delta S = -\frac{1}{c} \int \overset{\circ}{\nabla}_\nu \mathcal{T}^\nu{}_\lambda \, \epsilon^\lambda \sqrt{-g} \, d^4 x \, . \qquad (7.24)$$

Therefore, from both the invariance condition $\delta S = 0$ and the arbitrariness of ϵ^ρ, it follows that

$$\overset{\circ}{\nabla}_\nu \mathcal{T}^\nu{}_\lambda = 0 \, . \qquad (7.25)$$

It is important to remark that this is not a true conservation law in the sense that it does not lead to a *charge* conserved in time. Instead, it is an identity satisfied by the *gravitational* energy–momentum tensor, usually called Noether identity [8]. The sum of the energy–momentum of the gravitational field $t^\mu{}_\rho$ and the energy–momentum of the matter field $\mathcal{T}^\mu{}_\rho$ is a truly conserved quantity. In fact, this quatity satisfies

$$\partial_\mu \left[\sqrt{-g} \, (t^\mu{}_\rho + \mathcal{T}^\mu{}_\rho) \right] = 0 \, , \qquad (7.26)$$

which, by using Gauss theorem, yields the true conservation law

$$\frac{dq_\rho}{dt} = 0 \, , \qquad (7.27)$$

with

$$q_\rho = \int \left(t^0{}_\rho + \mathcal{T}^0{}_\rho \right) \sqrt{-g} \, d^3 x \qquad (7.28)$$

the conserved *charge*.

7.4 Examples

7.4.1 Scalar Field

Let us take the lagrangian of a scalar field in a Minkowski spacetime,

$$L_\phi = \frac{1}{2} \left[\eta^{ab} \, \partial_a \phi \, \partial_b \phi - \mu^2 \phi^2 \right] \, , \tag{7.29}$$

with μ given by (6.17). From Noether's theorem we find that the corresponding canonical energy–momentum and spin tensors are given respectively by

$$\theta^a{}_b = \partial^a \phi \, \partial_b \phi - \delta^a{}_b \, L_\phi \, , \tag{7.30}$$

and

$$\mathcal{S}^a{}_{bc} = 0 \, . \tag{7.31}$$

As a consequence of the vanishing of the spin tensor, the canonical energy–momentum tensor of the scalar field is symmetric, and is conserved in the ordinary sense:

$$\partial_a \theta^a{}_b = 0 \, . \tag{7.32}$$

In the presence of gravitation, the scalar field lagrangian is given by

$$\mathcal{L}_\phi = \frac{\sqrt{-g}}{2} \left[g^{\mu\nu} \, \partial_\mu \phi \, \partial_\nu \phi - \mu^2 \phi^2 \right] \, . \tag{7.33}$$

By using the identity

$$\delta \sqrt{-g} = \frac{\sqrt{-g}}{2} \, g_{\mu\nu} \, \delta g^{\mu\nu} \, ,$$

the *dynamical* energy–momentum tensor is found to be

$$\sqrt{-g} \, \mathcal{T}_{\mu\nu} = \sqrt{-g} \, \partial_\mu \phi \, \partial_\nu \phi - g_{\mu\nu} \, \mathcal{L}_\phi \, . \tag{7.34}$$

Like in the free case, it is symmetric, and conserved in the covariant sense:

$$\overset{\circ}{\nabla}_\mu \mathcal{T}^\mu{}_\nu = 0 \, . \tag{7.35}$$

7.4.2 Dirac Spinor Field

The Dirac spinor lagrangian in Minkowski spacetime is

$$L_\psi = \frac{ic\hbar}{2}\left(\bar\psi\gamma^a\partial_a\psi - \partial_a\bar\psi\gamma^a\psi\right) - mc^2\bar\psi\psi \ . \tag{7.36}$$

From the first Noether's theorem one finds that the corresponding canonical energy–momentum and spin tensors are given respectively by

$$\theta^a{}_b = \frac{ic\hbar}{2}\left(\bar\psi\gamma^a\partial_b\psi - \partial_b\bar\psi\gamma^a\psi\right) \ , \tag{7.37}$$

and

$$\mathcal{S}^a{}_{bc} = -\frac{c\hbar}{2}\left(\bar\psi\gamma^a S_{bc}\psi + \bar\psi S_{bc}\gamma^a\psi\right) , \tag{7.38}$$

with

$$S_{bc} = \frac{\sigma_{bc}}{2} = \frac{i}{4}[\gamma_b,\gamma_c] \ . \tag{7.39}$$

It should be noticed that, in contrast to the scalar field case, the canonical energy–momentum tensor for the Dirac spinor is not symmetric. As already discussed, however, we can use the Belinfante procedure [9] to construct a symmetric energy–momentum tensor for the spinor field, which is given by

$$\Theta^{ab} = \theta^{ab} - \frac{1}{2}\partial_c\varphi^{cab}, \tag{7.40}$$

with

$$\varphi^{cab} = -\varphi^{acb} = \mathcal{S}^{cab} + \mathcal{S}^{abc} - \mathcal{S}^{bca} \ . \tag{7.41}$$

In the presence of gravitation, the spinor field lagrangian is

$$\mathcal{L}_\psi = \sqrt{-g}\,c\left[\frac{i\hbar}{2}\left(\bar\psi\gamma^\mu\overset{\circ}{\mathcal{D}}_\mu\psi - \overset{\circ}{\mathcal{D}}{}^*{}_\mu\bar\psi\gamma^\mu\psi\right) - m\,c\,\bar\psi\psi\right] \ . \tag{7.42}$$

The *dynamical* energy–momentum tensor of the spinor field, according to the definition (7.11), is found to be

$$\mathcal{T}_{\rho\mu} = \tilde\theta_{\rho\mu} - \frac{1}{2}\overset{\circ}{\mathcal{D}}_\lambda\varphi_{\lambda\rho\mu} \ , \tag{7.43}$$

where

$$\tilde\theta_{\rho\mu} = \frac{ic\hbar}{2}\left(\bar\psi\gamma^\rho\mathcal{D}_\mu\psi - \mathcal{D}^*_\mu\bar\psi\gamma^\rho\psi\right) \tag{7.44}$$

176

is the canonical energy–momentum tensor modified by the presence of gravitation, and $\varphi_{\lambda\rho\mu}$ is still given by (7.6), but now written in terms of the spin tensor modified by the presence of gravitation:

$$\mathcal{S}^{\mu}{}_{bc} = \frac{c\hbar}{4} \left(\bar{\psi}\gamma^a h_a{}^\mu \sigma_{bc}\psi + \bar{\psi}\sigma_{bc}\gamma^a h_a{}^\mu\psi \right) . \tag{7.45}$$

Equation (7.43) is a generalization of the Belinfante procedure for the presence of gravitation. In fact, through a tedious but straightforward calculation we can show that

$$D_\mu \phi^{\mu\rho\lambda} = g^{\mu\lambda}\tilde{\theta}^\rho{}_\mu - g^{\mu\rho}\tilde{\theta}^\lambda{}_\mu , \tag{7.46}$$

from which we see that the *dynamical* energy–momentum tensor $\mathcal{T}^{\rho\lambda}$ of the Dirac field, that is, the Euler–Lagrange functional derivative of the spinor lagrangian (7.42) with respect to the metric, is always symmetric.

7.4.3 Electromagnetic Field

In Minkowski spacetime, the electromagnetic field is described by the lagrangian density

$$L_{em} = -\frac{1}{4} F_{ab}F^{ab} . \tag{7.47}$$

The corresponding canonical energy–momentum and spin tensors are given respectively by

$$\theta^a{}_b = -4\partial_b A_c F^{ac} + \delta^a{}_b F_{cd}F^{cd} , \tag{7.48}$$

and

$$\mathcal{S}^a{}_{bc} = F^a{}_b A_c - F^a{}_c A_b . \tag{7.49}$$

As in the spinor case, the canonical energy–momentum tensor $\theta^a{}_b$ of the electromagnetic field is not symmetric. By using the Belinfante procedure, however, it is possible to define the symmetric energy–momentum tensor,

$$\Theta^{ab} = \theta^{ab} - \frac{1}{2}\partial_c \varphi^{cab} , \tag{7.50}$$

where

$$\varphi^{cab} = -\varphi^{acb} = \mathcal{S}^{cab} + \mathcal{S}^{abc} - \mathcal{S}^{bca} . \tag{7.51}$$

As can be easily verified,

$$\varphi^{cab} = 2F^{ca}A^b \; . \tag{7.52}$$

Consequently,

$$\Theta^{ab} = 4\left[-F^{ac}\,F^b{}_c + \frac{1}{4}\,\eta^{ab}\,F_{cd}F^{cd}\right] \tag{7.53}$$

is in fact symmetric, and conserved in the ordinary sense:

$$\partial_a\theta^a{}_b = 0 \; . \tag{7.54}$$

In the presence of gravitation, the electromagnetic field lagrangian is

$$\mathcal{L}_{em} = -\frac{1}{4}\,\sqrt{-g}\,F_{\mu\nu}F^{\mu\nu} \; , \tag{7.55}$$

and the dynamical energy–momentum tensor

$$\mathcal{T}_{\mu\nu} = -\frac{2}{\sqrt{-g}}\,\frac{\delta\mathcal{L}_{em}}{\delta g^{\mu\nu}} \tag{7.56}$$

is found to be

$$\mathcal{T}_{\mu\nu} = \left[F_\mu{}^\rho\,F_{\nu\rho} - \frac{1}{4}\,g_{\mu\nu}\,F_{\rho\sigma}F^{\rho\sigma}\right] \; . \tag{7.57}$$

It is symmetric and covariantly conserved:

$$\overset{\circ}{\nabla}_\mu\,\mathcal{T}^{\mu\nu} = 0 \; . \tag{7.58}$$

Chapter 8

Closing Remarks

Gravitation differs from the other three known fundamental interactions of Nature by its more intimate relationship to spacetime. The other interactions (electromagnetic, weak and strong) are also described by (gauge) theories with a large geometrical content. However, while gravitation deals with changes of frames, the other interactions are concerned with changes of gauges in "internal" spaces. Gravitation relates to energy, while the other interactions cope with conserved quantities ("charges") which are independent of the events on spacetime — electric charge, weak isotopic spin and hypercharge, color.

In consequence, gravitation engender forces of inertial type, quite distinct from charge-produced forces. Hence its unique, universal character. Its presence is felt by all particles and fields in the same way — as if changing the very scene in which phenomena take place.

We hope to have given in these notes a first glimpse into the way these strange things happen.

Bibliography

[1] V. A. Fock, Z. Phys. **57**, 261 (1929).

[2] V. C. de Andrade and J. G. Pereira, Phys. Rev D **56**, 4689 (1997).

[3] R. Aldrovandi and J. G. Pereira, *An Introduction to Geometrical Physics* (World Scientific, Singapore, 1995).

[4] P. A. M. Dirac, in: *Planck Festscrift*, ed. W. Frank (Deutscher Verlag der Wissenschaften, Berlin, 1958).

[5] P. Ramond, *Field Theory: A Modern Primer*, 2nd edition (Addison-Wesley, Redwood, 1989).

[6] V. C. de Andrade and J. G. Pereira, Int. J. Mod. Phys. D **8**, 141 (1999).

[7] M. J. G. Veltman, *Quantum Theory of Gravitation*, in *Methods in Field Theory*, Les Houches 1975, Ed. by R. Balian and J. Zinn-Justin (North-Holland, Amsterdam, 1976).

[8] See, for example: N. P. Konopleva and V. N. Popov, *Gauge Fields* (Harwood, New York, 1980).

[9] F. J. Belinfante, Physica **6**, 687 (1939).

[10] K. Hayashi, Lett. Nuovo Cimento **5**, 529 (1972).

[11] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, Oxford, 1975).