Introduction to General Relativity

Janos Polonyi

University of Strasbourg, Strasbourg, France

(Dated: June 20, 2022)

Contents

I. INTRODUCTION

A distinguishing feature of gravitation compared with the other known interactions in Physics is that it touches upon the geometry of space and time. Correspondingly, the traditional approach to General Relativity is based on differential geometry. But there is another issue, locality which historically came up in connection with gravity first with fully recognized importance. Locality is evoked at two different stages in General Relativity. One is the independence of the equations of the theory from the choice of space-time coordinates. This renders the theory invariant under local reparametrization of the coordinate system. The equivalence principle, the existence of coordinate systems where gravitational forces disappear at a given point is another indication that locality plays more important role than thought before.

But locality actually represents a bridge to the other interactions. In fact, the interplay of locality and Special Relativity leads to gauge theories, a formal structure, common in all interactions known to us, gravity included. Inspired by this common structure General relativity is introduced below as a classical gauge field theory. After that a brief description of two applications is given, namely the Schwarzschild geometry and the Robertson-Walker geometry based cosmology.

Two appendices contain some complementary material, a short introduction into the formalism of classical field theory and the presentation of General Relativity as the gauge theory of the Poincaré group.

A. Equivalence principle

The unique status of gravitational forces among possible interactions is the surprising equivalence of the inertial mass appearing in Newton's third law in mechanics, m_{in}, and the coupling strength to a gravitational potential U_{gr}, the gravitational mass m_{gr},

$$m_{in}(\mathbf{a} - \mathbf{a}_{in}) = -m_{gr}\nabla U_{gr}(\mathbf{x}), \tag{1}$$

where $\mathbf{a}_{in}$ denotes the inertial acceleration, arising in non-Euclidean coordinate systems. Lorand Eötvös' measurement of the late nineteenth century and the improved versions performed later show the equivalence of gravitational and inertial masses with convincingly high accuracy. The Weak Equivalence Principle states that the world line of a small, free falling body is independent of its composition or structure, $m_{in} = m_{gr} = m$. As a result, the trajectory of a point particle is independent of its mass within Newton's theory.

The Strong Equivalence Principle consists of keeping the Weak Equivalence Principle valid in the presence of other, non-gravitational force, $\mathbf{F}_{ext}$,

$$m[\mathbf{a}-\mathbf{a}_{in}+\nabla U_{gr}(\mathbf{x})]=\mathbf{F}_{ext}. \tag{2}$$

This equation suggests a further equivalence, the identical origin of the gravitational and inertial forces. For instance, a homogeneous gravitational field, $U_{gr}(\mathbf{x})=gz$, can be eliminated by means of accelerating coordinate system, $z\to z-gt^2/2$, one is in a levitation, weightless state in the falling elevator.

The equivalence of inertial and gravitational forces is a local phenomenon, an inhomogeneous, time dependent gravitational potential can be eliminated at a given space-time point only by means of a suitable chosen space-time coordinates. In other words, the usual dynamics and symmetries, predicted by Special Relativity can be recovered in the absence of non-gravitational forces at any given point in space-time by means of a well chosen coordinate system. The Strong Equivalence Principle can be rephrased as the possibility of eliminating the gravitational forces at a given space-time point by a suitable chosen coordinate system. This principle holds only locally.

Yet another version of the Strong Equivalence Principle is that in a small enough region of the space-time it is impossible to detect the presence of a gravitational field and the laws of physics reduce to those of Special Relativity. In other words, the space-time and the physical laws can be made locally Lorentz-invariant. The mathematical details of locally observed Lorentz invariance will be spelled out below as a gauge symmetry.

Note that the gravitational forces are assumed to be originating from a simple scalar potential in this discussion, a restriction to be released later. But the principle is not flawless and has a limited validity, for instance quantum effects, related to the spin of the particles seem to represent an $\mathcal{O}(\hbar)$ violation. But quantum effects will be ignored in the rest of this notes. There are several, slightly different versions of the Equivalence Principle in the literature which confuses the picture even more. One may say that the Equivalence Principle is a rough guidance for our intuition which comes from the pre-gauge theory era of physics. Its simpler form and its limitation can easily be found when gravity is considered as a gauge theory.

B. Gravitation and geometry

The equivalence of the inertial and the gravitating mass makes that the only characteristic classical quantity of the point particle, its mass, drops out from the gravitational dynamics. But

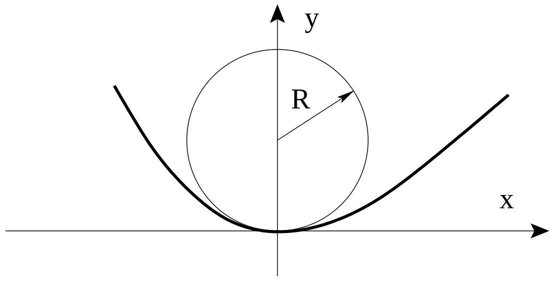


FIG. 1: The best matching circle of a curve.

what determines then the particle trajectory if not its mass? It has been suspected ever since the construction of Riemann-geometry that the gravitational force ought to be related to the curvature because it determines the geometry.

The curvature can be defined in its simplest form for a regular planar curve, $y(x)$, in the following manner: Let us look for approximations of the curve around a given point, x_0, with increasing precision. The zeroth order approximation is the constant, $y(x_0+\epsilon) \approx y(x_0)$. The linear approximation, $y(x_0+\epsilon) \approx y(x_0) + y'(x_0)\epsilon$, characterizes the tangent to the curve. The quadratic approximation, $y(x_0+\epsilon) \approx y(x_0) + y'(x_0)\epsilon + f''(x_0)\epsilon^2/2$, defines the curvature radius of the curve up to a sign as the radius of the best matching circle. The curvature radius has a non-trivial sign, it is positive (negative) if the circle in question is above (below) of the curve. To find the expression of the curvature radius let us imagine a circle, shown in Fig. 1. The equation of the circle,

$$R^2 = x^2 + (y - R)^2, \tag{3}$$

yelds

$$y = \frac{x^2}{2R} + \mathcal{O}\left(x^4\right) \tag{4}$$

for the lower hemisphere. If an at lest twice differentiable curve has vanishing value and slope at $x = 0$ then its curvature radius is

$$R = \frac{1}{\frac{d^2y}{dx^2}} \begin{cases} > 0 & \text{above,} \\ < 0 & \text{below.} \end{cases} \tag{5}$$

Let us now consider the free fall in the (x, y) plane, $x(t) = vt$, $y(t) = y_0 - gt^2/2$, with curvature radius

$$R = -\frac{v^2}{g}. \tag{6}$$

The gravitational force makes the particle to follow a trajectory with a given curvature. But the curvature contains the initial velocity, v, as well, hence it is not of purely geometric origin. It was Einstein's radically new point of view that the trajectory should be viewed in the space-time rather than in the space only where the curvature becomes independent of the initial conditions and can be assigned to geometry alone.

The inertial forces appear as a special manifestation of gravitational forces according to (1). This kind of gravitational forces appear without curvature. A simple example, discussed in Appendix A, is given by a family of uniformly accelerating world lines.

C. Static gravitational field

We have seen a distinguishing feature of gravitational forces, the equivalence principle above. We turn now to the similarity between static gravitational and electric forces, by comparing the Coulomb force

$$\mathbf{F}_C = \mathbf{r}\frac{e_1 e_2}{r^3}, \tag{7}$$

and Newton's gravitational law,

$$\mathbf{F}_g = -\mathbf{r}\frac{\tilde{m}_1\tilde{m}_2}{r^3}, \tag{8}$$

where e denote the electric charge and $\tilde{m} = m\sqrt{G}$, G being Newton's gravitational constant.

Despite their formal similarity the static electric and weak gravitational forces differ on two counts. One difference is that the gravitational charge, the mass can not be negative as opposed to the electric charge. This eliminates the possibility of screening of gravitational interaction. As a result, the gravitational forces remain long range like the unscreened Coulomb force and lead to instabilities and a number of interesting differences between the rules of statistical physics with and without gravitational forces. The other difference appears when the charges are not static but accelerate. The resulting electrodynamical and gravitational radiation are very different from each other.

The dynamical degree of freedom representing the electromagnetic interaction is the vector potential $A_\mu(x)$. The space-time geometry of classical General Relativity is characterized by the invariant length, defined by the metric tensor, $g_{\mu\nu}(x)$ introduced below. This tensor field can be considered as the dynamical degrees of freedom, responsible of gravitational interaction in classical physics.

The vector potential and the metric tensor describe elementary particles with spin one, $S = 1$, and spin two, $S = 2$, respectively. It is now understood that both the absence of gravitational repulsion and the complicated structure of gravitational radiation, compared to classical electrodynamics arise from the different spin of the carriers of the interaction.

An unifying concept, local gauge symmetries, streamlines the construction of classical theories for fields with non-vanishing spin. The minimalist version of a gauge theory, the Yang-Mills model will be introduced in the next Chapter.

D. Classical field theories

The need of classical fields appear as a way around an unexpected problem in Special Relativity. The non-relativistic equation of motion and the initial conditions for a system of N point particles, interacting by instantaneous, action-at-a-distance force are

$$\begin{aligned} m_a \frac{d^2\mathbf{x}_a(t)}{dt^2} &= \mathbf{F}_a(\mathbf{x}_1(t), \ldots, \mathbf{x}_N(t)), \\ \mathbf{x}_a(t_i) &= \mathbf{x}_{a,i}, \quad \frac{d\mathbf{x}_a(t)}{dt} = \mathbf{v}_{a,i}, \end{aligned} \tag{9}$$

where $a = 1, \ldots, N$, with initial conditions imposed at $t = t_i$. The relativistically covariant extension of these equations is given for the world lines, $x^\mu(s)$, parameterized by the invariant length s, called proper time, and read

$$\begin{aligned} m_a \ddot{x}^\mu_a(s_a) &= F^\mu_a(x_1(s_1), \ldots, x_N(s_N)), \\ x^\mu_a(s_{a,i}) &= x^\mu_{a,i}, \end{aligned} \tag{10}$$

where the dot stands for the derivation with respect to the proper time, the proper time of the particles is chosen in such a manner on the right hand side of the equation of motion that $x^0_a(s_a) = x^0_b(s_b)$ and the initial conditions are imposed at $x^0_a(s_i) = t_i$. The problem comes from the fact that that the four velocity preserves its length, $\dot{x}^2_a(s) = 1$ and the derivative with respect to the proper time gives the constraint $0 = \dot{x}_a\ddot{x}_a = \dot{x}_a F_a$ for each world line. One can show that there is no covariant function F^μ_a which satisfies this constraint.

The origin of the problem is that the instantaneous action-at-a-distance interaction requires propagation of signals with infinite velocities which is excluded by Special Relativity. The solution is to represent the interaction by means of a field variable, denoted by $\phi(\mathbf{x})$ for the sake of a simple example, a dynamical degrees of freedom at each space point. To make up the propagation of a signal we couple $\phi(\mathbf{x})$ to the fields of the neighboring points in a relativistically invariant manner.

A simple mechanical model of a scalar field in 1+1 dimensional space-time can be given by a chain of pendulums, coupled to their neighbors by a spring, depicted in Fig. 2. The Lagrangian is given by

$$L = \sum_n \left[\frac{mr^2}{2}\dot{\theta}_n^2 - \frac{kr_0^2}{2}(\theta_{n+1} - \theta_n)^2 - gr\cos\theta_n\right]. \tag{11}$$

The variable transformation $\theta_n(t) \to \Phi\theta_n(t) = \phi(t, x_n)$ brings it into the form

$$L = a\sum_n \left[\frac{1}{2c^2}(\partial_t\phi_n)^2 - \frac{1}{2}\left(\frac{\phi_{n+1} - \phi_n}{a}\right)^2 - \lambda\cos\frac{\phi_n}{\Phi}\right] \tag{12}$$

($\Phi = r_0\sqrt{ak}$, $c = a\frac{r_0}{r}\sqrt{\frac{k}{m}}$ and $\lambda = \frac{gr}{a}$). Next one goes into the continuum limit, $a \to 0$,

$$L = \int dx \left[\frac{1}{2c^2}(\partial_t\phi(x))^2 - \frac{1}{2}(\partial_x\phi(x))^2 - \lambda\cos\frac{\phi(x)}{\Phi}\right] \tag{13}$$

and finds the action of the sine-Gordon model,

$$S = \int dtdx \left[\frac{1}{2}\partial_\mu\phi(x)\partial^\mu\phi(x) - \lambda\cos\frac{\phi(x)}{\Phi}\right]. \tag{14}$$

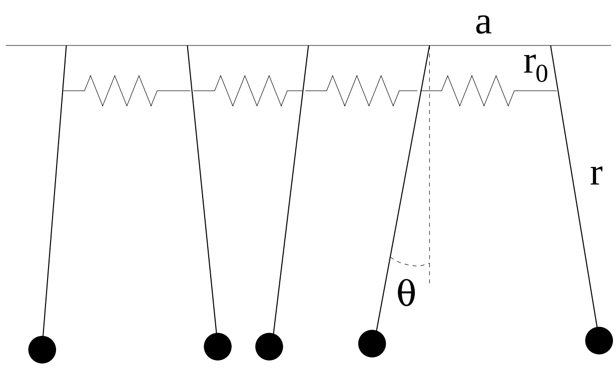


FIG. 2: The mechanical model of a 1+1 dimensional scalar field.

The local dynamics of a field is usually defined by means of its Lagrangian, cf. Appendix C. The simplest, the free real scalar field theory is defined by the Lagrangian

$$L = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2c^2}{2\hbar^2}\phi^2 \tag{15}$$

where $\partial_\mu = \frac{\partial}{\partial x^\mu}$ and $\lambda_C = \frac{\hbar}{mc}$ denotes the Compton wave-length of a particle of mass m. The action

$$S = \int dx \left[\frac{1}{2}\partial_\mu(x)\phi\partial^\mu\phi(x) - \frac{m^2c^2}{2\hbar^2}\phi^2(x)\right] \tag{16}$$

is manifestly Lorentz invariant. The normal modes, teh degrees of freedom which diagonalize the quadratic action, can be found by performing a Fourier transformation in space,

$$\begin{aligned}\phi_{\mathbf{k}}(x^0) &= \int d^3x\phi(x)e^{-i\mathbf{kx}},\\ \phi(x) &= \int \frac{d^3k}{(2\pi)^3}\phi_{\mathbf{k}}(x^0)e^{i\mathbf{kx}}\end{aligned} \tag{17}$$

and expressing the action in terms of the Fourier amplitude $\phi_{\mathbf{k}}(x^0)$,

$$\begin{aligned}S &= \int dx\frac{d^3kd^3q}{(2\pi)^6}\left[\frac{1}{2}\partial_0\phi_{\mathbf{k}}\partial_0\phi_{\mathbf{q}} + \frac{\mathbf{kq}}{2}\phi_{\mathbf{k}}\phi_{\mathbf{q}} - \frac{m^2c^2}{2\hbar^2}\phi_{\mathbf{k}}\phi_{\mathbf{q}}\right]e^{i(\mathbf{k}+\mathbf{q})\mathbf{x}}\\ &= \int dx^0\frac{d^3k}{(2\pi)^3}\left[\frac{1}{2}\partial_0\phi_{-\mathbf{k}}\partial_0\phi_{\mathbf{k}} - \frac{1}{2}\left(\mathbf{k}^2 + \frac{m^2c^2}{\hbar^2}\right)\phi_{-\mathbf{k}}\phi_{\mathbf{k}}\right],\end{aligned} \tag{18}$$

where the Fourier representation of the Dirac delta,

$$\delta(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}e^{i\mathbf{kx}}, \tag{19}$$

is used in the last equation. The field is real,

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3}\phi_{\mathbf{k}}(x^0)e^{i\mathbf{kx}} = \phi^*(x) = \int \frac{d^3k}{(2\pi)^3}\phi^*_{\mathbf{k}}(x^0)e^{-i\mathbf{kx}} \tag{20}$$

hence $\phi_{-\mathbf{k}}(x^0) = \phi^*_{\mathbf{k}}(x^0)$ and

$$S = \int dx^0\frac{d^3k}{(2\pi)^3}\left[\frac{1}{2}\partial_0\phi^*_{\mathbf{k}}\partial_0\phi_{\mathbf{k}} - \frac{1}{2}\left(\mathbf{k}^2 + \frac{m^2c^2}{\hbar^2}\right)\phi^*_{\mathbf{k}}\phi_{\mathbf{k}}\right]. \tag{21}$$

Relativistic normal modes are plane waves, $\phi_{\mathbf{k}}(x^0) = \phi_{1,\mathbf{k}}(x^0) + i\phi_{2,\mathbf{k}}(x^0)$, since

$$\begin{aligned}S[\phi_{\mathbf{k}}] &= \sum_{j=1,2}\int dx^0\left[\frac{1}{2}[\partial_0\phi_{j,\mathbf{k}}(x^0)]^2 - \frac{1}{2}\left(\mathbf{k}^2 + \frac{m^2c^2}{\hbar^2}\right)[\phi_{j,\mathbf{k}}(x^0)]^2\right]\\ &= \sum_{j=1,2}\int dx^0\left[\frac{M_{\mathbf{k}}}{2}[\partial_t\phi_{j,\mathbf{k}}(x^0)]^2 - \frac{M_{\mathbf{k}}\Omega^2_{\mathbf{k}}}{2}\phi_{j,\mathbf{k}}(x^0)\phi_{j,\mathbf{k}}(x^0)\right]\end{aligned} \tag{22}$$

with

$$M_{\mathbf{k}} = \frac{1}{c^2},\quad \Omega^2_{\mathbf{k}} = c^2\left(\mathbf{k}^2 + \frac{m^2c^2}{\hbar^2}\right),\quad E_{\mathbf{k}} = \hbar\Omega_{\mathbf{k}}. \tag{23}$$

The four-momentum of the normal mode, belonging to the wave vector $\mathbf{k}$, is

$$p = \left(\frac{E}{c},\mathbf{p}\right) = \left(\frac{\hbar c\sqrt{\mathbf{k}^2 + \frac{m^2c^2}{\hbar^2}}}{c},\hbar\mathbf{k}\right) = (\sqrt{m^2c^2 + \mathbf{p}^2},\mathbf{p}). \tag{24}$$

II. GAUGE THEORIES

We know four fundamental interactions in Nature, the gravitational, the electric, the weak and the strong forces. The last three appears in quantum version but we have so far no experimental evidence whatsoever indicating propagating classical gravitational field or any quantum effect in gravity. The remarkable fact is that all four interactions are described by the same type of models, by gauge field theories. These theories express a central assumption of physics in a distinguished manner, namely that the fundamental laws are local in the space-time. Non-gravitational gauge theories are introduced in this section in flat space-time.

A. Global symmetries

A transformation of the field configuration, $\phi^a(x) \to \phi'^a(x')$, where $x \to x'(x)$ is a symmetry if it changes the action $S[\phi]$ at most by total derivative,

$$S[\phi] \to S[\phi'] = S[\phi] + \int d^4x \partial_\mu \Lambda^\mu(x) \tag{25}$$

because such transformations leave the variational equations of motion unchanged.

It is advantageous to distinguish two different spaces in field theory. A field configuration of an n-component real field, $\phi(x) : E \to I$, is a mapping of the external space into the internal space, the former denoting the space-time $E = R^4$ and the latter standing for the set of values of the field, $I = R^n$.

A symmetry transformation can be internal, external or both. Internal symmetry transformations, $\phi^a(x) \to \phi'^a(x)$, act in the internal space only. Examples are charge conjugation, rotation in flavor or color space of quarks and leptons. External symmetry transformations change the space-time coordinates only, $\phi^a(x) \to \phi^a(x')$, eg. the Poincaré group, consisting of space-time translation and Lorentz transformations.

Continuous symmetries generate currents in classical field theories which satisfy the continuity equation according to Noether's theorem. The conserved quantity, the space integral of the time component of the Noether current is usually called charge in case of internal symmetry. The conserved quantities of external symmetries define energy-momentum (translation), angular momentum (space rotation) and a further vector (Lorentz boosts). Note that these symmetry transformations are global, meaning that they are characterized by the same parameters everywhere in the space-time.

B. Local symmetries

It has already been realized in the twenties by H. Weyl and been employed in constructing new theories by Yang and Mills in the sixties that the global symmetries of physics are in conflict with the spirit of special relativity. Let us consider for example a global internal symmetry, represented by the transformation

$$\psi(x) \to \psi^{\omega}(x) = \omega\psi(x) \tag{26}$$

acting an the multi-component field variable $\phi(x)$ with ω being an element of the symmetry group, typically $\omega \in G = O(n)$ (real fields) or $\omega \in G = U(n)$ (complex fields). The symmetry is global because we have to apply the same change of basis in the internal space anywhere and anytime in the Universe in order to keep the dynamics unchanged. Is the rigid application of the symmetry transformation really necessary in our world where special relativity holds? One can not exchange information between two locations in the space-time separated by space-like interval, $c^2\Delta t^2 - \Delta\mathbf{x}^2 < 0$ according to special relativity. How can then be a problem in using different bases in the description of physics at space-like separated regions? The symmetry transformations which seem to be in harmony with special relativity should concern change of basis in locations equipped with the possibility of the exchange of physical signals.

The suggestion is to give up any correlations among bases used at different space-time locations and to use local, so called gauge symmetries,

$$\psi(x) \to \psi^{\omega}(x) = \omega(x)\psi(x). \tag{27}$$

This is an extreme possibility, opposite to the global transformations. It creates obvious problems if applied to space-time regions with time-like separation, $\Delta t^2 - \Delta\mathbf{x}^2 > 0$ which can exchange signals. In particular, gauge symmetry makes it impossible to obtain any equation of motion for gauge non-invariant quantity which can freely be changed in an arbitrary time-dependent manner. This problem leads to the issue of gauge-fixing not pursued in this simple treatment.

The transformation rule (27) is homogeneous and has the virtue that any local equation

$$0 = F(\psi(x), \chi(x), \cdots) \tag{28}$$

which transforms in a homogeneous manner as in (27),

$$F^{\omega}(\psi(x), \chi(x), \cdots) = F(\psi^{\omega}(x), \chi^{\omega}(x), \cdots) = f(\omega(x))F(\psi(x), \chi(x), \cdots), \tag{29}$$

$f(\omega) \neq 0$, remains valid after any gauge transformation. These are called covariant or 'absolute' equations because they are valid in any convention, in other words in any gauge. We are interested in laws of Physics in as simple form as possible. Thus we seek absolute equations. Invariant quantities are called scalars and set of numbers or fields transforming in a homogeneous manner are usually called vectors or tensors. The rules of generating absolute equations consist of prescriptions of constructing scalars, vectors or tensors from scalars, vectors or tensors.

Let us, for the sake of example, consider an imaginary world consisting of two kind of particles, say particle 1 and 2, which participate in an identical manner in their interactions. The field variable has two components and the theory to start with displays a global symmetry group $G = O(2)$ or $G = U(2)$. The definition of the particle 1 or 2, amounts to a choice of a basis in the internal space. It is a convention used by physicists to construct models and communicate the results of their work.

Physicists at different laboratories may use different definitions, called in general conventions below. Experimental physicists need no basis since measurements are performed without making any reference to internal space. Nevertheless they need conventions as soon as they want to compare their findings with model predictions. In this imaginary world the physical phenomena are the same, independently of any choice of conventions. Hence there is a non-trivial condition that a function of the dynamical variabe be a measurable quantity, it must be gauge invariant.

C. Gauging

The main question for us in this section is to find the rules of modification of the theory in order to upgrade the global symmetry G into a local one. The result of this procedure, called gauging, is a theory with a gigantic symmetry group, $\mathcal{G} = \otimes \prod_x G_x$. We shall see that the price of such an enlargement of the symmetry is the introduction of a vector field, the gauge field.

Let us start with a theory defined by the Lagrangian $L(\phi, \partial\phi)$, cf. Appendix C, with global symmetry, $\omega \in G$. The Lagrangian has ultra-local terms, involving the field variable $\phi(x)$ at strictly the same space-time point, such as the mass term $\frac{1}{2}m^2\phi_a(x)\phi_a(x)$ or a local potential $U(\phi_a(x)\phi_a(x))$. There is no difference between global and local symmetry transformations as far as these terms are concerned. But pieces of the Lagrangian involving space-time derivative of the field are actually detecting the variation of the field on the space-time and are not strictly local. What is important from the point of view of the symmetry is that the transformation rule

$$\partial_\mu\phi(x) \to \partial_\mu\phi^\omega(x) = \partial_\mu\omega\phi(x) = \omega\partial_\mu\phi(x) \tag{30}$$

of the global symmetry transformation is modified for local symmetry briefly gauge transformations,

$$\partial_\mu \phi(x) \to \partial_\mu \phi^\omega(x) = \partial_\mu \omega(x)\phi(x) = \omega(x)\partial_\mu \phi(x) + (\partial_\mu \omega(x))\phi(x), \tag{31}$$

the trouble maker being the last term. It arises because the derivative compares the field values at neighboring points,

$$\partial_\mu \phi(x) = \lim_{\epsilon \to 0} \frac{\phi(x + \epsilon n_\mu) - \phi(x)}{\epsilon} \tag{32}$$

and this term represents the contribution due to the different conventions in different points. This contribution should not be there if by difference of the field variables we mean "physical" difference. We should transform the field variables into the same convention before subtraction. The expressing of the field at y into the convention of x, $\phi(y) \to \omega(x \leftarrow y)\phi(y)$, is a change of basis again. We are interested in this transformation for space-time points within each others vicinity when, the continuous dependence on the space-time coordinate assumed, this transformation is close to the identity. The possible moves of y into a neighboring x are characterized by an infinitesimal vector $\Delta x^\mu = x^\mu - y^\mu$ and the corresponding change of base,

$$\omega(y \leftarrow x) = 1\!\!1 - \Delta x^\mu A_\mu(x) + \mathcal{O}\left(\Delta^2 x\right), \tag{33}$$

is given in terms of four generators of the gauge group, $A_\mu(x)$ corresponding to the possible linearly independent moves of the point x. The use of a basis τ^a, $a = 1, \ldots, N$ for the Lie-algebra (generators) of an N-dimensional gauge group allows us to write

$$A_\mu(x) = A^a_\mu(x)\tau^a \tag{34}$$

where the basis matrices are assumed to satisfy the normalization conditions

$$\mathrm{tr}\ \tau^a \tau^b = -\frac{1}{2}\delta^{ab} \tag{35}$$

and commutation relations

$$\left[\tau^a, \tau^b\right] = f^{abc}\tau^c. \tag{36}$$

We have a vector field, a gauge field for each direction of the symmetry group following the change of basis one has to compensate for.

D. Covariant derivative

Once we have an expression for the compensation needed to bring the field around a space-time point into the convention at the same point we can define the covariant derivative as the derivative of the field $\phi(x)$ computed always in the convention at x,

$$\begin{aligned} D_\mu\phi(x) &= \lim_{\epsilon\to 0}\frac{e^{\epsilon n\cdot A(x+\epsilon n)}\phi(x+\epsilon n_\mu)-\phi(x)}{\epsilon} \\ &= \lim_{\epsilon\to 0}\frac{[1+\epsilon n\cdot A(x+\epsilon n)]\phi(x+\epsilon n_\mu)-\phi(x)}{\epsilon} \\ &= (\partial_\mu + A_\mu)\phi(x), \end{aligned} \tag{37}$$

giving

$$D_\mu = \partial_\mu + A_\mu. \tag{38}$$

The gauge field which appears in the definition of the covariant derivative is sometime called compensating field since its role is to compensate out the contributions of the inhomogeneous conventions from the derivative of a physical field.

Let us now find out the transformation rule for the gauge field $A_\mu(x)$ during the gauge transformation

$$\psi(x) \to \psi^\omega(x) = \omega(x)\psi(x). \tag{39}$$

The covariant derivative is the derivative of the field computed in fixed convention therefore $D_\mu\phi(x)$ transforms in the same way,

$$D_\mu\psi = (\partial_\mu + A_\mu)\psi \to D^\omega_\mu\psi^\omega = (\partial_\mu + A^\omega_\mu)\psi^\omega = \omega D_\mu\psi = \omega(\partial_\mu + A_\mu)\psi, \tag{40}$$

yielding

$$\omega(\partial_\mu + A_\mu)\psi = (\partial_\mu + A^\omega_\mu)\psi^\omega = (\partial_\mu\omega)\psi + \omega\partial_\mu\psi + A^\omega_\mu\omega\psi \tag{41}$$

and

$$A^\omega_\mu = -(\partial_\mu\omega)\omega^{-1} + \omega A_\mu\omega^{-1}. \tag{42}$$

Let us use the space-time derivative of the identity $\omega(x)\omega^{-1}(x) = \mathbb{1}$,

$$0 = (\partial_\mu\omega)\omega^{-1} + \omega\partial_\mu\omega^{-1}, \tag{43}$$

to write

$$A_\mu \to A_\mu^\omega = \omega(\partial_\mu + A_\mu)\omega^{-1}. \quad (44)$$

The transformation rule (40) gives the rule of replacing the partial derivative with covariant derivative in the Lagrangian,

$$L(\phi, \partial\phi) \to L(\phi, D\phi) = L(\phi, (\partial + A)\phi), \quad (45)$$

as the rule of gauging. The interaction induced in this manner between the particle described by the field ϕ and the gauge field is called minimal coupling. It should be clear by inspecting again the derivation of the conserved Noether-current in Chapter C 2 where the new coordinates, related to the global symmetry transformations are actually gauge transformation parameters that the minimal coupling involves the scalar product of the Noether-current and the gauge field, $A_\mu j^\mu$.

E. Parallel transport

Since two internal space vector residing at two different space-time locations can not be compared in their natural bases we need a definition what physically equivalent internal space vectors mean at different space-time points. This is achieved by the parallel transport, a generalization of the infinitesimal change of basis, given by (49).

Let us consider a continuously derivable path $\gamma^\mu : [0,1] \to \mathbb{R}^4$ in the space-time with $\gamma^\mu(0) = x_i^\mu$ and $\gamma^\mu(1) = x_f^\mu$ as initial and final points, respectively and a field $\phi(x)$ defined on this path. What is the condition that the values of this field long our path, $\phi(\gamma(s))$ are physically equivalent, despite the possible dependence of the components of $\phi(\gamma(s))$ on s when expressed in terms of the local basis? Suppose that we have a physical method to check the equivalence of the field along the path. The resulting field $\phi(\gamma(s))$ of a parallel transported internal space vector satisfies the equation

$$\phi(y) = W_\gamma(y, x)\phi(x) \quad (46)$$

where $W_\gamma(y, x)$ is a symmetry (basis) transformation which naturally depends on the choice of the points x and y and a somehow surprising manner will also depend on the path γ, too.

Since this function compensate the change of conventions along the path it should closely be related to the compensating field $A_\mu(x)$ introduced in defining the covariant derivative. In fact, parallel transport of a point of the internal space, vector in short, means that the "physical" components do not change along the path,

$$\frac{d\gamma^\mu}{ds} D_\mu \phi(\gamma(s)) = 0. \quad (47)$$

which can be written as an equation for the parallel transport transformation

$$\frac{d\gamma^\mu}{d\tau} D_{y^\mu} W_\gamma(y,x) = 0 \tag{48}$$

according to Eq. (46). The $\mathcal{O}(\Delta x)$ solution, (49), can be written by the help of the exponential map (B3) in the form

$$W_\gamma(x+\Delta x, x) = e^{-\Delta x^\mu A_\mu(x)}, \tag{49}$$

by assuming that the gauge field is constant in the straight line segment $[x, x+\Delta]$ of the path. The general solution, valid for a path of finite length is given in Appendix D.

It is easy to find out the transformation rule for parallel transport under gauge transformations. The starting point is that the product $\phi^\dagger(y)W_\gamma(y,x)\phi(x)$ or $\phi(y)W_\gamma(y,x)\phi(x)$ for complex or real ϕ, respectively, is gauge invariant. Hence the equation $\omega^{-1}(y)W^\omega_\gamma(y,x)\omega(x) = W_\gamma(y,x)$ follows and we have

$$W^\omega_\gamma(y,x) = \omega(y)W_\gamma(y,x)\omega^{-1}(x). \tag{50}$$

Two features of the parallel transport should be mentioned at this point:

- *Path dependence:* To understand the impact of the phase dependence of $W_\gamma(y,x)$ let us imagine first a gauge field $A_\mu(x)$ for which $W_\gamma(y,x)$ is *independent* of the path γ. In this case we may choose a reference point in space-time, say x_0 and extend its convention, its basis, to the rest of the space-time by performing the gauge transformation $\omega(x) = W^{-1}(x,x_0)$. In fact, this transformation renders all parallel transport trivial,

$$W'(y,x) = W^{-1}(y,x_0)W(y,x)W(x,x_0) = W(x_0,y)W(y,x_0) = \mathbb{1}. \tag{51}$$

  The gauge field which leads such a path independent parallel transports is called pure gauge because it can be canceled by an appropriate gauge transformation.

- *Parallel transport along closed paths:* An equivalent characterization of path independence of parallel transports is the triviality of parallel transport on any closed paths, $W_\gamma(x,x) = \mathbb{1}$. In fact, let us choose another point than x of the closed path γ what will be denoted by $y = \gamma(s)$, $0 < s < 1$. We introduce the fragments of γ from x to y and from y to x, as $\gamma_1(t) = \gamma(ts)$ and $\gamma_2(t) = \gamma(s + t(1-s))$ which satisfy the equation

$$W_\gamma(x,x) = W_{\gamma_1}(x,y)W_{\gamma_2}(y,x) = W_{\gamma_1}(x,y)W^{-1}_{\gamma_2}(x,y) \tag{52}$$

  the two parallel transports, $W_{\gamma_1}(x,y)$ and $W^{-1}_{\gamma_2}(y,x)$ correspond to two different paths connecting the events x and y therefore the path dependence is equivalent with the non-triviality of the parallel transports along closed paths.

F. Field strength tensor

The gauging, the upgrade of a global symmetry to a local one brings in a generator valued vector field. We are accustomed to the fact that fields corresponds to particles. Therefore the gauging of a symmetry suggests the presence of spin 1 bosons in the system. The dynamics of these particle can not come from the Lagrangian (45) because of the lack of the velocities $\partial_0 A_\mu$ in it. The simplest solution is the add a new term to the Lagrangian $L \to L + L_A$ where L_A satisfies the following conditions:

1. It should be **quadratic in the velocities**, $L_A = \mathcal{O}\left((\partial_0 A_\mu)^2\right)$.

2. It should be **Lorentz invariant**.

3. It should be **gauge invariant**.

The last property requires that L_A should be vanishing for pure gauge fields, it should depend on the non-pure gauge component of the field. This property suggests L_a, a local quantity, be constructed in terms of the deviation of parallel transports on infinitesimally small loops from the identity. Let us therefore consider the parallel transport of a field along a rectangle $x \to x+u \to x+u+v \to x+v \to x$ in space-time, u and v being infinitesimal, non-parallel vectors. The change of the field during the parallel transport is infinitesimal, as well, $\phi \to \phi + \delta\phi$ and should be linear in u, v and ϕ itself. Therefore one expects the relation

$$\delta\phi^a = -F^a_{\ b\mu\nu} u^\mu v^\nu \phi^b. \tag{53}$$

According to the definition of the parallel transport along an infinitesimal straight line, Eq. (49), the parallel transport along the rectangle is

$$U_\Box = e^{vA(x)} e^{uA(x+v)} e^{-vA(x+u)} e^{-uA(x)}. \tag{54}$$

We expand up to the displacement squares for each exponential functions gives

$$\begin{aligned} U_\Box &\approx \left(\mathbb{1} + vA(x) + \frac{1}{2}[vA(x)]^2\right)\left(\mathbb{1} + uA(x+v) + \frac{1}{2}[uA(x+v)]^2\right) \\ &\quad \times \left(\mathbb{1} - vA(x+u) + \frac{1}{2}[vA(x+u)]^2\right)\left(\mathbb{1} - uA(x) + \frac{1}{2}[uA(x)]^2\right) \end{aligned} \tag{55}$$

which can be simplified within this approximation to

$$\begin{aligned} U_\Box &\approx \mathbb{1} + (v\partial)uA - (u\partial)vA - (uA)(vA) + (vA)(uA) \\ &= \mathbb{1} - u^\mu v^\nu(\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]), \end{aligned} \tag{56}$$

resulting the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] = [D_\mu, D_\nu], \tag{57}$$

a generator valued field,

$$F_{\mu\nu} = F^a_{\mu\nu} \frac{\tau^a}{2i}, \tag{58}$$

with

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f^{abc} A^b_\mu A^c_\nu. \tag{59}$$

The gauge field strength is the measure of the non-triviality of the parallel transport of a state along a closed path.

The transformation rule for the parallel transport on a closed path,

$$\mathbb{1} - u^\mu v^\nu F_{\mu\nu}(x) \to \omega(x)[\mathbb{1} - u^\mu v^\nu F_{\mu\nu}(x)]\omega^{-1}(x) \tag{60}$$

gives the transformation rule

$$F_{\mu\nu}(x) \to \omega(x) F_{\mu\nu}(x) \omega^{-1}(x). \tag{61}$$

One can show that the triviality of the parallel transport over infinitesimal loops, the vanishing of the field strength tensor, assures the triviality of the parallel transport over arbitrary loops and restricts the gauge field to a pure gauge form, $A_\mu = \omega\partial_\mu\omega^{-1}$ in space-time without boundary conditions. When boundary conditions apply then the elimination of a pure gauge field configuration may be impossible. Such a configuration play an important role in the dynamics.

The unique solution of the constraints for L_A on a space-time with trivial topology is the Yang-Mills Lagrangian,

$$L_{YM} = \frac{1}{2g^2} \mathrm{tr}(F_{\mu\nu})^2, \tag{62}$$

fixed up to the coupling constant g. The normalization

$$\mathrm{tr}\ \tau^a \tau^b = -\frac{1}{2}\delta^{ab} \tag{63}$$

of the generators and commutation relations

$$[\tau^a, \tau^b] = f^{abc}\tau^c \tag{64}$$

of the Lea algebra of the gauge group yield

$$L_{YM} = -\frac{1}{4g^2}(F^a_{\mu\nu})^2. \tag{65}$$

It is advantageous to use the notation $A_\mu \to gA_\mu$ in perturbation expansion, giving rise to the Yang-Mills Lagrangian

$$L_{YM} = -\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu a}, \tag{66}$$

in terms of the field strength tensor

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu. \tag{67}$$

The coupling constant, appearing together with teh structure constant of teh gauge group indicates that the self-interaction of the gauge field is due to the non-commutativity of the gauge group.

G. Classical electrodynamics

Let us consider the electromagnetic interactions as the simplest toy model for General Relativity. case. We need the interaction for point charges, described by their world lines, $x^\mu_n(s)$, the index n identifying the particles. The electrodynamical interaction is the gauge theory which is based on local, gauged realization of the global phase symmetry of quantum mechanics, $\psi(x) \to e^{i\alpha}\psi(x)$. It is a $U(1)$ gauge theory and having a single gauge symmetry generator, i, the commutator term is vanishing in the field strength tensor (57). The action is

$$\begin{aligned} S &= -c\sum_n m_n \int ds_n - \frac{e}{c}\int dtd^3xj^\mu(x)A_\mu(x) - \frac{1}{16\pi}\int dtd^3xF_{\mu\nu}(x)F^{\mu\nu}(x) \\ &= -c\sum_n m_n \int ds_n - \frac{e}{c^2}\int d^4xj^\mu(x)A_\mu(x) - \frac{1}{16\pi c}\int d^4xF_{\mu\nu}(x)F^{\mu\nu}(x), \end{aligned} \tag{68}$$

where

$$j^\mu(x) = c\sum_n \int ds\delta(x - x_n(s))\dot{x}^\mu_n \tag{69}$$

stands for the electric current.

To find the equation of motion of the vector potential one writes the last term, Maxwell's action in the form

$$S_M = -\frac{1}{8\pi c}\int d^4x(\partial_\mu A_\nu\partial^\mu A^\nu - \partial_\mu A_\nu\partial^\nu A^\mu). \tag{70}$$

The variation equation for A_ν,

$$\begin{aligned} \frac{e}{c} j^\nu &= \frac{1}{4\pi} \partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) \\ &= \frac{1}{4\pi} \partial_\mu F^{\mu\nu}, \end{aligned} \tag{71}$$

is Maxwell's equation, to be generalized in general relativity to Einstein's equation.

To recover independent equations of motion for components of the the world line $x_n^\mu(s)$ we have to avoid the constraint $\dot{x}_n^2(s) = 1$ hence we start with a non-invariant length parametrization by replacing s by τ, $x(s) \to x(\tau)$ and write the first two terms of the action (68) in the form

$$S_{ch} = -\sum_n \int d\tau \left[m_n c \sqrt{\dot{x}^\mu(\tau) g_{\mu\nu} \dot{x}^\nu(\tau)} + \frac{e}{c} \dot{x}_n^\mu(\tau) A_\mu(x_n(\tau)) \right]. \tag{72}$$

The corresponding variational equation,

$$\begin{aligned} 0 &= -\frac{e}{c} \dot{x}_n^\nu(\tau) \partial_\mu A_\nu(x_n(\tau)) - \frac{d}{d\tau} \left[-mc \frac{\dot{x}_n^\mu(\tau)}{\sqrt{\dot{x}^\mu(\tau) g_{\mu\nu} \dot{x}^\nu(\tau)}} - \frac{e}{c} A_\mu(x_n(\tau)) \right] \\ &= mc \frac{\ddot{x}_n^\mu(\tau)}{\sqrt{\dot{x}^2(\tau)}} - \frac{e}{c} \dot{x}_n^\nu(\tau) [\partial_\mu A_\nu(x_n(\tau)) - \partial_\nu A_\mu(x_n(\tau))] + mc \frac{\dot{x}_n^\mu(\tau)}{[\dot{x}^2(\tau)]^{3/2}} \ddot{x}^\mu(\tau) \dot{x}_\mu(\tau), \end{aligned} \tag{73}$$

simplifies to the mechanical equation of motion with the Lorentz force,

$$mc \ddot{x}_n^\mu(s) = \frac{e}{c} F_{\mu\nu} \dot{x}_n^\nu(s), \tag{74}$$

when the invariant length is used to parameterized the world lines.

III. GRAVITY

It seems to be an essential feature of Nature that all known interactions belong to the class of gauge theories. For the electromagnetic, weak and strong interactions the internal space is independent of the space-time. The special feature of gravity is that it influences the geometry of the space-time therefore its internal space is not independent of its external space. The formalism of general relativity will be introduced below by underlying its origin in gauge theories.

A. Classical field theory on curved space-time

The mathematical view of a field configurations, $\phi(x)$, is a map $\phi : E \to I$ which describes "what" (I) happens "where" (E). Gravity and other interactions provide dynamics for E and I, respectively. Such a dynamics may generate singular x-dependence which can be avoided by renouncing the global, single valued nature of the dynamical fields. Instead, the fields are expected to produce a well defined image point in subsets of the space-time only and the resulting structure is called differentiable manifold. The regions $M_j \subset E$ where both the space-time coordinates and the fields have unique, well defined values are maps and the collection of maps, $\{M_j\}$, an atlas, is supposed to satisfy the following properties:

1. **Maps:** Each space-time point correspond at least to one map. ie. each space-time point can be identified by means of the coordinates.

2. **Coordinates:** There is a one-to-one mapping, $x_j : M_j \to V$, of each map into an open subset V of R^{d_E}. These d_E-dimensional functions, $x_j^\mu(p)$, $\mu = 1, \ldots, d_E$, $p \in E$, play the role of coordinates defined for each map and the space-time looks locally d_E-dimensional, in agreement with the Equivalence Principle. The inverse function, $p(x)$, labels the vicinity of p by the coordinates x. We assume $d_E = 4$ in what follows. Experimental devices are supposed to be available to measure the values of the coordinates at each space-time point.

3. **Coordinate transformations:** The space-time point of a given environment, U, may belong to several maps. If $p \in U$ can be represented by two maps, $x(p)$ and $x'(p)$ then the inverse function, $p(x)$, defined on an open set in R^4, establishes the coordinate transformation, $x'(x) = x'(p(x))$, which is supposed to be infinitely many times differentiable.

The **internal space**, a linear space consisting of the possible field values, is called tangent space in the case of gravity. This name shows the origin of the gravitational internal space T_p: the space of possible space-time directions at the given space-time point. A more formal, coordinate system independent definition is that T_p consists of the equivalence classes of world lines $x(s)$, crossing p, $x(s_p) = x(p)$ where $x(s)$ and $x'(s)$ are considered equivalent if they have the same tangent vector in p, $\dot{x}(s_p) = \dot{x}'(s_p)$. We assume that experimental devices are available at each space-time point to measure the four velocity of point particles and identify a point in the tangent space. The gravitational interaction determines the motion masses, therefore it governs the dynamics in T. An important result of this particular construction of the tangent space is that the infinitesimal vectors $u \in T_p$, can be related to the infinitesimal displacement, $p \to p'$,

$$x_j^\mu(p') = x_j^\mu(p) + u^\mu. \tag{75}$$

Let us choose a standard map, $X^\mu(p)$, for a given environment U. The basis vectors of T_p, corresponding to a given map $x(p)$, are

$$e_\mu = \frac{\partial X}{\partial x^\mu}, \tag{76}$$

in particular

$$e_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_{d-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{77}$$

in the case of the standard map. The change of the coordinate system, $x \to x' = x'(x)$, leads to a new basis in T_p, defined by the equation

$$e_\mu = \frac{\partial x}{\partial x^\mu} = \frac{\partial x}{\partial x'^\nu}\frac{\partial x'^\nu}{\partial x^\mu} = e'_\nu \frac{\partial x'^\nu}{\partial x^\mu}. \tag{78}$$

It is easy to check that the corresponding transformation rule of a contravariant vector is

$$e^\mu = \frac{\partial x^\mu}{\partial x'^\nu} e'^\nu \tag{79}$$

These transformation rules can easily be memorized by the following observations: One could have derived eq. (78) by using (76) by dropping $X(x)$. The result is that the covariant vectors transform as partial derivatives,

$$\partial_\mu = \partial'_\nu \frac{\partial x'^\nu}{\partial x^\mu}. \tag{80}$$

Since $dx^\mu \partial_\mu$ remains invariant under coordinate transformations the contravariant vectors transform as the infinitesimal coordinate changes,

$$\delta x^\mu = \frac{\partial x^\mu}{\partial x'^\nu} \delta x'^\nu. \tag{81}$$

This observation is the starting point of an elegant formalism in differential geometry. It sometime is called the absolute calculus since these objects can formally be defined without using a standard (local) coordinate system. Equations which preserve their form under the reparametrization of the coordinates are called covariant. The absolute calculus produces covariant equations, constructed in terms of functions (coordinates), their derivatives (covariant vector fields) and infinitesimal changes (contravariant vector fields). One arrives at nice and compact equations in this manner, showing the geometrical essence in as clear manner as possible. But this scheme has two drawbacks from the point of view of physics. One is that one is interested in the agreement between observed and theoretically predicted quantities and such a comparison can not be made without actually "dirtying our hand" with a given coordinate system. The other is that while the idea of representing vector fields by partial derivatives or by infinitesimal changes is correct and justified mathematically it leads to a wrong intuition in physics. I believe that it is more advantageous first to spend the time needed by improving our way of reading and understanding equations, written in terms of a coordinate system until we recognize the general structure. When the physics is properly expressed then, as a second step, one may go over the coordinate independent scheme if one wishes.

B. Geometry

There are three important properties of the space-time geometry which appears in gravitational interactions.

1. The **metric** property is related to the existence of an invariant distance and it is essential in establishing spatial and temporal distances between events. It is assumed that physical measurements provide us spatial distances, time intervals and angles to determine the metric. Due to the Equivalence Principle the metric structure must locally be compatible with Special Relativity. This is achieved by introducing the invariant length in a local manner,

$$ds^2(x) = dx^\mu g_{\mu\nu}(x) dx^\nu \tag{82}$$

   where the metric tensor $g_{\mu\nu}(x)$ is a symmetric tensor with three positive and a negative eigenvalues, i.e. with signature $+, -, -, -$.

2. The **affine** property controls parallelism by determining what directions can we consider parallel at different space-time points. The velocity and the internal angular momentum, the classical spin, of a small enough body is constant in the absence of external forces and the acceleration and the time derivative of the spin are vanishing. According to the Equivalence Principle this remains valid locally in a suitable chosen coordinate system in the presence of external gravitational field. Thus the velocity and the spin preserve their directions in a non-trivial geometry and can be used to establish the concept of parallelism. The affine structure is realized by the affine connection, alias compensating or gauge field of gauge theories.

3. The **torsion** of the space-time represents certain distortions of the space-time. This property seems to be important in describing the interaction of the quantum mechanical spin with the

gravitational field only. Note that thought the Equivalence Principle has been build in by point 2. in Section III A, it can eventually violated by implying the torsion in the dynamics because the distortions, characterized by the torsion can not be eliminated by using suitable local coordinate system. We consider General Relativity in classical physics and the torsion will be assumed to be absent.

C. Gauge group

It is instructive to recall that the absolute location in space-time becomes unobservable and the space-time location is relative when the dynamics is translation invariant. The spatial directions are relative in rotational invariant systems. The symmetry with respect to Lorentz boost makes the absolute velocity unobservable in Special Relativity. The acceleration and all higher order derivatives of the world line are rendered relative in General Relativity by imposing invariance of the dynamics under reparametrization of the space-time coordinates, called space-time diffeomorphism. Hence the choice of the coordinate system in the space-time is mere convention, the role of the coordinates is to identify space-time points only and the actual numerical values of the coordinate have no physical meaning.

The fundamental physical laws. eg. the Maxwell-equations, are supposed to be local in space-time, they can be expressed as equations among dynamical quantities, corresponding to the same space-time location. There is indeed no need of absolute coordinates for such a description. When we do not intend to follow and find the equation of motion for all dynamical degree of freedom then we seek an effective description. The motion of point particles or propagation of waves which are not followed by us relate and correlate dynamical quantities at different space-time location and are used to introduce direction and distance in space-time. We are interested here the fundamental laws therefore we assume that the equations of motions are expressed for each space-time location independently. This assumption leads to a rich gauge theory structure in an obvious manner and gravity can be founded as a gauge theory with either external or internal symmetries.

1. Space-time diffeomorphism

The reparametrization of the coordinates, the diffeomorphism of the external space, can be generated by the infinitesimal local translations in space time, $x^\mu(x) \to x'^\mu(x) = x^\mu(x) + \delta x^\mu(x)$. If the value of the coordinates are not relevant then the change of the coordinate system is represented by the change of the tangent vectors of world lines, in particular the coordinate axes. Therefore, the gauge group is $GL(4)$, consisting of 4×4 non-singular matrices describing the transformation of the coordinate axes $e_\mu(x)$ during general coordinate transformations,

$$e_\mu(x) \to e'_\mu(x) = M_\mu{}^\nu(x) e_\nu(x). \tag{83}$$

The affine structure of the space-time which is the central feature of this formalism handles the relation of directions at different space-time points by means of parallel transport in the $GL(4)$ gauge theory. The drawback of this line of thought is that the other independent geometrical structure, the metric which is a key player in the traditional approach to General Relativity is constructed in an indirect manner from the affine connection.

The internal space is chosen to be the tangent space, the directional vectors at each point of the space-time and its elements are contravariant vectors. The independent field variables are the metric tensor $g_{\mu\nu}$ and the $GL(4)$ gauge field, $\Gamma^\rho{}_{\mu\nu}$, the affine connection. The metric tensor is symmetrical, contains 10 independent component and has the signature $(+,-,-,-)$. Once the metric tensor is introduced the tangent spaces can be represented covariant vectors, too. Furthermore, the direct product of the tangent space givers rise to the space of local tensors, as well.

The representation of the diffeomorphism by the transformation of the tangent spaces, (83) with $M_\mu{}^\nu(x) = \delta_\mu^\nu - \partial_\mu \delta x^\nu$ raises a consistency issue, namely what conditions should the four vector fields, $e_\mu(x)$ satisfy to make up a coordinate basis? The necessary condition arises from the symmetry of the second partial derivatives, $\partial_\mu e_\nu = \partial_\mu \partial_\nu x' = \partial_\nu \partial_\mu x' = \partial_\nu e_\mu$. It is easy to check that this condition is sufficient. In fact,

$$\partial_\mu e_\nu = \partial_\nu e_\mu \tag{84}$$

is sufficient to assure the local existence and unicity of integral curves, the solution of the equations $\partial_\mu x'(x) = e_\mu(x)$. The tetrads, satisfying (84) and can be used to construct well defined local coordinates are called holonomic. The possibility of representing the reparametrization of the space-time by the transformation of the coordinate basis vectors stems from the preservation of the holonomy under diffeomorphism.

2. Internal Poincaré group

Both the metric and the affine structures can be derived in the gauge theory formalism by means of internal gauge symmetry. The starting point is the Equivalence Principle which assures the existence of a coordinate system with local Lorentz invariance.

Lorentz transformations: The local Lorentz coordinate axes, $e^a(x)$, can be considered as basis in the internal Lorentz space. But one might as well use another reference frame, obtained by a local Lorentz transformation,

$$e^a(x) \to e'^a(x) = \omega^a{}_b(x) e^b(x). \tag{85}$$

Thus one is led to propose Lorentz transformation as a local symmetry. A gauge field, $\omega^a{}_{b\mu}$, introduced to handle the compensations of the local Lorentz transformations defines the affine structure of the space-time.

Translations: The special feature of gravity is the relation between dynamics and geometry, a link between the internal and the external spaces. The external diffeomorphism, the coordinate reparametrization invariance can be generated by the infinitesimal local translations in space time, $x^\mu(x) \to x'^\mu(x) = x^\mu(x) + \delta x^\mu(x)$. The tangent space, T_p, consists of tangent vectors of world lines, $\dot{x}^\mu$, at p. By representing the tangent vectors in the local Lorentz reference frame, $\dot{x}^\mu = \dot{\xi}^a e^\mu_a$, we turn this latter into the tangent space, T_p. The internal space equivalent of local, infinitesimal translations,

$$\xi^a(x) \to \xi'^a(x) = \xi^a(x) + \delta\xi^a(x) \tag{86}$$

is now considered as a local gauge transformation. The gauge field compensating such a local modification of the coordinates of the local Lorentz spaces, $e^a_\mu(x)$ is called vierbein and provides the desired link between infinitesimal shifts in the internal Lorentz space and the space-time:

$$\delta\xi^a(x) = e^a_\mu \delta x^\mu(x). \tag{87}$$

We shall use Greek and Latin letters to denote vector indices in the Lorentz and in the space-time coordinate system, respectively. The translations (86), together with the Lorentz transformations (85) form the Poincaré group as gauge symmetry.

An unexpected bonus of the Poincaré group formalism is the possibility of treating fermions. In fact, fermions do not have well covariant transformations properties under the $GL(4)$ group, the general change of coordinates. They show well defined transformation rules for the Lorentz group only. Another advantageous feature of this formalism is the natural way torsion couples to angular momentum in the dynamics.

D. Gauge theory of diffeomorphism

The simpler framework for gravity, based on the external diffeomorphism as gauge group is introduced first.

1. *Covariant derivative*

The affine structure is defined by the connection Γ_μ, a 4×4 matrix valued vector field with 64 independent components. We shall use the notation $(\Gamma_\rho)^\mu{}_\nu = \Gamma^\mu{}_{\nu\rho}$ for the components of the connection. The covariant derivative,

$$D_\nu v^\mu = \partial_\nu v^\mu + \Gamma^\mu{}_{\rho\nu} v^\rho, \tag{88}$$

detects the 'real', physical changes of a vector field by projecting out the changes of the vector components arising from the changing conventions. By suppressing the indices in Eq. (88) we have for contravariant vectors

$$(D_\nu v)^\mu = (\partial_\nu v + \Gamma_\nu v)^\mu. \tag{89}$$

The metric structure is behind the reduplication of the vectors and suggests the definition

$$(D_\nu v)_\mu = (\partial_\nu v - v\Gamma_\nu)_\mu \tag{90}$$

for covariant vector field. The reason is that this definition allows us the contraction of indices within the covariant derivative and to recover the equivalence of the covariant and the partial derivatives for scalar fields,

$$D_\mu(u_\nu v^\nu) = (\partial_\mu u - u\Gamma_\mu)_\nu v^\nu + u_\nu(\partial_\mu v + \Gamma_\mu v)^\nu = \partial_\mu(u_\nu v^\nu) \tag{91}$$

The action of the covariant derivative is extended over any tensor field by performing the necessary compensation on each vector index, eg.

$$D_\nu v^\mu_\rho = \partial_\nu v^\mu_\rho + \Gamma^\mu_{\kappa\nu} v^\kappa_\rho - v^\mu_\kappa \Gamma^\kappa_{\rho\nu}. \tag{92}$$

It is easy to check that such an extension reproduces Leibnitz's rule,

$$D_\mu(u^\rho v^\sigma) = (D_\mu u^\rho) v^\sigma + u^\rho D_\mu(v^\sigma). \tag{93}$$

Notice that a coordinate transformation $x \to x'(x)$ induces the change

$$\Gamma^\mu{}_{\nu\rho} \to \Gamma'^\mu{}_{\nu\rho} = \frac{\partial x^\sigma}{\partial x'^\rho} \omega^\mu{}_\kappa (\delta^\kappa_\lambda \partial_\sigma + \Gamma^\kappa{}_{\lambda\sigma})(\omega^{-1})^\lambda{}_\nu. \tag{94}$$

according to the transformation rule (44) and the application of the second equation in (78) to the connection as a four-vector. The expressions $\omega^\mu{}_\kappa = \frac{\partial x'^\mu}{\partial x^\kappa}$ and $(\omega^{-1})^\mu{}_\kappa = \frac{\partial x^\mu}{\partial x'^\kappa}$ allow us to write

$$\begin{aligned} \Gamma^\mu{}_{\nu\rho} &\to \frac{\partial x^\sigma}{\partial x'^\rho}\frac{\partial x'^\mu}{\partial x^\kappa}\frac{\partial^2 x^\kappa}{\partial x'^\nu \partial x^\sigma} + \frac{\partial x^\sigma}{\partial x'^\rho}\frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa{}_{\lambda\sigma}\frac{\partial x^\lambda}{\partial x'^\nu} \\ &= \frac{\partial x'^\mu}{\partial x^\kappa}\frac{\partial^2 x^\kappa}{\partial x'^\nu \partial x'^\rho} + \frac{\partial x^\sigma}{\partial x'^\rho}\frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa{}_{\lambda\sigma}\frac{\partial x^\lambda}{\partial x'^\nu}, \end{aligned} \tag{95}$$

showing that the affine connection is not a tensor due to the inhomogeneous term in the gauge transformation (44), the first term on the right hand side of (95). But the antisymmetric part in the covariant indices, called torsion,

$$S^\rho_{\nu\mu} = \frac{1}{2}(\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu}). \tag{96}$$

is a tensor. The form (42) of the gauge transformation is sometime useful,

$$\begin{aligned} \Gamma^\mu{}_{\nu\rho} \to \Gamma'^\mu{}_{\nu\rho} &= -\frac{\partial x^\sigma}{\partial x'^\rho}\partial_\sigma \omega^\mu{}_\kappa \delta^\kappa_\lambda (\omega^{-1})^\lambda{}_\nu + \frac{\partial x^\sigma}{\partial x'^\rho}\omega^\mu{}_\kappa \Gamma^\kappa{}_{\lambda\sigma}(\omega^{-1})^\lambda{}_\nu \\ &= -\frac{\partial x^\kappa}{\partial x'^\nu}\frac{\partial^2 x'^\mu}{\partial x^\kappa \partial x^\sigma}\frac{\partial x^\sigma}{\partial x'^\rho} + \frac{\partial x^\sigma}{\partial x'^\rho}\frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa{}_{\lambda\sigma}\frac{\partial x^\lambda}{\partial x'^\nu}. \end{aligned} \tag{97}$$

It is advantageous to use harmonic gauge, defined by

$$\Gamma^\rho = g^{\mu\nu}\Gamma^\rho{}_{\mu\nu} = 0 \tag{98}$$

for solving the equations of motion. The name comes from the equation

$$\Box x^\mu = g^{\rho\nu} D_\rho D_\nu x^\mu = g^{\nu\rho} D_\rho \partial_\nu x^\mu = -\Gamma^\mu \tag{99}$$

where the second equation holds because x^μ is a scalar field for a given value of μ, stating that the coordinates are harmonic functions. Eq. (97) leads to the transformation rule

$$\begin{aligned} \Gamma^\mu \to g'^{\nu\rho}\Gamma'^\mu{}_{\nu\rho} &= g^{\tau\sigma}\frac{\partial x'^\nu}{\partial x^\tau}\frac{\partial x'^\rho}{\partial x^\sigma}\left(-\frac{\partial x^\kappa}{\partial x'^\nu}\frac{\partial^2 x'^\mu}{\partial x^\kappa \partial x^\sigma}\frac{\partial x^\sigma}{\partial x'^\rho} + \frac{\partial x^\sigma}{\partial x'^\rho}\frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa{}_{\lambda\sigma}\frac{\partial x^\lambda}{\partial x'^\nu}\right) \\ &= -g^{\tau\sigma}\frac{\partial^2 x'^\mu}{\partial x^\tau \partial x^\sigma} + \frac{\partial x^\mu}{\partial x'^\kappa}\Gamma^\kappa \end{aligned} \tag{100}$$

which gives the equation

$$g^{\tau\sigma}\frac{\partial^2 x'^\mu}{\partial x^\tau \partial x^\sigma} = \frac{\partial x^\mu}{\partial x'^\kappa}\Gamma^\kappa. \tag{101}$$

The harmonic gauge can always be reached by solving this equation for $x'^\mu(x)$ when the field Γ^κ is given.

The Equivalence Principle can be rephrased in a mathematical form. Let us first consider a gauge theory where the gauge field transforms according to eq. (44). It is easy to see that the gauge transformation, $\omega(x) = e^{(x^\mu - x_0^\mu)A_\mu(x_0)}$, defined by the help of a given, fixed space-time point, x_0, yields

$$\begin{aligned} A_\mu^\omega(x) &= \omega(x)(\partial_\mu + A_\mu(x))\omega^{-1}(x) \\ &= [\mathbb{1} + (x^\mu - x_0^\mu)A_\mu(x_0)](\partial_\mu + A_\mu(x))[\mathbb{1} - (x^\mu - x_0^\mu)A_\mu(x_0)] + \mathcal{O}\left((x - x_0)^2\right) \\ &= \mathcal{O}\left(x - x_0\right). \end{aligned} \tag{102}$$

In other words, the gauge field can be eliminated at x_0 by means of a suitable gauge transformation, rendering the covariant derivative locally equivalent with the partial derivative, and leaving only its space-time derivatives non-vanishing. In case of gravity the internal and external spaces are related and we first perform linear change of coordinates in such a manner that the metric tensor assumes its Minkowski form at x_0. After that we make a further nonlinear coordinate transformation, $x \to x'$, given by

$$x^\mu - x_0^\mu = x'^\mu - x_0'^\mu - \frac{1}{2}\Gamma^\mu{}_{\nu\rho}(x_0)(x'^\nu - x_0'^\nu)(x'^\rho - x_0'^\rho) \tag{103}$$

where

$$\frac{\partial x^\kappa}{\partial x'^\mu} = \delta^\kappa_\mu - \frac{1}{2}\Gamma^\kappa{}_{\mu\rho}(x_0)(x'^\rho - x_0'^\rho) - \frac{1}{2}\Gamma^\kappa{}_{\nu\mu}(x_0)(x'^\nu - x_0'^\nu), \tag{104}$$

and in particular

$$\frac{\partial x^\kappa}{\partial x'^\mu}_{|x'=x_0'} = \delta^\kappa_\mu. \tag{105}$$

The transformation rule of the affine connection,

$$\Gamma'^\mu{}_{\nu\rho}(x') = -\frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa{}_{\nu\rho}(x_0) + \frac{\partial x^\sigma}{\partial x'^\rho}\frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa{}_{\lambda\sigma}(x)\frac{\partial x^\lambda}{\partial x'^\nu}, \tag{106}$$

gives at $x = x_0$

$$\Gamma'^\mu{}_{\nu\rho}(x_0') = -\Gamma^\mu{}_{\nu\rho}(x_0) + \Gamma^\mu{}_{\nu\rho}(x_0) = 0. \tag{107}$$

In other words, the metric tensor can be brought into its flat space-time form and the affine connection can be made vanishing at any fixed space-time point by the use of appropriate coordinates. Such an elimination of the non-trivial geometry of the space time is a local feature because the second derivatives of the metric tensor and the first derivatives of the affine connection remain non-trivial in any coordinate system.

2. Lie derivative

The Lie derivative, the covariant derivative, generated by space-time diffeomorphism, gives the change of a field $\phi(x)$ during a space-time diffeomorphism $x^\mu \to x^\mu - w^\mu(x)$ expressed in the coordinate basis at x. The field is symmetric under such a diffeomorphism if its Lie derivative is vanishing.

Let us consider for the sake of simplicity a vector field $u^\nu(x)$. Its Lie derivative with respect to the diffeomorphism $w(x)$ is the sum of two terms. The first is the change $u \to u(x+w)$, induced by the coordinate transformation, $w^\nu \partial_\nu u^\mu$. The other contribution comes from the transformation of the vector $u(x+w)$ into the basis at x. The corresponding transformation (79) contains the matrix $\omega^\mu{}_\nu = \delta^\mu_\nu - \partial_\nu w^\mu$. The Lie derivative is therefore

$$\nabla_w u^\mu = w^\nu \partial_\nu u^\mu - \partial_\nu w^\mu u^\nu. \tag{108}$$

The form

$$\nabla_w u^\mu = w^\nu D_\nu u^\mu - u^\nu D_\nu w^\mu + 2S^\mu_{\rho\nu} u^\rho w^\nu \tag{109}$$

shows that the Lie derivative of vector field is a covariant vector field. The Lie derivative of a scalar is given by the partial derivative, $\nabla_w \phi = w^\nu \partial_\nu \phi$ and the generalization of (109) for covariant vectors and tensors is straightforward, eg.

$$\begin{aligned} \nabla_w u^\mu_\kappa &= w^\nu \partial_\nu u^\mu_\kappa - \partial_\nu w^\mu u^\nu_\kappa + \partial_\kappa w^\nu u^\mu_\nu \\ &= w^\nu D_\nu u^\mu_\kappa - u^\nu_\kappa D_\nu w^\mu + u^\mu_\nu D_\kappa w^\nu + 2S^\mu_{\rho\nu} u^\rho_\kappa w^\nu + 2S^\rho_{\kappa\nu} u^\mu_\rho w^\nu. \end{aligned} \tag{110}$$

3. Field strength tensor

The $GL(4)$ field strength tensor is

$$F_{\mu\nu} = [D_\mu, D_\nu] = [\partial_\mu + \Gamma_\mu, \partial_\nu + \Gamma_\nu] = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + [\Gamma_\mu, \Gamma_\nu] \tag{111}$$

which is antisymmetric in the space-time indices,

$$F_{\mu\nu} = -F_{\nu\mu}. \tag{112}$$

The field strength tensor is called the curvature tensor and reads

$$R^\mu{}_{\nu\rho\sigma} = (F_{\rho\sigma})^\mu{}_\nu = \partial_\rho \Gamma^\mu{}_{\nu\sigma} - \partial_\sigma \Gamma^\mu{}_{\nu\rho} + \Gamma^\mu{}_{\kappa\rho} \Gamma^\kappa{}_{\nu\sigma} - \Gamma^\mu{}_{\kappa\sigma} \Gamma^\kappa{}_{\nu\rho} \tag{113}$$

with all indices shown. According to the remark, made after Eq. (61) the vanishing of the curvature tensor is equivalent with the absence of gravitational forces in a space-time without boundary.

A useful identity for the curvature tensor obtained from symmetrical connection is

$$R^{\rho}{}_{\kappa\mu\nu} + R^{\rho}{}_{\mu\nu\kappa} + R^{\rho}{}_{\nu\kappa\mu} = 0. \tag{114}$$

An important relation for the curvature tensor follows from the Bianchi identity for commutators,

$$0 = [A, [B, C]] + [B, [C, A]] + [C, [A, B]], \tag{115}$$

which yields

$$\begin{aligned} 0 &= [D_\mu, [D_\nu, D_\rho]] + [D_\nu, [D_\rho, D_\mu]] + [D_\rho, [D_\mu, D_\nu]] \\ &= [D_\mu, F_{\nu\rho}] + [D_\nu, F_{\rho\mu}] + [D_\rho, F_{\mu\nu}] \\ &= D_\mu F_{\nu\rho} + D_\nu F_{\rho\mu} + D_\rho F_{\mu\nu}, \end{aligned} \tag{116}$$

or

$$0 = D_\mu R^{\sigma}{}_{\kappa\nu\rho} + D_\nu R^{\sigma}{}_{\kappa\rho\mu} + D_\rho R^{\sigma}{}_{\kappa\mu\nu}, \tag{117}$$

by writing all indices explicitly.

The Lagrangian of a gauge field is usually a quadratic expression of the the field strength tensor, $\mathrm{tr} F_{\mu\nu} F^{\mu\nu}$. The distinguished feature of gravity is that its internal space, the tangent space of external space, is related to the space-time. This feature allows us to contract internal index with external one and to construct invariant expressions which are linear in the field strength. We may make three different contractions, the Ricci tensors, defined as

$$\begin{aligned} R_{\nu\sigma} &= R^{\rho}{}_{\nu\rho\sigma} \\ &= \partial_\rho \Gamma^{\rho}{}_{\nu\sigma} - \partial_\sigma \Gamma^{\rho}{}_{\nu\rho} + \Gamma^{\rho}{}_{\kappa\rho} \Gamma^{\kappa}{}_{\nu\sigma} - \Gamma^{\rho}{}_{\kappa\sigma} \Gamma^{\kappa}{}_{\nu\rho} \end{aligned} \tag{118}$$

$R^{\rho}{}_{\nu\sigma\rho} = -R_{\nu\sigma}$, and

$$\begin{aligned} R'_{\rho\sigma} &= R^{\mu}{}_{\mu\rho\sigma} \\ &= \partial_\rho \Gamma^{\mu}{}_{\mu\sigma} - \partial_\sigma \Gamma^{\mu}{}_{\mu\rho} + \Gamma^{\mu}{}_{\kappa\rho} \Gamma^{\kappa}{}_{\mu\sigma} - \Gamma^{\mu}{}_{\kappa\sigma} \Gamma^{\kappa}{}_{\mu\rho} \\ &= \partial_\rho \Gamma^{\mu}{}_{\mu\sigma} - \partial_\sigma \Gamma^{\mu}{}_{\mu\rho}. \end{aligned} \tag{119}$$

We have relied so far the $GL(4)$ gauge field only. The other independent field variable, the metric tensor can be used to construct the scalar curvatures, $R = g^{\mu\nu} R_{\mu\nu}$ and $R' = g^{\mu\nu} R'_{\mu\nu}$. The scalar curvature is actually trivial, $R' = 0$, because $R'_{\mu\nu}$ is antisymmetric and its contraction with the symmetric $g^{\mu\nu}$ is vanishing. Note that the curvature tensor is different in the presence of torsion and R' becomes non-trivial.

E. Metric admissibility

The metric and the affine properties are in principle independent as are the metric tensor, $g_{\mu\nu}(x)$, and the affine connection, $\Gamma^{\mu}_{\nu\rho}$. A natural relation can be established between the affine connection and the metric by following a geometric argument. This argument will be replaced by a variational equation later, when the Einstein equation is derived by the variational principle.

The geometric argument goes as follows. Let us consider two vector fields, $u^\mu(x)$ and $v^\mu(x)$, which are parallel transported along a path $\gamma(s)$,

$$\dot\gamma(s)D_\mu u = \dot\gamma(s)D_\mu v = 0. \tag{120}$$

The connection between the symmetric part of the affine connection, $\{{\rho \atop \mu\nu}\} = \frac{1}{2}(\Gamma^{\rho}_{\nu\mu} + \Gamma^{\rho}_{\mu\nu})$, called Christoffel symbol and the metric structure is achieved by imposing the condition that the scalar product of parallel transported vectors is preserved,

$$\dot\gamma(s)D_\mu u^\nu g_{\nu\rho} v^\rho = u^\nu v^\rho \dot\gamma(s) D_\mu g_{\nu\rho} = 0 \tag{121}$$

which amounts to the covariant equation,

$$Dg = 0, \tag{122}$$

the metric admissibility condition which fixes the torsion free affine connection. In order to find a more explicit form we write down this equation with all indices shown together with the relations obtained by performing cyclic permutations on the indices,

$$\begin{aligned}
D_\rho g_{\mu\nu} &= \partial_\rho g_{\mu\nu} - g_{\kappa\nu}\Gamma^{\kappa}_{\mu\rho} - g_{\mu\kappa}\Gamma^{\kappa}_{\nu\rho} \\
D_\mu g_{\nu\rho} &= \partial_\mu g_{\nu\rho} - g_{\kappa\rho}\Gamma^{\kappa}_{\nu\mu} - g_{\nu\kappa}\Gamma^{\kappa}_{\rho\mu} \\
D_\nu g_{\rho\mu} &= \partial_\nu g_{\rho\mu} - g_{\kappa\mu}\Gamma^{\kappa}_{\rho\nu} - g_{\rho\kappa}\Gamma^{\kappa}_{\mu\nu}.
\end{aligned} \tag{123}$$

By taking the sum of the last two equations minus the first one arrives at

$$\Gamma_{\rho\mu\nu} + \Gamma_{\rho\nu\mu} = \partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu} \tag{124}$$

with $\Gamma_{\rho\mu\nu} = g_{\rho\kappa}\Gamma^{\kappa}_{\mu\nu}$ and the relation

$$\{{\rho \atop \mu\nu}\} = \frac{1}{2} g^{\rho\sigma}(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}). \tag{125}$$

The antisymmetric part of the affine connection, the torsion tensor is left free after imposing the metric admissibility condition.

The curvature tensor, given in terms of the metric, occurs frequently in applications. We have

$$\begin{aligned} R_{\rho\kappa\nu\mu} &= g_{\rho\lambda}\left[\Gamma^{\lambda}_{\kappa\mu,\nu} + \Gamma^{\lambda}_{\sigma\nu}\Gamma^{\sigma}_{\kappa\mu} - (\mu \Longleftrightarrow \nu)\right] \\ &= \frac{1}{2}g_{\rho\lambda}\left\{\left[g^{\lambda\sigma}(g_{\sigma\kappa,\mu} + g_{\mu\sigma,\kappa} - g_{\kappa\mu,\sigma})\right]_{,\nu} + \Gamma^{\lambda}_{\sigma\nu}\Gamma^{\sigma}_{\kappa\mu} - (\mu \Longleftrightarrow \nu)\right\} \\ &= \frac{1}{2}g_{\rho\lambda}g^{\lambda\sigma}_{,\nu}(g_{\sigma\kappa,\mu} + g_{\mu\sigma,\kappa} - g_{\kappa\mu,\sigma}) + \frac{1}{2}(g_{\mu\rho,\kappa\nu} - g_{\kappa\mu,\rho\nu}) + g_{\rho\lambda}\Gamma^{\lambda}_{\sigma\nu}\Gamma^{\sigma}_{\kappa\mu} - (\mu \Longleftrightarrow \nu) \end{aligned} \tag{126}$$

where the notation $f_{,\mu} = \partial_\mu f$ is used.The relation

$$g_{\rho\lambda}g^{\lambda\sigma}_{,\nu} = -g^{\lambda\sigma}g_{\rho\lambda,\nu} = -g^{\lambda\sigma}(\Gamma_{\nu\rho\lambda} + \Gamma_{\nu\lambda\rho}) \tag{127}$$

allows us to write

$$\begin{aligned} R_{\rho\kappa\nu\mu} &= -\frac{1}{2}g^{\lambda\sigma}(\Gamma_{\nu\rho\lambda} + \Gamma_{\nu\lambda\rho})(g_{\sigma\kappa,\mu} + g_{\mu\sigma,\kappa} - g_{\kappa\mu,\sigma}) + \frac{1}{2}(g_{\mu\rho,\kappa\nu} - g_{\kappa\mu,\rho\nu}) + g_{\rho\lambda}\Gamma^{\lambda}_{\sigma\nu}\Gamma^{\sigma}_{\kappa\mu} \\ &\quad -(\mu \Longleftrightarrow \nu) \\ &= -\Gamma^{\lambda}_{\kappa\mu}(\Gamma_{\nu\rho\lambda} + \Gamma_{\nu\lambda\rho}) + \frac{1}{2}(g_{\mu\rho,\kappa\nu} - g_{\kappa\mu,\rho\nu}) + g_{\rho\lambda}\Gamma^{\lambda}_{\sigma\nu}\Gamma^{\sigma}_{\kappa\mu} - (\mu \Longleftrightarrow \nu) \\ &= \frac{1}{2}(g_{\mu\rho,\kappa\nu} - g_{\nu\rho,\kappa\mu} - g_{\kappa\mu,\rho\nu} + g_{\kappa\nu,\rho\mu}) - g_{\sigma\lambda}\Gamma^{\lambda}_{\kappa\mu}\Gamma^{\sigma}_{\rho\nu} - g_{\sigma\rho}\Gamma^{\lambda}_{\kappa\mu}\Gamma^{\sigma}_{\lambda\nu} + g_{\rho\lambda}\Gamma^{\lambda}_{\sigma\nu}\Gamma^{\sigma}_{\kappa\mu} \\ &\quad +g_{\sigma\lambda}\Gamma^{\lambda}_{\kappa\nu}\Gamma^{\sigma}_{\rho\mu} + g_{\sigma\rho}\Gamma^{\lambda}_{\kappa\nu}\Gamma^{\sigma}_{\lambda\mu} - g_{\rho\lambda}\Gamma^{\lambda}_{\sigma\mu}\Gamma^{\sigma}_{\kappa\nu} \\ &= \frac{1}{2}(g_{\mu\rho,\kappa\nu} - g_{\nu\rho,\kappa\mu} - g_{\kappa\mu,\rho\nu} + g_{\kappa\nu,\rho\mu}) - g_{\sigma\lambda}\Gamma^{\lambda}_{\kappa\mu}\Gamma^{\sigma}_{\rho\nu} + g_{\sigma\lambda}\Gamma^{\lambda}_{\kappa\nu}\Gamma^{\sigma}_{\rho\mu}. \end{aligned} \tag{128}$$

Let us consider a two-dimensional sphere as a simple example. The invariant length $ds^2 = r^2(d\theta^2 + \sin^2\theta d\phi^2)$ gives the metric tensor

$$g_{\mu\nu} = r^2\begin{pmatrix}1 & 0\\ 0 & \sin^2\theta\end{pmatrix}, \quad g^{\mu\nu} = \frac{1}{r^2}\begin{pmatrix}1 & 0\\ 0 & \frac{1}{\sin^2\theta}\end{pmatrix}. \tag{129}$$

The non-vanishing matrix elements of the Christoffel symbol, $\left\{{\theta \atop \phi\phi}\right\} = -\sin\theta\cos\theta$, $\left\{{\phi \atop \theta\phi}\right\} = \left\{{\phi \atop \phi\theta}\right\} = -\cot\theta$, give rise

$$\begin{aligned} R^{\theta}_{\phi\theta\phi} &= \partial_\theta\left\{{\theta \atop \phi\phi}\right\} - \partial_\phi\left\{{\theta \atop \theta\phi}\right\} + \left\{{\theta \atop \theta\mu}\right\}\left\{{\mu \atop \phi\phi}\right\} - \left\{{\theta \atop \phi\mu}\right\}\left\{{\mu \atop \theta\phi}\right\} \\ &= -\partial_\theta\sin\theta\cos\theta + \sin\theta\cos\theta\cot\theta = \sin^2\theta \end{aligned} \tag{130}$$

and $R^{\phi}_{\theta\theta\phi} = g^{\phi\phi}R_{\phi\theta\theta\phi} = -g^{\phi\phi}R_{\theta\phi\theta\phi} = -g^{\phi\phi}g_{\theta\theta}R^{\theta}_{\phi\theta\phi} = -1$. The Ricci tensor is diagonal, $R_{\theta\theta} = 1$, $R_{\phi\phi} = \sin^2\theta$ and the scalar curvature is $R = \frac{2}{r^2}$.

The symmetries of the curvature tensor in addition to (112) and (114) in the metric admissible case are

$$R_{\rho\kappa\mu\nu} = -R_{\kappa\rho\mu\nu} = R_{\mu\nu\rho\kappa} \tag{131}$$

and the number of independent components is $256 \to 20$. We note that the curvature is vanishing for flat space only.

The contraction of two indices in the Bianchi-identity (117) gives

$$0 = D_\mu R_{\kappa\rho} + D_\nu R^\nu{}_{\kappa\rho\mu} - D_\rho R_{\kappa\mu}. \tag{132}$$

A further contraction of indices κ and μ by means of the metric tensor gives for the metric admissible curvature tensor

$$0 = 2D_\mu R^\mu{}_\rho - D_\rho R \tag{133}$$

which expresses the covariant conservation law for the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \tag{134}$$

as

$$D_\mu G^\mu{}_\nu = 0. \tag{135}$$

Einstein's original argument to establish the equation of motion for the space time geometry was based on a trial and error method of finding a covariant equation whose source term is the energy-momentum tensor,

$$X^{\mu\nu} = T^{\mu\nu}. \tag{136}$$

Due to the conservation law, $D_\mu T^{\mu\nu} = 0$, we have $D_\mu X^{\mu\nu} = 0$. Since $X^{\mu\nu}$ is expected to be linear in the curvature tensor the choice $X = \kappa G$, κ being a constant, is a natural one.

The metric admissibility simplifies the condition, expressing the invariance of the metric tensor under space-time diffeomorphism. In fact, the vanishing of the Lie derivative of the metric tensor,

$$\begin{aligned} 0 = \nabla_w g_{\mu\nu} &= w^\kappa \partial_\kappa g_{\mu\nu} + \partial_\mu w^\kappa g_{\kappa\nu} + \partial_\nu w^\kappa g_{\mu\kappa} \\ &= D_\mu w_\nu + D_\nu w_\mu + 2S^\rho_{\nu\kappa} g_{\rho\mu} w^\kappa \end{aligned} \tag{137}$$

and the diffeomorphism $w^\mu(x)$, satisfying this condition for a given metric tensor is called Killing field, representing a symmetry of the metric in question.

F. Invariant integral

Two problems still have to addressed before embarking the variational equations of General Relativity. One is the integral measure in the action,

$$S = \int d^4 x L. \tag{138}$$

In the case of relativistic field theories in flat space-time the Lorentz invariance of the action is realized by using Lorentz invariant Lagrangian, L, and integration measure, d^4x. The gauge invariance of the action is assured in a similar manner, by the gauge invariance of the action (non-trivial) and the integral measure (trivial). The gauge invariance in General Relativity is the invariance under (non-linear) change of coordinates. The integral measure d^4x is obviously non-invariant. Hence we have to find an integral measure which remains invariant under the change of coordinates. The other problem to settle is presented by the modified form of the continuity equation in gauge theory. The usual continuity equation in flat space-time, $\partial_\mu j^\mu = 0$ is replaced by the covariant equation, $D_\mu j^\mu = 0$ in General Relativity. The affine connection, appearing in the covariant derivative, prevent us to arrive at the balance equation, expressing the change of the charge, enclosed in a given volume as a surface integral on the boundary of the volume. We shall see that this latter problem disappears when the invariant integral measure is used in the space-time.

The metric tensor transforms as

$$g_{\mu\nu} = \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} g'_{\rho\sigma} \tag{139}$$

during the coordinate change $x^\mu \to x'^\mu$ and its determinant, $g = \det g_{\mu\nu}$, changes as

$$g = g' \left(\det \frac{\partial x'}{\partial x} \right)^2 \tag{140}$$

showing that the integral measure

$$d_{\rm inv} x = dx\sqrt{-g} \to dx' \det \frac{\partial x}{\partial x'} \sqrt{-g'} \det \frac{\partial x'}{\partial x} \tag{141}$$

stays invariant. Thus one speaks of scalar, vector and tensor densities constructed by means of the determinant of the metric tensor as $S\sqrt{-g}$, $v^\mu\sqrt{-g}$, $T^{\mu\nu}\sqrt{-g}$, etc.

A useful and simple expression can be obtained for the divergence of a vector field in the absence of torsion by using the equation

$$\delta g = \frac{\partial g}{\partial g_{\sigma\mu}} \delta g_{\sigma\mu} = g g^{\sigma\mu} \delta g_{\sigma\mu} \tag{142}$$

which holds because $gg^{\sigma\mu}$ is actually the minor corresponding to the matrix element $g_{\sigma\mu}$. (The minor M_A of an $n \times n$ matrix A,

$$(M_A)_{j,k} = (-1)^{j+k} d_{k,j}, \tag{143}$$

is defined in terms of the determinant $d_{j,k}$ of the $n-1 \times n-1$ matrix, obtained by omitting the j-th row and the k-th column of A. The determinant of A, expanded along the k-th row of A can be written as

$$\det[A] = \sum_j A_{k,j}(-1)^{j+k} d_{k,j}. \tag{144}$$

The equation for $\ell \neq j$

$$0 = \sum_j A_{k,j}(-1)^{j+k} d_{\ell,j} \tag{145}$$

expresses the vanishing of the determinant of a matrix whose j-th and ℓ-th rows are identical (When the determinant is expanded along the ℓ-th row then the matrix element $A_{\ell,j}$ enters only as the coefficient of the sub-determinant $d_{\ell,j}$. Thus this equation can be interpreted as the determinant of a matrix whose ℓ-th row is $A_{k,j}$, the same as the k-th row.) Eqs. (144) and (145) can be summarized by the expression

$$A^{-1} = \frac{M_A}{\det[A]} \tag{146}$$

for the inverse matrix. Finally, eq. (142) follows from

$$\frac{\partial \det[A]}{\partial A_{k,j}} = (-1)^{j+k} d_{k,j} = (M_A)_{j,k} \tag{147}$$

and (146) for the symmetric metric tensor.)

Eq. (142) gives $\partial_\nu g = g g^{\sigma\mu} \partial_\nu g_{\sigma\mu}$ allows us to write

$$\Gamma^\mu_{\nu\mu} = \frac{1}{2} g^{\sigma\mu}(\partial_\nu g_{\sigma\mu} + \partial_\mu g_{\nu\sigma} - \partial_\sigma g_{\nu\mu}) = \frac{1}{2} g^{\sigma\mu} \partial_\nu g_{\sigma\mu} = \frac{\partial_\nu g}{2g} = \frac{\partial_\nu \sqrt{-g}}{\sqrt{-g}}. \tag{148}$$

The Gauss' theorem can finally be obtained for the divergence

$$D_\mu v^\mu = \partial_\mu v^\mu + \Gamma^\mu_{\nu\mu} v^\nu = \partial_\mu v^\mu + \frac{\partial_\mu \sqrt{-g}}{\sqrt{-g}} v^\mu = \frac{1}{\sqrt{-g}} \partial_\mu(\sqrt{-g} v^\mu) \tag{149}$$

as

$$\int dx \sqrt{-g} D_\mu v^\mu = \int dx \partial_\mu(\sqrt{-g} v^\mu) = \int ds_\mu \sqrt{-g} v^\mu. \tag{150}$$

A particularly useful application of this relation is for covariantly conserved currents, $D_\mu j^\mu = 0$ which yield ordinarily conserved current density, $\partial_\mu(\sqrt{-g} j^\mu) = 0$ and conserved charge $Q = \int d^3x \sqrt{-g} j^0$.

Metric admissibility renders the definition of the D'Alambertian unique. In fact, we find

$$D_\mu D^\mu = g^{\mu\nu} D_\mu D_\nu = D_\mu g^{\mu\nu} D_\nu = D_\mu D_\nu g^{\mu\nu}, \tag{151}$$

and its action on a scalar is particularly simple,

$$D_\mu D^\mu \phi = D_\mu \partial^\mu \phi = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} \partial^\mu \phi) \tag{152}$$

As an example consider a d-dimensional Euclidean space, parametrized by polar coordinates,

$$x^j = \begin{pmatrix} r \\ r\cos\phi^1 \\ r\sin\phi^1\cos\phi^2 \\ r\sin\phi^1\sin\phi^2\cos\phi^3 \\ \vdots \\ r\sin\phi^1\cdots\sin\phi^{n-2}\cos\phi^{n-1} \\ r\sin\phi^1\cdots\sin\phi^{n-2}\sin\phi^{n-1} \end{pmatrix}, \tag{153}$$

$0 \le \phi^j \le \pi$, $j = 1, \ldots, d-2$, $0 \le \phi^{d-1} \le 2\pi$ where the metric tensor is of the form

$$g_{jk} = r^2 \begin{pmatrix} 1 & 0 \\ 0 & g_{S^{d-1}} \end{pmatrix}. \tag{154}$$

and

$$D^2 = \frac{1}{r^{d-1}} \partial_r r^{d-1} \partial_r + \frac{1}{r^2} D^2_{S^{d-1}}. \tag{155}$$

Another example, the Laplace operator on the two-sphere with metric (129) is

$$\Delta_{S^2} = \frac{1}{r^2} \left[\frac{1}{\sin\theta} \partial_\theta \sin\theta \partial_\theta + \frac{1}{\sin^2\theta} \partial^2_\phi \right]. \tag{156}$$

The relation between the affine connection and the metric tensor, the metric admissibility, was imposed as a reasonable assumption. We shall see below that General Relativity supports this relation and the metric admissibility will be the result of the equation of motion for the affine connection.

G. Dynamics

The equations of motion of Einstein's general relativity can be obtained as the Euler-Lagrange equations of the Einstein action

$$S_E = -\frac{1}{16\pi G} \int dx \sqrt{-g} (R + 2\Lambda) = -\frac{1}{16\pi G} \int dx \sqrt{-g} (g^{\mu\nu} R_{\mu\nu} + 2\Lambda) \tag{157}$$

with G and Λ being the gravitational and the cosmological constants. The action is considered as the functional of the independent field variables g and Γ. The metric admissibility will be derived rather than assumed.

The affine connection is not a tensor and it will be advantageous to use tensor fields as independent variables. Thus we separate the Christoffel symbol from the affine connection by writing

$$\Gamma^{\rho}_{\mu\nu} = \left\{{\rho \atop \mu\nu}\right\} + C^{\rho}_{\mu\nu} \tag{158}$$

and consider $C^{\rho}_{\mu\nu}$, a tensor, rather than the whole $\Gamma^{\rho}_{\mu\nu}$ as the independent variable controlling the affine structure.

Notice the following slight complication about the choice of the independent components of the metric tensor used for the variational procedure. The variation of the relation $\delta^{\mu}_{\rho} = g^{\mu\nu}g_{\nu\rho}$ gives

$$\begin{aligned} 0 &= \delta g^{\mu\nu} g_{\nu\rho} + g^{\mu\nu}\delta g_{\nu\rho} \\ \delta g_{\mu\nu} &= -g_{\mu\rho}g_{\nu\sigma}\delta g^{\rho\sigma} \end{aligned} \tag{159}$$

thus the metric can not be used in this calculation to change the position of the indices. Instead, we keep $g^{\mu\nu}$ as the independent tensor field for the metric.

A notation which will serve us is the metric admissible covariant derivative,

$$\tilde{D}_{\mu}v^{\nu} = \partial_{\mu}v^{\nu} + \left\{{\nu \atop \rho\mu}\right\} v^{\rho} \tag{160}$$

which is independent of the tensor field $C^{\rho}_{\mu\nu}$.

We can now write the $GL(4)$ field strength tensor, the Riemann curvature tensor, expressed in term of the 'background field' covariant derivative $\tilde{D}$ as

$$F_{\mu\nu} = [D_{\mu}, D_{\nu}] = [\tilde{D}_{\mu} + C_{\mu}, \tilde{D} + C_{\nu}]. \tag{161}$$

The variation when $C \to C + \delta C$ generates

$$\begin{aligned} \delta F_{\mu\nu} &= [D_{\mu} + \delta C_{\mu}, D_{\nu} + \delta C_{\nu}] - [D_{\mu}, D_{\nu}] \\ &= (D_{\mu} + \delta C_{\mu})(D_{\nu} + \delta C_{\nu}) - (D_{\nu} + \delta C_{\nu})(D_{\mu} + \delta C_{\mu}) - D_{\mu}D_{\nu} + D_{\nu}D_{\mu} \\ &= D_{\mu}\delta C_{\nu} + \delta C_{\mu}D_{\nu} - D_{\nu}\delta C_{\mu} - \delta C_{\nu}D_{\mu} + \mathcal{O}\left(\delta C^2\right) \\ &= (D_{\mu}\delta C_{\nu}) - (D_{\nu}\delta C_{\mu}) + \mathcal{O}\left(\delta C^2\right) \end{aligned} \tag{162}$$

which gives

$$\delta R^{\mu}{}_{\nu\rho\sigma} = D_{\rho}\delta C^{\mu}{}_{\nu\sigma} - D_{\sigma}\delta C^{\mu}{}_{\nu\rho} \tag{163}$$

when the indices are written explicitly and

$$\delta R_{\nu\sigma} = D_\rho \delta C^\rho{}_{\nu\sigma} - D_\sigma \delta C^\rho{}_{\nu\rho}. \tag{164}$$

The variation of the curvature scalar, $g^{\nu\sigma}\delta R_{\nu\sigma}$ is a sum of terms which are proportional to δC or $\partial\delta C$. It is a scalar therefore the terms proportional to $\partial\delta C$ can be extended the replacement $\partial\delta C \to \tilde{D}\delta C$ by the expense of modifying the terms proportional to δC. The result is the expression

$$g^{\nu\sigma}\delta R_{\nu\sigma} = \tilde{D}_\rho v^\rho + \delta C^\kappa{}_{\nu\rho} K_\kappa{}^{\nu\rho}, \tag{165}$$

K being linear in C. The actual calculation

$$\begin{aligned}
g^{\nu\sigma}\delta R_{\nu\sigma} &= g^{\nu\sigma}(\tilde{D}_\rho \delta C^\rho{}_{\nu\sigma} - \tilde{D}_\sigma \delta C^\rho{}_{\nu\rho}) + C^\rho_{\kappa\rho}\delta C^\kappa{}_\nu{}^\nu - \delta C^\rho{}_\kappa{}^\nu C^\kappa_{\nu\rho} - \delta C^\rho{}_{\nu\kappa} C^{\kappa\nu}{}_\rho - C^\rho{}_\kappa{}^\nu \delta C^\kappa{}_{\nu\rho} + \delta C^\rho{}_{\kappa\rho} C^\kappa{}_\nu{}^\nu + \delta C^\rho{}_{\nu\kappa} C^\kappa{}_\rho{}^\nu \\
&= \tilde{D}_\rho v^\rho + C^\rho_{\kappa\rho}\delta C^\kappa{}_\nu{}^\nu + \delta C^\rho{}_{\kappa\rho} C^\kappa{}_\nu{}^\nu - \delta C^\rho{}_\kappa{}^\nu C^\kappa_{\nu\rho} - \delta C^\rho{}_{\nu\kappa} C^{\kappa\nu}{}_\rho - C^\rho{}_\kappa{}^\nu \delta C^\kappa{}_{\nu\rho} + \delta C^\rho{}_{\nu\kappa} C^\kappa{}_\rho{}^\nu \\
&= \tilde{D}_\rho v^\rho + C^\rho_{\kappa\rho}\delta C^\kappa{}_\nu{}^\nu + \delta C^\rho{}_{\kappa\rho} C^\kappa{}_\nu{}^\nu - \delta C^\kappa{}_\nu{}^\rho C^\nu_{\rho\kappa} - \delta C^\kappa{}_{\nu\rho} C^{\rho\nu}{}_\kappa - C^\rho{}_\kappa{}^\nu \delta C^\kappa{}_{\nu\rho} + \delta C^\kappa{}_{\nu\rho} C^\rho{}_\kappa{}^\nu \\
&= \tilde{D}_\rho v^\rho - \tilde{D}_\sigma (g^{\nu\sigma}\delta C^\rho{}_{\nu\rho}) + \delta C^\kappa{}_\nu{}^\nu C^\rho_{\kappa\rho} + \delta C^\rho{}_{\kappa\rho} C^\kappa{}_\nu{}^\nu - \delta C^\kappa{}_{\nu\rho}(C^{\nu\rho}{}_\kappa - C^{\rho\nu}{}_\kappa - C^\rho{}_\kappa{}^\nu + C^\rho{}_\kappa{}^\nu) \\
&= \tilde{D}_\rho v^\rho + \delta C^\kappa{}_{\nu\rho}(g^{\nu\rho} C^\rho_{\kappa\rho} + g^{\kappa\rho} C^\nu{}_\lambda{}^\lambda - C^{\nu\rho}{}_\kappa + C^{\rho\nu}{}_\kappa)
\end{aligned} \tag{166}$$

yields

$$v^\rho = g^{\nu\sigma}\delta C^\rho{}_{\nu\sigma} - g^{\nu\rho}\delta C^\sigma{}_{\nu\sigma}, \tag{167}$$

and

$$K_\kappa{}^{\nu\rho} = g^{\nu\rho} C^\rho_{\kappa\rho} + g^{\kappa\rho} C^\nu{}_\lambda{}^\lambda - C^{\nu\rho}{}_\kappa + C^{\rho\nu}{}_\kappa. \tag{168}$$

After these preparation one can easily calculate the variation of the integrand in the action

$$\delta[\sqrt{-g}(R_{\mu\nu}g^{\mu\nu} + 2\Lambda)] = \delta\sqrt{-g}(R_{\mu\nu}g^{\mu\nu} + 2\Lambda) + \sqrt{-g}\delta R_{\mu\nu}g^{\mu\nu} + \sqrt{-g}R_{\mu\nu}\delta g^{\mu\nu}. \tag{169}$$

The first term contains

$$\begin{aligned}
\delta\sqrt{-g} &= -\frac{g}{2\sqrt{-g}} g^{\mu\nu}\delta g_{\mu\nu} \\
&= -\frac{1}{2}\sqrt{-g} g^{\mu\nu} g_{\mu\rho} g_{\nu\sigma} \delta g^{\rho\sigma} \\
&= -\frac{1}{2}\sqrt{-g} g_{\nu\sigma}\delta g^{\nu\sigma},
\end{aligned} \tag{170}$$

the second is given by Eq. (165) thus we have

$$\begin{aligned}
\delta[\sqrt{-g}(R_{\mu\nu}g^{\mu\nu} + 2\Lambda)] &= -\frac{1}{2}\sqrt{-g} g_{\nu\sigma}\delta g^{\nu\sigma}(R + 2\Lambda) + \sqrt{-g}\tilde{D}_\rho v^\rho + \sqrt{-g}\delta C^\kappa{}_{\nu\rho} K_\kappa{}^{\nu\rho} + \sqrt{-g} R_{\mu\nu}\delta g^{\mu\nu} \\
&= \sqrt{-g}\left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - \Lambda g_{\mu\nu}\right)\delta g^{\mu\nu} + \sqrt{-g}\delta C^\kappa{}_{\nu\rho} K_\kappa{}^{\nu\rho} + \sqrt{-g}\tilde{D}_\rho v^\rho
\end{aligned} \tag{171}$$

The last contribution can be ignored being a surface term which is ignored from the point of view of the equations of motion. The variation of the affine connection yields $K = 0$ which gives $C = 0$, the metric admissibility condition for the affine connection. Finally, the variation of the metric tensor leads to the Einstein equation

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = 0 \tag{172}$$

in the absence of matter.

IV. COUPLING TO MATTER

After having obtained the Lagrangian and the equations of motion for the gravitation field in the absence of matter let us turn to the question of introducing matter in our description.

A. Point particle in an external gravitational field

First we consider the simplest problem of the motion of a massive point particle in a fixed gravitational field, ie. in a fixed geometry where the trajectory, identified by the equation of motion is the generalization of the straight line of the Minkowskian, flat space-time.

1. Equivalence Principle

The simplest way to find the equation of motion for a point particle is to use the Equivalence principle. A free particle follows the straight trajectory $\xi^a(s)$ satisfying the equation of motion

$$\frac{d\dot{\xi}^a(s)}{ds} = 0 \tag{173}$$

with $\dot{\xi} = \frac{d\xi}{ds}$ in flat space-time, in the absence of gravity. When an external gravitational field is introduced then the trajectory $x^\mu(s)$ is not a straight line anymore but the Equivalence Principle allows us to recover the same equation of motion locally, at a given space-time point, by an appropriate choice of the coordinate system. Eq. (173) shows that the four velocity, assumed as a vector field, $u^\mu(x) = \dot{x}^\mu$, filling up a region of the space-time, remains unchanged on the world line. The unique covariant extension of such a parallel transport on a non-flat geometry is

$$u^\nu D_\nu u^\mu = \dot{u}^\mu + u^\rho \Gamma^\mu{}_{\rho\nu} u^\nu = 0. \tag{174}$$

Let us now assume that an external, non-gravitational force acts on the particle and Eq. (174) is replaced by

$$\dot{u}^\mu + u^\rho \Gamma^\mu{}_{\rho\nu} u^\nu = \frac{F^\mu}{mc}. \tag{175}$$

The form

$$mc\dot{u}^\mu = F^\mu + F^\mu_{gr} \tag{176}$$

of this equation with

$$F^\mu_{gr} = -mu^\mu \Gamma^\mu{}_{\rho\nu} u^\nu \tag{177}$$

shows that the gravitational field generates a force F_{gr} which is linear in the velocities in a manner similar to the Lorentz force of electrodynamics where

$$\begin{aligned} mc\ddot{x}^\mu &= F^\mu + F^\mu_{ed} \\ F^\mu_{ed} &= \frac{e}{c} F^\mu{}_\nu u^\nu, \end{aligned} \tag{178}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor.

2. Spin precession

It is instructive to generalize this argument for a gyroscope, a point particle with an angular momentum, spin, given by the four-vector $S^\mu = (0, \mathbf{S})$ in the rest-frame. In the absence of external forces we have time independent spin,

$$\frac{dS^\mu}{ds} = 0. \tag{179}$$

The covariant generalization for an external gravitational field of the form $S^\mu = (0, \mathbf{S})$ of the spin vector is the equation of motion

$$\dot{S}^\mu + \Gamma^\mu{}_{\rho\nu} S^\rho \dot{x}^\nu = 0, \tag{180}$$

together with the auxiliary condition

$$S_\mu \dot{x}^\mu = 0, \tag{181}$$

expressing the structure $S^\mu = (0, \mathbf{S})$, found in the rest-frame, in a covariant manner.

We furthermore assume that the external force F_{ext} does not exert torque on the system. The spin will still be conserved, $d\mathbf{S}/dt = 0$, in the co-moving frame which can be reproduced in a covariant form by requiring that the vector $\dot{S}$ be proportional to $\dot{x}$,

$$\dot{S} = a\dot{x}. \tag{182}$$

The covariant derivative of the orthogonality condition (181) along the world line, $0 = a\dot{x}_\mu\dot{x}^\mu + S_\mu\ddot{x}^\mu$, yields

$$a = -S_\mu\ddot{x}^\mu = -S_\mu\frac{F^\mu}{mc}. \tag{183}$$

The covariant generalization of the equation of motion (182),

$$\dot{S}^\mu = -S_\nu\frac{F^\nu}{mc}\dot{x}^\mu, \tag{184}$$

is the Fermi-Walker transport of the spin which reduces to parallel transport in the absence of external force, $F = 0$.

3. *Variational equation of motion*

A direct derivation of the trajectory of a point particle, without referring to the Equivalence Principle starts with the action of a free massive point particle,

$$S = -mc\int\sqrt{\dot{x}^\mu g_{\mu\nu}(x)\dot{x}^\nu}d\tau \tag{185}$$

where $\dot{x}^\mu = dx^\mu(\tau)/d\tau$. The corresponding Euler-Lagrange equation,

$$\frac{\partial L}{\partial x^\rho} - \frac{d}{d\tau}\frac{\partial L}{\partial\dot{x}^\rho} = 0 \tag{186}$$

containing the terms

$$\begin{aligned}\frac{\partial L}{\partial x^\rho} &= -mc\frac{\dot{x}^\mu\partial_\rho g_{\mu\nu}\dot{x}^\nu}{2\sqrt{\dot{x}^\mu g_{\mu\nu}\dot{x}^\nu}},\\ \frac{\partial L}{\partial\frac{dx^\rho}{d\tau}} &= -mc\frac{g_{\rho\nu}\dot{x}^\nu}{\sqrt{\dot{x}^\mu g_{\mu\nu}\dot{x}^\nu}},\end{aligned} \tag{187}$$

reads

$$\begin{aligned}0 &= -\frac{\dot{x}^\mu\partial_\rho g_{\mu\nu}\dot{x}^\nu}{2\sqrt{\dot{x}^\mu g_{\mu\nu}\dot{x}^\nu}} + \frac{d}{d\tau}\frac{g_{\rho\nu}\dot{x}^\nu}{\sqrt{\dot{x}^\mu g_{\mu\nu}\dot{x}^\nu}}\\ &= \frac{1}{\sqrt{\dot{x}^\mu g_{\mu\nu}\dot{x}^\nu}}\left[-\frac{1}{2}\dot{x}^\mu\partial_\rho g_{\mu\nu}\dot{x}^\nu + \dot{x}^\kappa\partial_\kappa g_{\rho\nu}\dot{x}^\nu + g_{\rho\nu}\ddot{x}^\nu + g_{\rho\nu}\dot{x}^\nu\frac{d}{d\tau}\frac{1}{\sqrt{\dot{x}^\mu g_{\mu\lambda}\dot{x}^\lambda}}\right]\end{aligned} \tag{188}$$

We symmetrize in the indices κ and ν the factor $\partial_\kappa g_{\rho\nu}$ in the second term of the last line and find

$$\begin{aligned}0 &= -\frac{1}{2}\dot{x}^\mu\partial_\rho g_{\mu\nu}\dot{x}^\nu + \frac{1}{2}\dot{x}^\kappa(\partial_\kappa g_{\rho\nu} + \partial_\nu g_{\rho\kappa})\dot{x}^\nu + g_{\rho\nu}\ddot{x}^\nu + g_{\rho\nu}\frac{dx^\nu}{d\tau}\frac{d}{d\tau}\frac{1}{\sqrt{\dot{x}^\mu g_{\mu\lambda}\dot{x}^\lambda}}\\ &= g_{\rho\sigma}(\ddot{x}^\sigma + \Gamma^\sigma_{\nu\kappa}\dot{x}^\nu\dot{x}^\kappa + \dot{x}^\sigma\dot{f})\end{aligned} \tag{189}$$

with $f(\tau) = 1/\sqrt{\dot{x}^\mu g_{\mu\lambda}\dot{x}^\lambda}$. This equation assumes the simplest form (174) when $f(\tau)$ is constant, ie. $\tau = s$, the invariant length of the world line and its solution is called geodesic. For $\tau \neq s$ the term, proportional to $\dot{f}$ adjusts the length of the four-velocity without changing its direction.

4. Geodesic deviation

The dynamics of an infinitesimal deviation from a solution of the Newton equation of a particle moving in a given potential is that of a harmonic oscillator with time dependent frequency. In fact, let us consider a trajectory $\xi(t)$ which obeys the equation of motion, $m\ddot{\xi}^j = -\nabla^j U$ and consider a neighboring trajectory $\xi(t) + \delta\xi(t)$. The infinitesimal shift $\delta\mathbf{x}$ satisfies the linear, time dependent equation $m\delta\ddot{\xi}^j = -\delta\xi^k\nabla^k\nabla^j U$.

The generalization of the equation of motion of infinitesimal deviation $\delta\xi^\mu$ from a geodesic consists of a straightforward linearization of (174) in the deformation. But the result is easier to find by embedding our world line and its deformed partner into a family of world lines, solutions of Eq. (174) which fill up the space-time in the vicinity of our observation point and using a curve $\gamma^\mu(\tau)$ which crosses our world line and its deformed partner. The integral curves of the four-velocity vector field, $u(x) = \frac{d}{ds}\xi(x) = \dot{\xi}(x)$, are the world lines, hence $\frac{d}{ds}\phi = u^\nu D_\nu\phi = \dot{\phi}$ holds for any field $\phi(x)$.

Let us consider the surface in the space-time which is swept through by the world lines which cross γ and use the coordinates s and τ to identify its points $\xi^\mu(s,\tau)$. The coordinate basis vector fields, $u = \partial_s\xi(s,\tau)$, $v = \partial_\tau\xi(s,\tau)$ are holonomic, $\partial_s v = \partial_\tau u$ according to Eq. (84), and we have

$$u^\nu D_\nu v^\mu = v^\nu D_\nu u^\mu. \tag{190}$$

Let us calculate finally the acceleration of the deformation $\delta\xi = \epsilon\partial_\tau\xi = \epsilon v$,

$$\ddot{v} = u^\mu D_\mu(u^\nu D_\nu v) \tag{191}$$

Two successive applications of the holonomy condition give

$$\begin{aligned}
\ddot{v} &= u^\mu D_\mu(v^\nu D_\nu u) \\
&= u^\mu(D_\mu v^\nu)D_\nu u + u^\mu v^\nu D_\mu D_\nu u \\
&= v^\mu(D_\mu u^\nu)D_\nu u + u^\mu v^\nu[D_\mu, D_\nu]u + u^\mu v^\nu D_\nu D_\mu u \\
&= v^\mu(D_\mu u^\nu)D_\nu u + u^\mu v^\nu[D_\mu, D_\nu]u + v^\nu D_\nu(u^\mu D_\mu u) - v^\nu(D_\nu u^\mu)D_\mu u \\
&= u^\mu v^\nu[D_\mu, D_\nu]u + v^\nu D_\nu(u^\mu D_\mu u)
\end{aligned} \tag{192}$$

The four-acceleration of the world lines is vanishing, therefore

$$\ddot{v}^\rho = R^\rho{}_{\kappa\mu\nu}u^\kappa u^\mu v^\nu, \tag{193}$$

the acceleration satisfies a linear equation whose coefficient matrix depends is a quadratic function of the four-velocity.

It is instructive to see what happens in electrodynamics where we may start with the the equation of motion, (178), written for the deformed world line $x + \delta x$ is

$$mc(\ddot{x}^\mu + \delta\ddot{x}^\mu) = \frac{e}{c}F^\mu{}_\nu(x + \delta x)(\dot{x}^\nu + \delta\dot{x}^\nu). \tag{194}$$

The linearization in the deformation δx yields immediately

$$mc\delta\ddot{x}^\mu = \frac{e}{c}\delta x^\rho \partial_\rho F^\mu{}_\nu \dot{x}^\nu + \frac{e}{c}F^\mu{}_\nu \delta\dot{x}^\nu. \tag{195}$$

This equation can be obtained from Eq. (192) by first making the replacement $D_\mu \to \partial_\mu$,

$$mc\ddot{v} = v^\nu \partial_\nu \dot{u}, \tag{196}$$

followed by the use of the equation of motion, Eq. (178),

$$mc\ddot{v}^\mu = \frac{e}{c}v^\nu \partial_\nu F^\mu{}_\rho u^\rho + \frac{e}{c}F^\mu{}_\rho v^\nu \partial_\nu u^\rho. \tag{197}$$

Eq. (195) follows by noting $v^\nu \partial_\nu u^\rho = \delta x^\nu \partial_\nu \dot{x}^\rho = \delta\tau \partial_\tau \partial_s x^\rho = \partial_s(\delta\tau \partial_\tau x^\rho) = \delta\dot{x}^\nu$.

5. *Newtonian limit*

It is instructive to consider the Newtonian limit where the static gravitational field is weak and the motion of the test particle is slow by writing

$$g_{\mu\nu} = \eta_{\mu\nu} + \gamma_{\mu\nu} \tag{198}$$

and assuming that γ is an infinitesimal tensor. The slow motion of the test particle leads to $\frac{dx^\mu}{ds} \approx (1, 0, 0, 0)$ which together with $g^{\mu\nu} = \eta^{\mu\nu} - \gamma^{\mu\nu}$ gives

$$\ddot{x}^\mu \approx -\Gamma^\mu_{00} = \frac{1}{2}\frac{\partial\gamma_{00}}{\partial x_\mu}, \tag{199}$$

or

$$\ddot{\mathbf{x}} = -\nabla\phi \tag{200}$$

where the static Newtonian potential is

$$\phi = \frac{\gamma_{00}}{2}. \tag{201}$$

B. Interacting matter-gravity system

The variation of the the Einstein action S_E with respect to the metric tensor leads to the vacuum Einstein equation, Eq. (172). When matter is included then the action becomes the sum of the gravitational and the matter actions,

$$S = S_E + S_M. \tag{202}$$

Therefore the matter contribution to the Einstein equation will be given by the expression

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi G M_{\mu\nu}(x) \tag{203}$$

with

$$M_{\mu\nu}(x) = \frac{2}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}(x)} \tag{204}$$

The source of the gravitational interaction is the mass. According to special relativity this corresponds to energy and covariance makes the whole energy-momentum tensor as the source. Thus one expects that the quantity (204) is proportional to the energy momentum tensor. But the Einstein equation (203) then expresses the vanishing of the full energy-momentum tensor. The vanishing of the total energy-momentum tensor is understandable because this latter is defined by carrying out space-time translations, an operation which becomes ill-defined without a background space-time.

A nontrivial condition, satisfied by the Einstein equation in the presence of matter is Eq. (135), amounts to the energy-momentum conservation, $D_\mu M^\mu_{\ \nu} = 0$, for any theory with action of the form (202). Hence the energy-momentum of the matter and the gravitation field are conserved separately. It is remarkable is that the gravity-matter interaction does not violate the conservation of the matter energy-momentum and the energy-momentum density and flux, the matrix elements of the energy-momentum tensor, are identical for the gravity and mater, except their sign.

We review briefly the energy-momentum tensor of a system of point particles, ideal fluid and a scalar field.

1. Point particle

Let us suppose that we have a particle of mass m moving along the world lines x^μ. The action is

$$S_M = -mc \int ds \sqrt{\dot{x}^\mu g_{\mu\nu} \dot{x}^\nu} \tag{205}$$

whose variation

$$\begin{aligned} \delta S_M &= -\frac{1}{2}mc\int ds\frac{\dot{x}^\mu \delta g_{\mu\nu}\dot{x}^\mu}{\sqrt{\dot{x}^\mu g_{\mu\nu}\dot{x}^\nu}} \\ &= -\frac{1}{2}mc\int ds\dot{x}^\mu\dot{x}^\nu\delta g_{\mu\nu} \\ &= -\frac{1}{2}\int dx\int ds\frac{p^\mu(s(t))p^\nu(s(t))}{mc}\delta(x-x(s))\delta g_{\mu\nu}(x), \end{aligned} \tag{206}$$

where the parameter of the world line is chosen to be the invariant length in the second equation, after having completed the variation. The definition (204) yields

$$\sqrt{-g}M^{\mu\nu}(x) = \int ds\frac{p^\mu(s(t))p^\nu(s(t))}{mc}\delta(x-x(s)). \tag{207}$$

The density of the four-momentum is

$$T^{\mu 0}(t,\mathbf{x}) = p^\mu(s(t))\delta(\mathbf{x}-\mathbf{x}(t)). \tag{208}$$

The tensor which reduces to this expression is

$$T^{\mu\nu}(x) = \frac{p^\mu(s(t))p^\nu(s(t))}{p^0(s(t))}\delta(\mathbf{x}-\mathbf{x}(t)), \tag{209}$$

and it can be written in a manifestly covariant form as

$$\begin{aligned} T^{\mu\nu}(x) &= \int ds\frac{p^\mu(s(t))p^\nu(s(t))}{p^0(s(t))}c\frac{dt}{ds}\delta(x-x(s)) \\ &= \int ds\frac{p^\mu(s(t))p^\nu(s(t))}{mc}\delta(x-x(s)), \end{aligned} \tag{210}$$

establishing $\sqrt{-g}M = T$.

2. *Ideal fluid*

It is worthwhile mentioning that in a more realistic situation one assumes a continuous distribution of matter. For homogeneous and isotropic matter in the rest frame we have

$$T^{\mu\nu} = \begin{pmatrix} \epsilon & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{211}$$

which can be written in a covariant manner as

$$T^{\mu\nu} = (p+\epsilon)\dot{x}^\mu\dot{x}^\nu - pg^{\mu\nu}, \tag{212}$$

because the relation $\dot{x}^\mu = (1, 0, 0, 0)$ holds in the rest frame. For ideal fluid where mean free path and times are short enough to maintain isotropy we have

$$T^{\mu\nu}(x) = (p(x) + \epsilon(x))u^\mu(x)u^\nu(x) - p(x)g^{\mu\nu}, \tag{213}$$

as the source term to the Einstein equation where $u^\mu(x) = \dot{x}^\mu(x)$ is the four-velocity of the fluid particles at the space-time point x.

3. Classical fields

We consider finally a simple scalar field theory with the action

$$S_M = \int dx\sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}\partial_\mu\phi(x)\partial_\nu\phi(x) - V(\phi(x))\right]. \tag{214}$$

The calculation of the variation with respect to the metric tensor is greatly simplified by the fact that the Lagrangian depends on the metric tensor and not its space-time derivatives,

$$M_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\partial\sqrt{-g}}{\partial g^{\mu\nu}}L + 2\frac{\partial L}{\partial g^{\mu\nu}}. \tag{215}$$

The use of Eq. (142) which gives $g^{\mu\nu}$ as the coefficient of the Lagrangian in the first term and the relation

$$\frac{\partial L}{\partial g^{\mu\nu}} = \frac{1}{2}g_{\nu\kappa}\frac{\partial L}{\partial\partial_\kappa\phi}\partial_\mu\phi \tag{216}$$

establishes the identity of $M_{\mu\nu}$ with the energy-momentum tensor

$$T_{\mu\nu} = \frac{\partial L}{\partial\partial^\mu\phi}\partial_\nu\phi - g_{\mu\nu}L, \tag{217}$$

given by (C41).

V. GRAVITATIONAL RADIATION

A gauge theory has two distinct sectors. One the one hand, there are dynamical degrees of freedom distributed in space-time in such a manner that they support a retarded or advanced signal, generated by external charges, like the far or radiation field in electrodynamics. One the other hand, there are "slave" modes which follow the motion of the external charges in an algebraic manner, without any non-trivial dynamics, like the near or Coulomb field of electrodynamics. We turn now to the former and mention few rudimentary features of gravitational radiation.

Let us start with the state without radiation: the flat Minkowski space-time solves trivially the Einstein equation. A weak radiation field,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{218}$$

with $|h_{\mu\nu}| \ll 1$ should not change the space-time geometry in a fundamental manner, an assumption which allows to consider General Relativity on the flat background space-time as a relativistic classical field theory. We shall discuss the plane wave solutions of the linearized Einstein equation in what follows.

A. Linearization

We have up to $\mathcal{O}(h)$

$$\begin{aligned}
\Gamma^{\rho}{}_{\mu\nu} &= \frac{1}{2}\eta^{\rho\sigma}(\partial_\mu h_{\nu\sigma} + \partial_\nu h_{\sigma\mu} - \partial_\sigma h_{\mu\nu}) = \frac{1}{2}(\partial_\mu h^\rho_\nu + \partial_\nu h^\rho_\mu - \partial^\rho h_{\mu\nu}),\\
R^{\mu}{}_{\nu\rho\sigma} &= \partial_\rho \Gamma^{\mu}{}_{\nu\sigma} - \partial_\sigma \Gamma^{\mu}{}_{\nu\rho} = \frac{1}{2}\partial_\rho(\partial_\nu h^\mu_\sigma + \partial_\sigma h^\mu_\nu - \partial^\mu h_{\nu\sigma}) - (\rho \leftrightarrow \sigma),\\
R_{\nu\sigma} &= \frac{1}{2}\partial_\mu(\partial_\nu h^\mu_\sigma + \partial_\sigma h^\mu_\nu - \partial^\mu h_{\nu\sigma}) - \frac{1}{2}\partial_\sigma(\partial_\nu h^\mu_\mu + \partial_\mu h^\mu_\nu - \partial^\mu h_{\nu\mu})\\
&= \frac{1}{2}(\partial_\nu\partial_\mu h^\mu_\sigma + \partial_\sigma\partial_\mu h^\mu_\nu - \Box h_{\nu\sigma} - \partial_\sigma\partial_\nu h),\\
R &= \partial_\nu\partial_\mu h^{\mu\nu} - \Box h,
\end{aligned} \tag{219}$$

where the indices are raised and lowered by $\eta^{\mu\nu}$ and $\eta_{\mu\nu}$, respectively as in an ordinary relativistic field theory and $h = h^{\mu}{}_{\mu}$.

A $GL(4)$ gauge transformation, an external diffeomorphism $x^\mu \to x^\mu + \xi^\mu(x)$ induces

$$h_{\mu\nu} \to h'_{\mu\nu} = h_{\mu\nu} + \partial_\mu\xi_\nu + \partial_\nu\xi_\mu. \tag{220}$$

The linearization in $h_{\mu\nu}$ suppresses the commutator of the field strength tensor in the third equation which now looks as an Abelian field strength. As a result our expression for the Riemann, Ricci and the Einstein tensors, as well as the scalar curvature are gauge invariant.

We shall use harmonic gauge (98) where

$$\partial^\nu h_{\mu\nu} = \frac{1}{2}\partial_\mu h. \tag{221}$$

This condition is satisfied after the gauge transformation which solves

$$\Box\xi_\nu = \frac{1}{2}\partial_\nu h - \partial_\rho h^\rho_\nu \tag{222}$$

since

$$\frac{\partial}{\partial x'^{\nu}}h'^{\nu}_{\mu} = \partial^{\nu}(h_{\mu\nu} + \partial_{\mu}\xi_{\nu} + \partial_{\nu}\xi_{\mu}) = \partial^{\nu}h_{\mu\nu} + \partial^{\nu}\partial_{\mu}\xi_{\nu} + \frac{1}{2}\partial_{\mu}h - \partial_{\rho}h^{\rho}_{\mu} = \frac{1}{2}\partial_{\mu}h + \partial_{\mu}\partial^{\nu}\xi_{\nu} = \frac{1}{2}\frac{\partial}{\partial x'^{\mu}}h'. \tag{223}$$

Note that the harmonic gauge condition is preserved by further gauge transformations which correspond to harmonic functions, $\Box\xi^{\mu} = 0$.

B. Wave equation

The linearized Einstein equation, (203),

$$\frac{1}{2}(\partial_{\nu}\partial_{\mu}h^{\mu}_{\sigma} + \partial_{\sigma}\partial_{\mu}h^{\mu}_{\nu} - \Box h_{\nu\sigma} - \partial_{\sigma}\partial_{\nu}h - \eta_{\nu\sigma}\partial_{\rho}\partial_{\mu}h^{\mu\rho} + \eta_{\nu\sigma}\Box h) - \Lambda h_{\nu\sigma} = 8\pi G T_{\nu\sigma} \tag{224}$$

contains the $\mathcal{O}\left(h^0\right)$, Minkowski energy-momentum tensor because $h = \mathcal{O}\left(G\right)$. The wave equation reads in harmonic gauge as

$$\Box h_{\nu\sigma} - \frac{1}{2}\eta_{\nu\sigma}\Box h + 2\Lambda h_{\nu\sigma} = -16\pi G T_{\nu\sigma}. \tag{225}$$

The equations simplify when expressed in terms of

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \tag{226}$$

where $\bar{h}^{\mu}_{\mu} = \bar{h} = -h$ because the gauge condition and the Einstein equation are

$$\partial_{\nu}\bar{h}^{\nu}_{\mu} = 0, \tag{227}$$

and

$$\Box\bar{h}_{\nu\sigma} + \Lambda(2\bar{h}_{\nu\sigma} - \eta_{\nu\sigma}\bar{h}) = -16\pi G T_{\nu\sigma}, \tag{228}$$

respectively. The cosmological constant, Λ, plays the role of a mass which makes the radiation field short ranged and will be ignored below. The retarded solution is

$$\bar{h}_{\nu\sigma}(t, \mathbf{x}) = 4G\int d^3x' \frac{T_{\nu\sigma}(t - |\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|} \tag{229}$$

which is automatically given in harmonic gauge due to $\partial_{\nu}T^{\nu\mu} = 0$.

C. Plane-waves

The linearized Einstein equation in vacuum is satisfied by plane waves,

$$\bar{h}_{\mu\nu} = e_{\mu\nu}e^{ikx} + e^{*}_{\mu\nu}e^{-ikx}, \tag{230}$$

given in terms of the polarization tensor $e_{\mu\nu}$. The wave vector is light-like, $k^2 = 0$ and the gravitational wave propagates on the line cone. The symmetric polarization tensor contains 10 parameters but the gauge fixing,

$$k^\mu e_{\mu\nu} = 0 \tag{231}$$

decreases the number of the independent parameters to 6.

We may perform a further gauge transformation, $x^\mu \to x^\mu + \xi^\mu$ with $\Box\xi^\mu = 0$ without leaving the harmonic gauge. The choice

$$\xi^\mu = ia^\mu e^{ikx} - ia^{*\mu} e^{-ikx} \tag{232}$$

transforms the polarization tensor into

$$e_{\mu\nu} \to e'_{\mu\nu} = e_{\mu\nu} + k_\mu a_\nu + k_\nu a_\mu \tag{233}$$

and leaves 2 independent parameters. Some simplification of the covariant expressions can be achieved by choosing the longitudinal and transverse component b^μ in such a manner that $e^\mu_\mu \to e^\mu_\mu + 2k^\mu a_\mu = 0$ and $e_{\mu 0} \to e_{\mu 0} + k_\mu a_0 + k_0 a_\mu = 0$, respectively.

It is instructive to compare this situation with electrodynamics in Lorentz gauge, $\partial^\mu A_\mu = 0$, where the plane waves satisfy the wave equation $\Box A_\mu = 0$. The solution

$$A_\mu(x) = e_\mu e^{ikx} + e^*_\mu e^{-ikx}, \tag{234}$$

$k^2 = 0$, has 4 independent parameters in the polarization vector e_μ and this number is reduced to 3 by the gauge condition, $k^\mu e_\mu = 0$. The gauge transformation, $A_\mu \to A_\mu + \partial_\mu\phi$, performed by a harmonic function $\Box\phi = 0$ leaves the Lorentz gauge condition unchanged and the choice $\phi(x) = iae^{ikx} - ia^* e^{-ikx}$ transforms the polarization vector, $e_\mu \to e_\mu - ak_\mu$, and can be used to reduce the free parameters to 2.

D. Polarization

Let us now consider a simple application, the effect of a gravitational plane wave on the motion of point particles, namely the equation of motion for the deformation of the geodesics, discussed in Section IV A 4. We assume stationary unperturbed particles, $u^\mu = (1, \mathbf{0}) + \mathcal{O}(h)$, $v = \mathcal{O}(h)$ and write Eq. (193) as

$$\partial_0^2 v^\rho = R^\rho{}_{00\nu} v^\nu = \frac{1}{2}\partial_0^2 h^\rho_\nu v^\nu, \tag{235}$$

where we used $h_{\mu 0}=0$ in the second equation. The gauge condition (231) shows that the deformation is transverse, $k_\mu v^\mu=0$. Since $u_\mu v^\mu=0$ the deformation is transverse in the spatial directions, as well, $\mathbf{kv}=0$, with $k^\mu=(k^0,\mathbf{k})$ and $v^\mu=(0,\mathbf{v})$.

We shall use coordinates where $k^\mu=(k,0,0,k)$. We know that (i) $k^\mu e_{\mu\nu}=0$, (ii) $e^\mu_\mu=0$ and (iii) $e_{0\nu}=0$. The equations (i) and (iii) imply $e_{3\nu}=0$, leaving $e_{jk}\neq 0$ for $j,k=1,2$ only. Due to (ii) the symmetric matrix $e_{\mu\nu}$ is traceless, leaving two independent components, $e_{11}=-e_{22}$ and $e_{12}=e_{21}$. When $e_{12}=0$ one finds

$$\partial_0^2 v^j=-\frac{1}{2}k^{02}(e_{11}e^{ikx}+e^*_{11}e^{-ikx})\epsilon^{jk}v^k, \tag{236}$$

yielding the solution

$$\begin{pmatrix}v^1(x)\\ v^2(x)\end{pmatrix}=\begin{pmatrix}[1+\frac{1}{2}(e_{11}e^{ikx}+e^*_{11}e^{-ikx})]v^1(0,\mathbf{x})\\ [1-\frac{1}{2}(e_{11}e^{ikx}+e^*_{11}e^{-ikx})]v^2(0,\mathbf{x})\end{pmatrix}. \tag{237}$$

A ring of particle in the $(1,2)$ plane oscillates horizontally and vertically, as shown in Fig. 3 (a). For $e_{11}=0$ we have

$$\partial_0^2 \overset{1}{\underset{2}{v}}=-\frac{1}{2}k^{02}(e_{12}e^{ikx}+e^*_{12}e^{-ikx})\overset{2}{\underset{1}{v}}, \tag{238}$$

and

$$\begin{pmatrix}v^1(x)\\ v^2(x)\end{pmatrix}=\begin{pmatrix}[1+\frac{1}{2}(e_{12}e^{ikx}+e^*_{12}e^{-ikx})]v^2(0,\mathbf{x})\\ [1+\frac{1}{2}(e_{12}e^{ikx}+e^*_{12}e^{-ikx})]v^1(0,\mathbf{x})\end{pmatrix}. \tag{239}$$

The direction of the deformation of the ring in the $(1,2)$ plane rotated by $\pi/4$ compared to the previous case as shown in Fig. 3 (b). The direction independent, isotope deformation is due to the monopole term of the multipole expansion, the deformation along a fixed direction indicate the presence of the dipole term and these deformations, carried out in two directions belong to the quadrupole order.

In the case of electrodynamics the plane wave (234) corresponds to the field strength tensor

$$F_{\mu\nu}=ik_\mu(e_\nu e^{ikx}-e^*_\nu e^{-ikx})-(\mu\leftrightarrow\nu), \tag{240}$$

which gives for Eq. (195)

$$\begin{aligned} mc\delta\ddot{x}^\mu &= \frac{e}{c}\{i[k^\mu(e_\nu e^{ikx}-e^*_\nu e^{-ikx})-k_\nu(e^\mu e^{ikx}-e^{*\mu}e^{-ikx})]\delta\dot{x}^\nu \\ &\quad -\delta x^\rho k_\rho[k^\mu(e_\nu e^{ikx}+e^*_\nu e^{-ikx})-k_\nu(e^\mu e^{ikx}+e^{*\mu}e^{-ikx})]\dot{x}^\nu\}. \end{aligned} \tag{241}$$

A plane wave, propagating in the spatial 3 direction is polarized in the $(1,2)$ plane and one finds deformations, characteristic of dipole field.

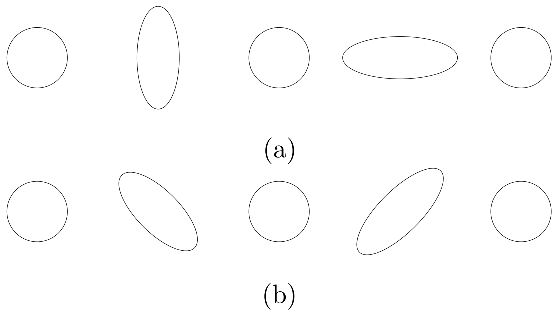


FIG. 3: The shape of a circle of particles in the $(1,2)$ plane as a function of time for (a): $e_{12} = 0$ (b): $e_{12} = 0$.

Gravitational radiation is too weak to be seen in a direct manner. But an indirect evidence is known, the slowing of the PSR1913+16 binary system, pulsar, is consistent with the energy loss, caused by the power, radiated.

VI. SCHWARZSCHILD SOLUTION

After a short digression into the dynamics of the propagating gravitational field let us now turn to a simpler problem and inquire about the physical phenomena of the non-propagating, near field sector. The simplest electrodynamics problem for fixed external charges is that of a point charge with the solution of the Coulomb force. The analogous problem, the gravitational field created by a static point mass is the Schwarzschild solution. We use the notation $x^0 = ct \to t$ in this section.

A. Metric

The symmetry of the corresponding space-time is time independence and rotational invariance. The rotational invariance requires that (i) the space-time can be foliated by a family of two-dimensional surfaces, $\Sigma(t,r)$ and (ii) that any pair of points of a hyper-surface there is a spatial rotation bringing one point into the other. The time independence assures that there is a time-like unit vector field $n(x)$, $n^2(x) = 1$, generating an infinitesimal transformation of the space-time, $x^\mu \to x^\mu + \epsilon n^\mu(x)$, which leaves the geometry, the metric tensor in particular, invariant. The spatial rotations generate displacement orthogonal to the time direction defined by the vector field $n^\mu(x)$ therefore the most general static, rotational invariant metric is of the form

$$ds^2 = f(r)dt^2 - \sum_{j,k=1}^{3} x^j h_{jk}(r) x^k \tag{242}$$

where $r = \sqrt{\mathbf{x}^2}$, $h_{jk}(r)$ is a three-dimensional, rotational invariant metric. Let us chose the radial coordinate

$$r = \sqrt{\frac{A}{4\pi}} \tag{243}$$

where A is the time independent area of the surface $\Sigma(t, r)$ and parameterize this surface by the polar angle parameters θ and ϕ, giving

$$ds^2 = f(r)dt^2 - h(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{244}$$

The non-vanishing Christoffel symbols of this metric are

$$\begin{aligned}
\Gamma^t{}_{rt} &= \frac{f'}{2f}, \quad \Gamma^r{}_{tt} = \frac{f'}{2h} \\
\Gamma^r{}_{rr} &= \frac{h'}{2h}, \quad \Gamma^r{}_{\theta\theta} = -\frac{r}{h}, \quad \Gamma^r{}_{\phi\phi} = -\frac{r\sin^2\theta}{h}, \quad \Gamma^\theta{}_{r\theta} = \Gamma^\phi{}_{r\phi} = \frac{1}{r} \\
\Gamma^\theta{}_{\phi\phi} &= -\sin\theta\cos\theta, \quad \Gamma^\phi{}_{\theta\phi} = \cot\theta
\end{aligned} \tag{245}$$

and the other non-vanishing components can be obtained by exchanging the covariant indices. The Ricci tensor is diagonal,

$$\begin{aligned}
R_{tt} &= -\frac{f''}{2h} + \frac{f'}{4h}\left(\frac{f'}{f} + \frac{h'}{h}\right) - \frac{f'}{rh} \\
R_{rr} &= \frac{f''}{2f} - \frac{f'}{4f}\left(\frac{f'}{f} + \frac{h'}{h}\right) - \frac{h'}{rh} \\
R_{\theta\theta} &= -1 + \frac{r}{2h}\left(\frac{f'}{f} - \frac{h'}{h}\right) + \frac{1}{h} \\
R_{\phi\phi} &= R_{\theta\theta}\sin^2\theta.
\end{aligned} \tag{246}$$

The vacuum Einstein equation for $r \neq 0$ where $R = 0$ are

$$R_{tt} = R_{rr} = R_{\theta\theta} = 0. \tag{247}$$

Since

$$0 = \frac{R_{tt}}{f} + \frac{R_{rr}}{h} = -\frac{1}{rh}\left(\frac{f'}{f} + \frac{h'}{h}\right) \tag{248}$$

we have

$$\frac{f'}{f} + \frac{h'}{h} = 0 \tag{249}$$

requiring

$$hf = A, \tag{250}$$

A being a constant. The condition that the metric approaches the flat one for $r \to \infty$ gives $A = 1$ and find for the last equation in (247)

$$0 = rf' + f - 1 \tag{251}$$

or

$$\frac{drf}{dr} = 1. \tag{252}$$

The solution is

$$f = 1 + \frac{B}{r} \tag{253}$$

where B is a constant. The parametrization $B = -2GM/c^2$ yields the metric

$$ds^2 = \left(1 - \frac{r_s}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{r_s}{r}} - r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{254}$$

where

$$r_s = \frac{2GM}{c^2} \tag{255}$$

is the Schwarzschild-radius (with $c \neq 1$ restored). It is advantageous to remove the dimension of the coordinates by the intrinsic scale r_s, $t \to tr_s$ and $r \to rr_s$ and write the dimensionless invariant length square as

$$ds^2 = \left(1 - \frac{1}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{1}{r}} - r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{256}$$

Few remarks are in order at this point.

1. The gravitational field is weak for $r \gg 1$ and Newton's gravitational law applies approximately according to Eq. (201).

2. Consider two stationary observers, two signals are emitted by one at time t_{e1} and t_{e2} from radius r_e and received by the other at time t_{r1} and t_{r2} and radius r_r. The geometry is time independent hence $t_{r1} - t_{e1} = t_{r2} - t_{e2}$, implying $t_{r2} - t_{r1} = t_{e2} - t_{e1}$. The proper time passed between the two signals,

$$\Delta s_e = \sqrt{1 - \frac{1}{r_e}}(t_{e2} - t_{e1}), \quad \Delta s_r = \sqrt{1 - \frac{1}{r_r}}(t_{r2} - t_{r1}), \tag{257}$$

lead to the redshift

$$\frac{\Delta s_e}{\Delta s_r} = \frac{\nu_r}{\nu_e} = \sqrt{\frac{1 - \frac{1}{r_e}}{1 - \frac{1}{r_r}}}. \tag{258}$$

3. The metric shows two singularities, one at the Schwarzschild radius

$$r_s \approx \begin{cases} 2.8\frac{M}{M_{sun}} \text{ km} \\ 2.4\frac{M}{M_{proton}} \cdot 10^{-52} \text{ cm} \end{cases} \tag{259}$$

and another at $r = 0$. The former turns out to be a singularity of this coordinate system because the curvature tensor remain regular and can be eliminated by means of appropriately defined coordinates. The latter is a true singularity.

4. A stationary observer's four-velocity is $u_t = \frac{1}{\sqrt{1-\frac{1}{r}}}$, $u_r = u_\theta = u_\phi = 0$. Its four-acceleration,

$$a^\mu = u^\nu D_\nu u^\mu = u^\nu \partial_\nu u^\mu + \Gamma^\mu{}_{\rho\nu} u^\rho u^\nu = \frac{\Gamma^\mu{}_{tt}}{1 - \frac{1}{r}}, \tag{260}$$

has a single non-vanishing component, $a^r = \frac{1}{2r^2} = \phi'(r)$ where $\phi(r) = -\frac{1}{2r}$ is the Newtonian potential according to Eq. (201), canceling the gravitational force, in agreement with the Equivalence Principle. But the regularity of the acceleration is misleading because the physical, gauge invariant acceleration is obtained by multiplying it by $\sqrt{g_{rr}}$,

$$\sqrt{-a^\mu g_{\mu\nu} a^\nu} = \frac{1}{2r^2\sqrt{1 - \frac{1}{r}}}, \tag{261}$$

and it diverges at $r = 1$: An extended system, bound by elementary particles is torn into pieces as it approaches $r = 1$ from above.

5. Despite the absence of a singularity something dramatic happens at the Schwarzschild radius. The light cone corresponds to the infinitesimal changes

$$\frac{dt}{dr} = \pm\frac{1}{1 - \frac{1}{r}}, \tag{262}$$

the light cones have space-dependent orientation which becomes singular at $r = 1$, the causal structure changes discontinuously at the Schwarzschild radius. Another pathology is seen by considering the motion of a massive particle where $ds^2 > 0$. This inequality is consistent with constant r, the particle can be at rest compared to the Schwarzschild radius if $r > 1$. But this is not possible anymore for $r < 1$ where the role of the time t as a coordinate with positive unit vector is taken over the the radius r and thereby is forced to change during the motion.

6. The solution requires that the mass be concentrated at $r = 0$. For any realistic, non-singular mass distribution the solution is more complicated in the space region with non-vanishing

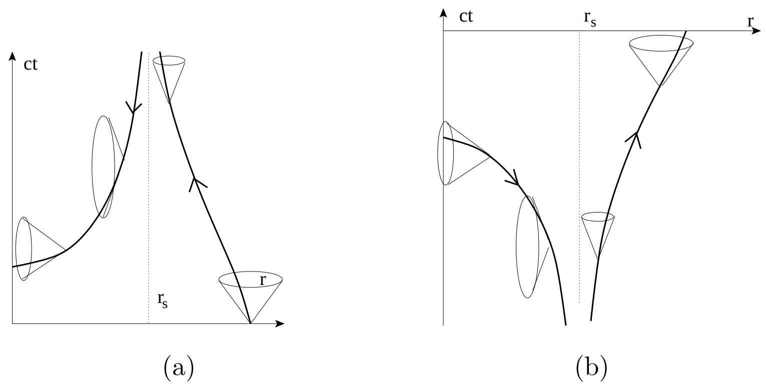


FIG. 4: The light-cone structure of time-like geodesics and an inward (a) and outward (b) massive particle world line.

mass density. The Schwarzschild-radius is naked and visible only for mass distributions which are vanishing for $r \geq 1$. the numerical values given in Eq. (259) suggest that the Schwarzschild-radius might be found experimentally in astrophysics rather than in particle physics.

7. A massive point particle has two characteristic length scales: The Compton wavelength, $\lambda_C = \hbar/mc$, denotes the maximal localization the particle can have because its restriction into a region with shorter size leads to pair creation and the loosing sight of the original particle. In other words, a point particle is surrounded with a virtual particle-anti particle cloud of the size λ_C. The other length scale, the Schwarzschild radius, increases with the mass and the two scales coincide at $m = m_{Pl}/\sqrt{2}$ where $m_{Pl} = \sqrt{\hbar c/G} \sim 2.1 \times 10^{-5}$g denotes the Planck-mass. A point particle which is lighter or heavier than the Planck mass is surrounded by a cloud of virtual pairs (quantum effect in approximately flat space-time) or appears as a Schwarzschild sphere (strong gravitational field effect).

8. The solution remains the same when time independence is not assumed at the beginning, namely the spherically symmetric solutions of the vacuum Einstein equation are static (Kirchhoff's theorem). This holds in the Newtonian theory in an obvious manner since the mass can be concentrated at the origin for spherically symmetrical field. This theorem excludes the spherically symmetric s-waves from gravitational radiation field.

B. Geodesics

Let us consider the motion of a massive particle in the Schwarzschild geometry where the Lagrangian

$$L = -mc\sqrt{\dot{x}^\mu g_{\mu\nu}(x)\dot{x}^\nu} \tag{263}$$

can be written as

$$L = -mc\sqrt{\left(1-\frac{1}{r}\right)\dot{t}^2 - \frac{\dot{r}^2}{1-\frac{1}{r}} - r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2)}. \tag{264}$$

The motion is always planar, equation of motion for θ,

$$r^2\sin\theta\cos\theta\dot{\phi}^2 = \frac{d}{ds}r^2\dot{\theta} \tag{265}$$

is satisfied by $\theta(s) = \pi/2$, the case considered hereafter. The coordinates t and ϕ are cyclic therefore the corresponding generalized momentums are conserved,

$$\begin{aligned} -\frac{1}{mc}\frac{\partial L}{\partial \dot{t}} &= \left(1-\frac{1}{r}\right)\dot{t} = E, \\ \frac{1}{mc}\frac{\partial L}{\partial \dot{\phi}} &= r^2\dot{\phi} = \ell, \end{aligned} \tag{266}$$

One usually solves the non-relativistic radial equation of motion by exploiting the energy conservation. Since the temporal component of the relativistic equation of motion is the energy conservation such a starting point corresponds to the use of the equation $\dot{x}^2 = \kappa$, giving

$$\left(1-\frac{1}{r}\right)\dot{t}^2 - \frac{\dot{r}^2}{1-\frac{1}{r}} - r^2\dot{\phi}^2 = \kappa. \tag{267}$$

We set $\kappa = 1$ in this calculation of the orbit of a massive object and find the radial equation of motion

$$\frac{E^2 - \dot{r}^2}{1-\frac{1}{r}} - \frac{\ell^2}{r^2} = \kappa \tag{268}$$

what we write

$$\dot{r}^2 + V(r) = E^2 \tag{269}$$

in terms of the effective potential

$$V(r) = \left(1-\frac{1}{r}\right)\left(\kappa + \frac{\ell^2}{r^2}\right). \tag{270}$$

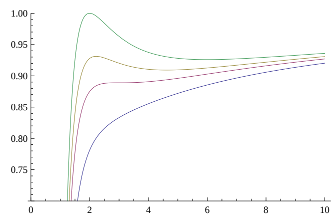


FIG. 5: The effective potential, (270), as the function of r/r_s for $\ell^2/r_s^2 = 1.5, \sqrt{3}, 1.85, 2.0$ in increasing order.

The contribution proportional to $\kappa = 1$ is the Newtonian effective potential and the remaining $\mathcal{O}\left(r^3\right)$ piece represents relativistic effects.The motion with $|E| < 1$ is bounded, $0 < r < r_{max}$, because $V(0) = -\infty$ and $V(\infty) = 1$. The extremes of the potential satisfy the equation

$$\frac{dV(r)}{d\frac{1}{r}} = -3\frac{\ell^2}{r^2} + 2\frac{\ell^2}{r} - 1 = 0. \tag{271}$$

The potential is monotonous when $\ell < \sqrt{3}$ and the particle falls into the center. For $\ell > \sqrt{3}$ the potential displays a local maximum at are

$$r_{max} = \frac{3}{1 + \sqrt{1 - \frac{3}{\ell^2}}}, \tag{272}$$

and a local minimum at

$$r_{min} = \frac{3}{1 - \sqrt{1 - \frac{3}{\ell^2}}}, \tag{273}$$

cf. Fig. 5 and there are stable orbits in certain range of E.

The presence of the $\mathcal{O}\left(r^{-3}\right)$ relativistic term in the effective potential (270) violates Kepler's law for planetary motion, in particular it induces a perihelion motion. This brought the first decisive victory for General relativity when Einstein could reproduce the perihelion motion of Mercury, known since in 1859.

The Euler-Lagrange equation of the Lagrangian of a massive particle, (189), gives a geodesics. One expects that the world line of a light particle approaches the motion of a massless particle when the mass tends to zero. This can be easily established in an obvious manner for photons in the geometrical limit by means of Fermat's principle. By assuming here that massless particles follow null-geodesics the previous consideration remains valid with $\kappa = 0$. Photons with $\kappa = 0$ do not feel the Newtonian gravitational potential as expected but their orbital angular momentum is coupled to gravitation. This is not surprising since the affine connection term in the covariant derivative couples the polarization of the electromagnetic radiation to gravity. The polarization follows the change of the direction of propagation and couples the orbital angular momentum to gravity. The deflection of light around the Sun has been observed first in 1919 and was the second major support of General Relativity.

C. Space-like hyper-surfaces

The Schwarzschild geometry is static and its non-trivial features are captured by the constant time hyper-surfaces, a one-dimensional manifold of curved three-dimensional spaces, parametrized by t. Each three-dimensional hyper-surface can be embedded into a four dimensional Euclidean space. To simplify matters we consider the two-dimensional $\theta = \pi/2$ section of the hyper-surfaces, obeying the metric

$$-ds^2 = \frac{dr^2}{1-\frac{1}{r}} + r^2 d\phi^2. \tag{274}$$

This surface can easily be embedded into a three-dimensional Euclidean space by means of the cylindrical coordinates (z, r, ϕ). The surface $z = z(r)$ in the Euclidean three-space space of metric

$$-ds^2 = dz^2 + dr^2 + r^2 d\phi^2 \tag{275}$$

has the invariant length

$$-ds^2 = \left[1 + z'^2(r)\right] dr^2 + r^2 d\phi^2, \tag{276}$$

where $z'(r) = \frac{dz(r)}{dr}$. The comparison with (274) gives

$$1 + z'^2(r) = \frac{1}{1-\frac{1}{r}}, \tag{277}$$

and

$$z(r) = \int_1^r \frac{dr'}{\sqrt{r'-1}} = 2\sqrt{r-1}, \tag{278}$$

defined for $r > 1$ only.

D. Around the Schwarzschild-horizon

The causal structure of space-time is determined by the local light cones because any signal or interaction can propagate on their surface or within them. The light cones, given by Eq. (262) become narrow and narrow as the Schwarzschild-radius is approached from above, as shown in

Fig. 4. This indicates that the free fall motion, seen by a stationary observer slows down as the horizon is approached. This can be understood as a manifestation of the red-shift, mentioned in point 2. of SectionVI A. The role of the radial and the time coordinate of the metric is exchanged for a time-like geodesic as it traverses $r = 1$ and the light-cones are oriented horizontally. The Kruskal-Szekeres coordinate system, introduced below shows that such an orientation of the light-cones makes that no physical object of signal can crosses $r = 1$ from below. Therefore the sphere $r = 1$ is a horizon which can be traversed inside only and appears for the outside observers as a black hole.

The metric (256) has a coordinate singularity only at the Schwarzschild-radius because the non-vanishing components of the curvature tensor

$$R_{t\theta t\theta} = R_{t\phi t\phi} = -R_{r\theta r\theta} = -R_{r\phi r\phi} = \frac{1}{2}R_{\theta\phi\theta\phi} = -\frac{1}{2}R_{trtr} = \frac{1}{r^3} \tag{279}$$

make the eigenvalues, the invariant content of the curvature tensor, regular. Therefore a point particle experiences nothing irregular or special when crossing the horizon. But an object which is extended in the radial direction suffers strong tidal forces at the horizon. This is because a small, finite separation, Δr, corresponds to diverging invariant length, $\sqrt{-\Delta s^2} = \Delta r/\sqrt{1 - 1/r}$, as $r \to 1$.

1. *Falling through the horizon*

It is instructive to follow the radial free fall of a point particle through the horizon. The equation of motion of a massive particle for $\ell = 0$ is

$$\dot{r}^2 = \frac{1}{r} + E^2 - 1. \tag{280}$$

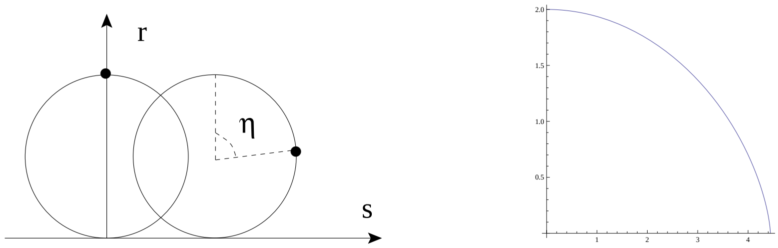
Let us suppose that the motion starts with vanishing velocity at $r = r_0 > 1$ and write

$$ds = \frac{dr}{\sqrt{\frac{1}{r} - \frac{1}{r_0}}}. \tag{281}$$

Such a $r(s)$ function can easily be obtained in a parametrized form,

$$\begin{aligned} r &= \frac{r_0}{2}(1 + \cos\eta), \\ s &= \frac{r_0^{3/2}}{2}(\eta + \sin\eta), \end{aligned} \tag{282}$$

known from the description of the motion of a point on a circle, rolling with constant speed. Nothing special happens at $r = 1$ and the total proper time of falling into the center at $\eta = \pi$ is $s = \frac{\pi}{2}r_0^{3/2}$, cf. Fig. 6.



(a) (b)

FIG. 6: (a): The geometrical origin of the parametrization (282). (b): The radius r as the function of the proper time, s, for the free fall from $r = 2$.

2. Stretching the horizon

We have seen so far, that the free fall, followed by its proper time shows no particular singularity at the horizon. But what does a stationary observer finds which uses the Schwarzschild coordinates t and r? The Schwarzschild time can be found from the first equation of (266),

$$\dot{r} = \frac{dr}{dt}\dot{t} = \frac{dr}{dt}\frac{E}{1-\frac{1}{r}}, \tag{283}$$

indicating a singularity for such an observer when the horizon is crossed. Such a singularity might be avoided by the use of the coordinate r^* satisfying

$$dr^* = \frac{dr}{1-\frac{1}{r}}, \tag{284}$$

because the singular factor is absorbed in the new coordinate and our equation for the radius now reads as

$$\dot{r} = \frac{dr^*}{dt}E. \tag{285}$$

The solution of Eq. (284),

$$r^* = r + \ln|r-1|, \tag{286}$$

replaced into the equation of motion, (280), gives

$$E^2\left[1-\left(\frac{dr^*}{dt}\right)^2\right] = 1-\frac{1}{r}. \tag{287}$$

If $r \to 1$ from above then $r^* \to -\infty$,

$$\frac{dr^*}{dt} \to -1 \tag{288}$$

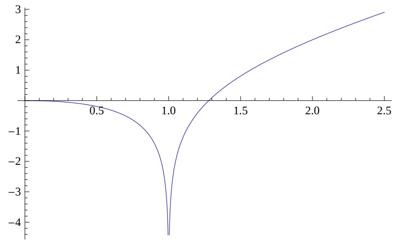


FIG. 7: The tortoise coordinate, r^* as function of r.

and $t \sim -r^* \to \infty$, it takes infinitely long time to fall through the Schwarzschild-radius. It is easy to understand that this is the result of the apparent singularity in the Schwarzschild metric, $g_{tt} \to 0$, $g_{rr} \to \infty$ with $g_{tt}g_{rr} = 1$ as $r \to 1$.

The radius r^*, defied by Eq. (286) is called tortoise coordinate. We are told that Achilles could not pass a tortoise because each time he reached the point where the tortoise was before it was already ahead. The singularity of the Schwarzschild metric, the factor $1/(1-1/r)$ multiplying dr^2 in the expression of the invariant length (256), signals that the scale of the radius should be refined, r should be allowed to decrease beyond zero to describe the free fall through the horizon. We have $r^* \sim r$ and $r^* \to -\infty$ as $r \to 1$, shown qualitatively in Fig. 7, stretching out conveniently the approach of the horizon from either side.

3. Szekeres-Kruskall coordinate system

The Schwarzschild time diverges on both sides of the horizon as indicated in Fig. 4. To resolve the traverse the horizon we need better suited coordinates. Instead of the coordinates t and r one may use $u = t - r$ and $v = t + r$ in Minkowski flat space-time, labeling the out- and in-going light rays, respectively. The expression of the invariant length is $ds^2 = dt^2 - dr^2 = dudv$, showing clearly that the new coordinates correspond to light cones because $ds^2 = 0$ for $du = 0$ or $dv = 0$. We can preserve the light cone structure of the Schwarzschild geometry by the help of the new coordinates,

$$u^* = t - r^*, \quad v^* = t + r^*. \tag{289}$$

The relation

$$dr^2 = \left(1 - \frac{1}{r}\right)^2 dr^{*2} \tag{290}$$

yields the metric

$$ds^2 = \left(1 - \frac{1}{r}\right) du^* dv^* - r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{291}$$

which still displays a singularity at the horizon.

However this singularity is now less troublesome and a simple rescaling,

$$\begin{aligned}(I) \quad u' &= -e^{-\frac{u^*}{2}} = -\sqrt{r-1}e^{\frac{r-t}{2}}, \quad v' = e^{\frac{v^*}{2}} = \sqrt{r-1}e^{\frac{r+t}{2}},\\ (II) \quad u' &= e^{-\frac{u^*}{2}} = \sqrt{1-r}e^{\frac{r-t}{2}}, \quad v' = e^{\frac{v^*}{2}} = \sqrt{1-r}e^{\frac{r+t}{2}},\end{aligned} \tag{292}$$

where the transformations (I) and (II) apply outside and inside of the Schwarzschild sphere, respectively, removes the singularity since the metric in terms of the new dimensionless coordinate,

$$ds^2 = \frac{4}{r}e^{-r}du'dv' - r^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{293}$$

is regular at $r = 1$. The coordinates u' and v' correspond to null-directions, it is more natural to use the time- and space-like coordinates,

$$\rho = \frac{v' - u'}{2}, \quad \tau = \frac{v' + u'}{2}, \tag{294}$$

given by

$$\begin{aligned}(I) \quad \rho &= \sqrt{r-1}e^{\frac{r}{2}}\cosh\frac{t}{2}, \quad \tau = \sqrt{r-1}e^{\frac{r}{2}}\sinh\frac{t}{2},\\ (II) \quad \rho &= \sqrt{1-r}e^{\frac{r}{2}}\sinh\frac{t}{2}, \quad \tau = \sqrt{1-r}e^{\frac{r}{2}}\cosh\frac{t}{2},\end{aligned} \tag{295}$$

yielding the metric

$$ds^2 = \frac{4}{r}e^{-r}(d\tau^2 - d\rho^2) - r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{296}$$

The final problem to overcome is that this transformation covers a part of the space-time only because it gives $v' > 0$. The regions with $v' < 0$ can be obtained by using

$$\begin{aligned}(III) \quad u' &= e^{-\frac{u^*}{2}} = \sqrt{r-1}e^{\frac{r-t}{2}}, \quad v' = -e^{\frac{v^*}{2}} = -\sqrt{r-1}e^{\frac{r+t}{2}},\\ (IV) \quad u' &= -e^{-\frac{u^*}{2}} = \sqrt{1-r}e^{\frac{r-t}{2}}, \quad v' = -e^{\frac{v^*}{2}} = -\sqrt{1-r}e^{\frac{r+t}{2}},\end{aligned} \tag{297}$$

instead of (292), yielding

$$\begin{aligned}(III) \qquad \rho &= -\sqrt{r-1}e^{\frac{r}{2}}\cosh\frac{t}{2}, \quad \tau = -\sqrt{r-1}e^{\frac{r}{2}}\sinh\frac{r}{2},\\ (IV) \qquad \rho &= -\sqrt{1-r}e^{\frac{r}{2}}\sinh\frac{t}{2}, \quad \tau = -\sqrt{1-r}e^{\frac{r}{2}}\cosh\frac{t}{2}.\end{aligned} \tag{298}$$

The transformation, given by eqs. (295) and (298) defines the Kruskal-Szekeres coordinates, satisfying

$$(r-1)e^r = \rho^2 - \tau^2 \ (I),\ (II),\ (III) \text{ and } (IV), \quad t = \begin{cases} 2\text{arcth}\frac{\tau}{\rho} & (I) \text{ and } (III),\\ 2\text{arcth}\frac{\rho}{\tau} & (II) \text{ and } (IV),\end{cases} \tag{299}$$

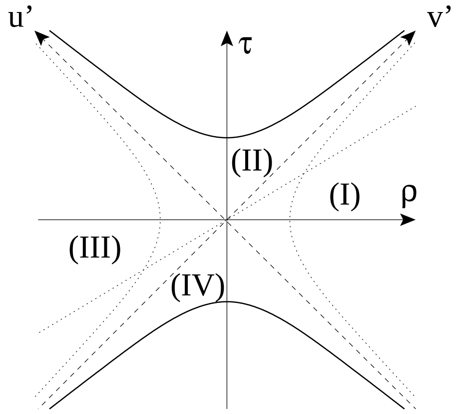


FIG. 8: The mapping of two Schwarzschild space-time onto the Kruskal-Szekeres geometry. The constant time and radius curves are radial or horizontally opening hyperbolic curves, respectively, such as the dotted lines and $t = -\infty \to \rho = -\tau$, $t = 0 \to \tau = 0$, $t = \infty \to \rho = \tau$. The outside and inside of the horizon $r = r_s \to \rho = \pm\tau$, is mapped into (I)-(III) and (II)-(IV), respectively and the Schwarzschild geometry, (I)-(II) is reduplicated into (III)-(IV).

cf. Fig. 8.

The matching of the two space-times in the outer region can be demonstrated by the Einstein-Rose bridge, the extension of the embedding, described in Section VI C. One chooses an Euclidean three-space, parametrized by by the coordinates (z, u, ϕ) and equipped by the metric

$$-ds^2 = dz^2 + d\rho^2 + r^2 d\phi^2. \tag{300}$$

The surface $z = z(\rho)$ has the induced metric

$$-ds^2 = [1 + z'^2(\rho)]d\rho^2 + r^2 d\phi^2, \tag{301}$$

which is to be matched to (296),

$$-ds^2 = \frac{4}{r}e^{-r}d\rho^2 + r^2 d\phi^2. \tag{302}$$

Hence the equation

$$1 + z'^2(\rho) = \frac{4}{r}e^{-r} \tag{303}$$

follows. The solution, as a function of the Schwarzschild radius r,

$$z(r) = \int_1^r dr' \frac{d\rho}{dr}\sqrt{\frac{4}{r'}e^{-r'} - 1}, \tag{304}$$

is sketched qualitatively in Fig. 9. It remains a challenge to understand the effects of such a dramatic reduplication of the Universe at each point particle.

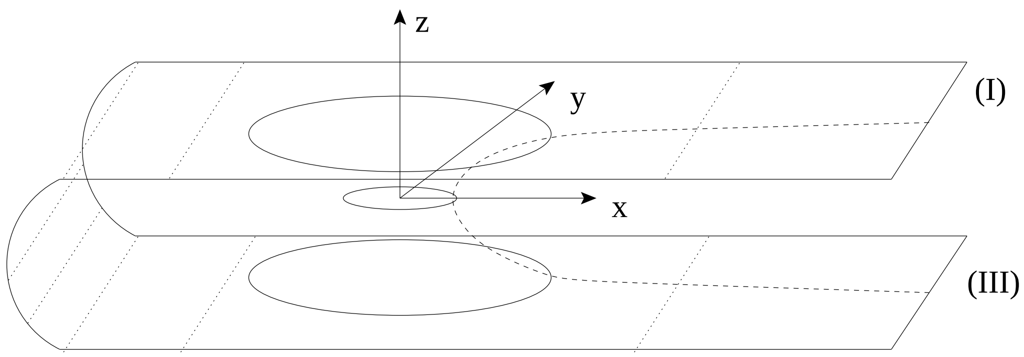


FIG. 9: The Einstein-Rosen bridge, connecting the two space-times in the $(z, \rho\cos\phi, \rho\sin\phi)$ coordinate system. The regions (I) and (III) belong to $z > 0$ and $z < 0$, respectively. A strip of the surface $\rho = \infty$ is indicated by the dotted lines, they form the horizontal part of a bended plane. The bending represents the analytical continuation between the regions (I) and (III). It is supposed to be at infinitely far, leaving the throat, indicated by the circles, ϕ-independent, rotation invariant. The dashed line follows the path $0 < \rho < \infty$ in (I) and (III).

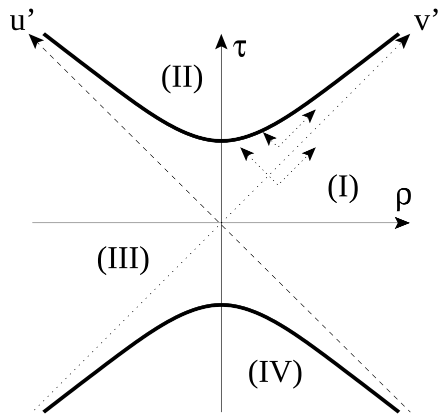


FIG. 10: The future light cones on the Kruskal-Szekeres space-time, indicated by the dotted lines running parallel with the u' and v' axes. The fat solid line represents the singular point $r = 0$.

4. Causal structure

The causal structure can easily recognized in the Kruskal-Szekeres space-time because τ is a time-like coordinate everywhere and the light-cones preserve their direction, contrary to the Schwarzschild parametrization, displayed in Fig. 4. The equal time lines indicate that the Schwarzschild time t runs in opposite direction in the two space-times, (I)-(II) and (III)-(IV). On can see immediately that the two outer regions, (I) and (III) are causally disconnected.

Though time and space coordinates appear in equal footing in relativity, the time has a distinguished role, it parametrizes the motion and has an orientation, encoded by the time arrow.

The radial ($L = 0$) null geodesic, line cones, are parallel to the coordinate axes u' and v' and two forward oriented light cones of a particle, slightly before and after crossing the Schwarzschild radius are indicated in Fig. 10. One can see that before and after the crossing the ingoing light rays fall into the singularity but the outward oriented light rays stay outside of the horizon or fall later into the center, respectively. The massive particle world lines remain within the future light cones, therefore no particle, either massive or massless, is able to emerge from the region $r < 1$ to $r > 1$: The Schwarzschild-sphere is impenetrable from inside. Note that the horizon is free of singularities and its one-way oriented, irreversible passage results from an appropriate rearrangement of the future light cones without strong forces.

VII. HOMOGENEOUS AND ISOTROPIC COSMOLOGY

The assumption that we do not occupy any special location in the Universe suggests that there should be a coordinate system in which the matter distribution in the Universe appears homogeneous on large enough distance scales. Furthermore, the absence of a preferred direction suggests isotropy, as well. The astrophysical observations supports these assumptions with an astonishing precision.

A gauge or coordinate system independent way of stating spatial homogeneity is to impose the existence of a one dimensional family of hyper-surfaces, $\Sigma(t)$ which (i) foliate the space-time, ie. any space-time point corresponds to one and only one hyper-surface and (ii) for any pair of points $p, q \in \Sigma(t)$ there is an isometry, a scalar product preserving mapping of the space-time which sends p into q.

Isotropy states that for any pair of space-like unit vectors, $v(p)$ and $v'(p)$ at a given space-time point p there is an isometry of the space-time which rotates $v(p)$ into $v'(p)$. The Universe appears isotropic in good approximation.

We introduce the concept of standard observers which find the the distant galaxies at rest. Their world lines provides a time-like congruence, system of time-like curves which fill up the space-time.

It is worthwhile mentioning two remarks at this point. (i) The tangent vectors of the time-like congruence are orthogonal to the tangent vectors of the space-like hyper-surfaces of the homogeneity assumption at each space-time point. In fact, otherwise one could construct a preferred spatial direction which contradicts isotropy. (ii) The space-time metric g induces a three-dimensional metric h on the hyper-surfaces. The isometry which maps a point of a hyper-surface $\Sigma(t)$ into another one according to the assumption of homogeneity is clearly an isometry of this three-dimensional induced metric, too. Furthermore, according to the previous remark this three-dimensional geometry has no preferred directions.

A. Maximally symmetric spaces

The homogeneous and isotropic three-space at a given time appears as symmetric as possible. To make this concept more precise let us consider geometries with maximal number of symmetries, Killing vectors.

The reparametrization $x^\mu \to x'^\mu = x^\mu - w^\mu(x)$ is a symmetry of the metric tensor if the deformation $w(x)$ is a Killing field, satisfying the Killing equation, (137). The parallel transport along an infinitesimal rectangle gives the equation

$$D_\mu D_\nu w_\rho - D_\nu D_\mu w_\rho = -R^\lambda{}_{\rho\mu\nu} w_\lambda. \tag{305}$$

There are two similar equations, obtained by cyclic permutation of the indices,

$$\begin{aligned} D_\nu D_\rho w_\mu - D_\rho D_\nu w_\mu &= -R^\lambda{}_{\mu\nu\rho} w_\lambda, \\ D_\rho D_\mu w_\nu - D_\mu D_\rho w_\nu &= -R^\lambda{}_{\nu\rho\mu} w_\lambda. \end{aligned} \tag{306}$$

the sum of these three equation can be written for a Killing field as

$$D_\mu D_\nu w_\rho - D_\nu D_\mu w_\rho + D_\rho D_\mu w_\nu = 0, \tag{307}$$

due to cyclic symmetry (114). The expression (305) of the parallel transport allows us to rewrite this equation in the form

$$D_\rho D_\mu w_\nu = R^\lambda{}_{\rho\mu\nu} w_\lambda, \tag{308}$$

stating that second (covariant) derivative of the Killing field can be expressed in terms of the Killing field itself, w. The x-dependence is therefore given by the value of the Killing field and it first derivative at a given point, $w^\mu(x_0)$, $D_\nu w^\mu(x_0)$. We need yet another property of the vector fields, a set of fields $\{w^\mu_n(x)\}$ is called independent if the the vanishing of the linear superposition, made up by constant coefficient,

$$\sum_n c_n w^\mu_n(x) = 0 \tag{309}$$

implies $c_n = 0$.

We have d independent vectors $w^\mu(x_0)$ and $d(d-1)/2$ independent anti-symmetric tensors $D_\nu w_\mu(x_0) - D_\mu w_\nu(x_0)$ at each point in d-dimensions hence the maximally symmetric space has

$d + d(d-1)/2 = d(d+1)/2$ independent Killing vectors. For instance the d-dimensional Euclidean space has d translational and $d(d-1)/2$ rotational symmetries. Furthermore, homogeneous and isotropic spaces have maximal symmetry.

B. Robertson-Walker metric

Let us consider now the three-dimensional curvature tensor, $\tilde{R}^j_{k\ell m}$, more precisely the tensor

$$\tilde{R}^{jk}_{\ell m} = \tilde{R}^j_{n\ell m} h^{kn} \tag{310}$$

which can be thought as a transformation of second order antisymmetric contravariant tensors,

$$T^{\ell k} = -T^{k\ell} \to \tilde{R}^{k\ell}_{mn} T^{mn}. \tag{311}$$

The matrix $\tilde{R}$ which acts on the tensor space is symmetric according to the second line in Eqs. (131) therefore it can be diagonalized. Isotropy of the tree-space, the absence of preferred direction requires that the matrix of this map be degenerate,

$$\tilde{R}^{jk}_{\ell m} = K(\delta^j_\ell \delta^k_m - \delta^k_\ell \delta^j_m). \tag{312}$$

The degenerate eigenvalue K is related to the scalar curvature,

$$\tilde{R} = \tilde{R}^{jk}_{jk} = Kd(d-1) = k|\tilde{R}|, \tag{313}$$

with $k = -1, 0, 1$ in d spatial dimensions. The corresponding Ricci tensor is

$$\tilde{R}^k_m = \tilde{R}^{jk}_{jm} = K(d-1)\delta^k_m = \frac{\tilde{R}}{d}\delta^k_m \tag{314}$$

The spatial homogeneity makes $\tilde{R}$ constant within each spatial hyper-surface $\Sigma(t)$.

We shall now construct the metric on a spatial hyper-surface with homogeneous curvature. Spaces with positive curvature, $k = 1$, can be obtained by embedding into R^4,

$$x^2 + y^2 + z^2 + w^2 = a^2, \tag{315}$$

yielding

$$0 = xdx + ydy + zdz + wdw \tag{316}$$

and the induced metric

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 + dw^2 \\ &= dx^2 + dy^2 + dz^2 + \frac{(xdx + ydy + zdz)^2}{a^2 - x^2 - y^2 - z^2} \end{aligned} \tag{317}$$

written in polar coordinates as

$$\begin{aligned} ds^2 &= dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) + \frac{r^2 dr^2}{a^2 - r^2} \\ &= \frac{dr^2}{1 - \frac{r^2}{a^2}} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \end{aligned} \tag{318}$$

For negative curvature the embedding is given by

$$x^2 + y^2 + z^2 - w^2 = -a^2, \tag{319}$$

which yields

$$0 = xdx + ydy + zdz - wdw \tag{320}$$

and the induced metric is

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 - dw^2 \\ &= dx^2 + dy^2 + dz^2 - \frac{(xdx + ydy + zdz)^2}{a^2 + x^2 + y^2 + z^2}. \end{aligned} \tag{321}$$

We find in polar coordinates

$$\begin{aligned} ds^2 &= dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) - \frac{r^2 dr^2}{a^2 + r^2} \\ &= \frac{dr^2}{1 + \frac{r^2}{a^2}} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \end{aligned} \tag{322}$$

The spatially homogeneous and isotropic space-time provides time coordinate axis orthogonal to the space directions, therefore the four dimensional metric is of Robertson-Walker,

$$ds^2 = d\tau^2 - a^2(\tau)\left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{323}$$

where τ is the proper time measured by clocks in rest in the space-like hyper-surface, the coordinate r is made dimensionless by means of the scale factor $r \to a(\tau)r$ which is an arbitrary constant for flat space and $k = \mathrm{sign}K$. The three-space is of finite volume for positive curvature, $k = 1$ and infinite for $k = 0, -1$.

One may write the metric as

$$ds^2 = d\tau^2 - a^2(\tau)\left[d\chi^2 + h^2(\chi)(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{324}$$

with

$$r = h(\chi) = \begin{cases} \sin\chi & k = 1 \\ \chi & k = 0 \\ \sinh\chi & k = -1 \end{cases} \tag{325}$$

which justifies the use of the combination

$$r_{ph} = a(\tau)h(\chi) \tag{326}$$

as a cosmic distance parameter. Yet another useful form of the metric is

$$ds^2 = a^2(\eta)\left[d\eta^2 - d\chi^2 - h^2(\chi)(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{327}$$

where

$$\eta = \int \frac{d\tau}{a(\tau)}. \tag{328}$$

The metric tensor

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & -a^2(t)\tilde{g}_{ij} \end{pmatrix}, \qquad \tilde{g}_{ij} = \begin{pmatrix} \frac{1}{1-kr^2} & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}, \tag{329}$$

$\mu, \nu = (\tau, r, \theta, \phi)$,

$$\sqrt{-g} = a^3(\tau)\frac{r^2\sin\theta}{\sqrt{1-kr^2}} \tag{330}$$

yields the Christoffel symbols with two or more indices τ vanishing and the further non-vanishing components are

$$\Gamma^\tau_{ij} = \dot{a}a\tilde{g}_{ij}, \qquad \Gamma^i_{\tau j} = \frac{\dot{a}}{a}\tilde{g}^i_j, \qquad \Gamma^i_{jk} = \tilde{\Gamma}^i_{jk} = \frac{1}{2}\tilde{g}^{i\ell}(\partial_j\tilde{g}_{k\ell} + \partial_k\tilde{g}_{\ell j} - \partial_\ell\tilde{g}_{jk}) \tag{331}$$

with $\dot{a} = \frac{da}{d\tau}$. The Ricci tensor of the metric tensor $\tilde{g}$ is proportional to $\tilde{g}$ itself and the proportionality constant turns out to be after some calculation $k(d-1)$ for a d-dimensional homogeneous and isotropic space. Thus we have

$$\tilde{R}_{jm} = 2k\tilde{g}_{jm} \tag{332}$$

for the three-space and

$$R_{\mu\nu} = \begin{pmatrix} R_{00} & 0 \\ 0 & \tilde{R}_{jk} + \tilde{g}_{jk}(a\ddot{a} + 2\dot{a}^2) \end{pmatrix} = \begin{pmatrix} -3\frac{\ddot{a}}{a} & 0 \\ 0 & \tilde{g}_{jk}(a\ddot{a} + 2\dot{a}^2 + 2k) \end{pmatrix} \tag{333}$$

for the four dimensional space-time. The scalar curvature is

$$R = R_{00} - \frac{1}{a^2}\tilde{g}^{jk}R_{jk} = -6\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right). \tag{334}$$

C. Equation of motion

In order to find the Einstein equation we approximate the matter, averaged over long distances, as an ideal fluid which is rest in the cosmic coordinate system of the Robertson-Walker metric, ie. the energy-momentum tensor is given by

$$T^{\mu\nu} = (\rho c^2 + p)u^\mu u^\nu - pg^{\mu\nu}, \tag{335}$$

where u^μ is the four-velocity of the matter, the unit tangent vector field of the time-like congruence. The metric (323) is based on the time coordinate τ therefore $u^\mu = (1, 0, 0, 0)$. We have $T^\mu_\mu = \rho c^2 - 3p = 0$ for scale invariant case like EM radiation, $p = 0$ for matter at rest, like cosmic dust and $T^{\mu\nu} \sim g^{\mu\nu}$, $p = -\rho c^2$ in the vacuum. Notice that $\rho c^2 + 3p \geq 0$ in each case.

The first two terms in the divergence of a tensor $A^{\mu\nu}$

$$D_\nu A^{\mu\nu} = \partial_\nu A^{\mu\nu} + \Gamma^\nu_{\rho\nu} A^{\mu\rho} + \Gamma^\mu_{\rho\nu} A^{\rho\nu} \tag{336}$$

looks as the covariant divergence of a four-vector which can be written in a simpler manner according to (149),

$$D_\nu A^{\mu\nu} = \frac{1}{\sqrt{-g}} \partial_\nu (\sqrt{-g} A^{\mu\nu}) + \Gamma^\mu_{\rho\nu} A^{\rho\nu}. \tag{337}$$

Thus the expression (335) leads to the energy-momentum conservation law

$$0 = -\partial_\nu p g^{\mu\nu} + \frac{1}{\sqrt{-g}} \partial_\nu [\sqrt{-g}(\rho c^2 + p)u^\mu u^\nu] + \Gamma^\mu_{\rho\nu}(\rho c^2 + p)u^\rho u^\nu \tag{338}$$

where the metric admissibility, $Dg = 0$, was used, too. The rest frame condition, $u^\mu = (1, 0, 0, 0)$, renders the spatial components, $\mu = 1, 2, 3$ of this equation trivial and the temporal part $\mu = 0$ reads as

$$a^3 \dot{p} = \frac{d}{d\tau}[a^3(\rho c^2 + p)], \tag{339}$$

giving

$$0 = \frac{d}{d\tau}(a^3 \rho c^2), \tag{340}$$

$\rho \sim 1/a^3$ for dust. In the case of radiation we write

$$0 = \frac{4}{3}\frac{d}{d\tau}(a^3 \rho c^2) - \frac{1}{3} a^3 \dot{\rho} c^2 = \frac{1}{a}\frac{d}{d\tau}(a^4 \rho c^2), \tag{341}$$

resulting in $\rho \sim 1/a^4$. The density drops faster in the latter case during the expansion of the universe (growing a) than for dust. Though the radiation represents a negligible component in the actual universe, it was dominant in an earlier phase.

The Einstein equations read finally as

$$
\begin{aligned}
R_{00} - \frac{1}{2}R - \Lambda &= 3\frac{\dot{a}^2 + k}{a^2} - \Lambda \\
&= 8\pi G T_{\tau\tau} = 8\pi G \rho c^2
\end{aligned}
\tag{342}
$$

for the components 00 and

$$
\begin{aligned}
\frac{1}{a^2\tilde{g}_{rr}}\left[R_{rr} - \frac{1}{2}g_{rr}(R + 2\Lambda)\right] &= \frac{1}{a^2\tilde{g}_{rr}}\left[R_{rr} + \frac{a^2\tilde{g}_{rr}}{2}(R + 2\Lambda)\right] \\
&= \frac{\ddot{a}}{a} + 2\frac{\dot{a}^2}{a^2} + 2\frac{k}{a^2} - 3\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) + \Lambda \\
&= -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} - \frac{k}{a^2} + \Lambda \\
&= \frac{8\pi G}{a^2\tilde{g}_{rr}}T_{rr} = 8\pi G p
\end{aligned}
\tag{343}
$$

for rr.

We can express the acceleration $\ddot{a}$ by forming a suitable linear superposition of these two equations,

$$
\frac{\ddot{a}}{a} = \frac{\Lambda}{3} - \frac{4}{3}\pi G(3p + \rho c^2). \tag{344}
$$

The cosmological constant introduces a pressure in the absence of matter and leads to violation of the Newtonian gravitational law in the slow motion, weak gravitational field limit. We shall set $\Lambda = 0$ in the rest of the discussion for simplicity.

The first remark is that there is no static solution, $\ddot{a} < 0$, for $3p + \rho c^2 > 0$. The rate of change of spatial physical distances, $\ell_{ph} = \tilde{\ell}a$ with $\tilde{\ell}$ constant,

$$
v = \frac{d\ell_{ph}}{d\tau} = \ell_{ph}\frac{\dot{a}}{a} = H\ell_{ph}, \tag{345}
$$

where

$$
H = \frac{\dot{a}}{a} \tag{346}
$$

called Hubble-constant, though its value has slow time dependence on astrophysical time scale. It was supposed to be around 600km/s/Mpc according as proposed by Hubble in the '30s, its present value is around 70km/s/Mpc. Note that $v > c$ for large enough separation. This is not in contradiction with special relativity which considers velocities at the same space-time point but leads to the appearance of horizons as we shall see later.

The universe is expanding at the present, $\dot{a} > 0$ but in view of $\ddot{a} < 0$ the expansion rate must have been faster in the past. By assuming a constant expansion rate, $a(\tau) = \dot{a}(\tau_0)\tau$, where

$\dot{a}(\tau_0) = H(\tau_0)a(\tau_0)$, τ_0 being our time, the life-time of the Universe, we have $\tau_0 = 1/H$. Due to the slowing expansion rate the Big Bang must have occurred less time before and the inverse Hubble-constant gives only an order of magnitude estimate of the lifetime of the universe. The zero size signals a singularity in the time evolution which prevents us to inquire about the earlier state of the Universe. The so called singularity theorems of general relativity assures that the singularity at the Big Bang is present even without assuming homogeneity and isotropy. Naturally the classical equations of General Relativity do not allow us to inquire about the state of the Universe when its size was smaller than Planck's length.

For the flat or open universe, $k = 0$ or $k = -1$, respectively $\dot{a} \neq 0$ according to Eq. (342) which can be written as

$$\dot{a}^2 = \frac{8\pi G}{3} a^2 \rho c^2 - k \tag{347}$$

and the expansion continues forever. In fact, $\rho c^2 = \mathcal{O}\left(a^{-3}\right)$ or $\rho c^2 = \mathcal{O}\left(a^{-4}\right)$ for dust or radiation dominated universe, $\rho c^2 a^2 \to 0$ as $\tau \to \infty$ and $\dot{a}$ approaches zero from above. For closed universe, $k = 1$, the matter contribution to Eq. (347) decreases compared to k during the expansion and there is a maximal value of a, $a \leq a_0$. But the maximal value can not be approached asymptotically because $\ddot{a}$ does not tend to zero according to Eq. (344) but instead a big crunch occurs at some finite time where $a = 0$ is reached and the universe ceases to exists.

The 00 component of the Einstein equation (342) for $\Lambda = 0$ shows that the universe is closed or open if $\rho > \rho_c$ or $\rho < \rho_c$, respectively where

$$\rho_c c^2 = \frac{3H^2}{8\pi G}. \tag{348}$$

The actual observational and theoretical background suggests that the cosmological constant Λ actually plays an important role in determining the age of the universe, in particular the choice $\rho_{matter} \approx 0.27\rho_c$, $\rho_\Lambda \approx 0.73\rho_c$, $\rho_{matter} + \rho_\Lambda \approx \rho_c$ is preferred.

D. Frequency shifts

The precision spectroscopic measurement of the electromagnetic radiation from stars is an easy way of collecting information about distant astrophysical objects. Three different mechanisms are known to change the observed value of the characteristic frequency of a physical system. The first is observed in flat space-time when the source of a wave and its observer moves with respect to each other. The second and the the third mechanisms are for sources and observers at rest but in the presence of a static or time-dependent gravitational field, respectively.

1. Let us consider a monochromatic plane wave with wave vector $k^\mu = (\omega_0/c, \mathbf{k})$ with $k^2 = m^2c^2/\hbar^2$ whose source is moving with velocity $\mathbf{v}$ with respect to an observer in flat space-time. The time component of the wave vector in the observer's reference frame is

$$\omega = \frac{\omega_0 - \mathbf{vk}}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}}. \tag{349}$$

By means of the relation $|\mathbf{k}| = \sqrt{\frac{\omega_0^2}{c^2} - k^2}$ we write

$$\mathbf{vk} = v\frac{\omega_0}{c}\sqrt{1 - \frac{m^2c^4}{\hbar^2\omega_0^2}}\cos\theta \tag{350}$$

and find

$$\omega = \omega_0 \frac{1 - \frac{v}{c}\sqrt{1 - \frac{m^2c^4}{\hbar^2\omega_0^2}}\cos\theta}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} \tag{351}$$

This is the relativistic Doppler effect, its non-relativistic analogy for a simple wave propagating with speed $\tilde{c}$ with $v, \tilde{c} \ll c$ is

$$\omega = \omega_0 \left(1 - \frac{v}{\tilde{c}}\cos\theta\right). \tag{352}$$

The Doppler shift is particularly useful in the observation of the broadening of spectral lines of light arriving from stars because it allows us to estimate the random velocity of the atoms or molecules and the temperature on the surface of the star.

2. Time independent gravitational field changes the frequency, being proportional with the energy. Let us consider two electron-positron pairs at rest far from a star of mass M. One pair annihilates into a photon and the other starts to fall freely towards the star. Let us suppose that at distance r form the star the second pair annihilates into a photon. The photon arising from the first pair, at large distance form the star, is of frequency $\omega_1(\infty) = E_k(\infty)/\hbar = 2m_0c^2/\hbar$. The photon which is created at distance r by the second pair is of the frequency $\omega_2(r) = E_k(r)/\hbar$ where

$$E_k = 2\left(m_0c^2 + \frac{Gm_0M}{r}\right) = 2m_0c^2\left(1 + \frac{GM}{rc^2}\right) \tag{353}$$

is the kinetic energy of the pair at the instance of the annihilation. Let us suppose that we send the photon arising from the annihilation of the first pair at the location where the annihilation of the second pair took place by means of static mirrors which leave the photon frequency unchanged and denote its frequency $\omega_1(r)$ there. Energy conservation requires $\omega_1(r) = \omega_2(r)$ since otherwise one can extract energy form static gravitational field by annihilation of electron-positron pairs at certain distance form the star, sending the photons arising from the process at a different distance and recombining them into electron-positron pair again. Thus the frequency of the photon changes as

$$\omega(r) = \omega(\infty)\left(1 + \frac{GM}{rc^2}\right) \tag{354}$$

due to the presence of the gravitational field. The relative change,

$$\frac{\Delta\omega}{\omega} = \frac{\omega(r) - \omega(\infty)}{\omega(\infty)} = \frac{GM}{rc^2} \tag{355}$$

is approximately $2 \cdot 10^{-6}$ at the surface of the Sun but becomes comparable to one on the surface of a neutron star and can be used to estimate the total mass.

3. Time dependent gravitational field leads the the change of frequency in a trivial manner because the time of emission of a photon is different than the time of its absorption. Let us suppose that a light signal is emitted at the point $p_0 = (ct_e, \mathbf{r})$ which is received at $p = (ct_o, \mathbf{0})$. The propagation is along a null-geodesic,

$$cdt = a(ct)d\chi, \tag{356}$$

yielding

$$\int_{t_e}^{t_o} \frac{cdt}{a(ct)} = \chi. \tag{357}$$

The source emits $dn = dt_e\omega_e$ periods during the time interval dt_e of the emission with frequency ω_e. It is the same number of periods, $dn = dt_o\omega_o$ which is observed in time dt_o and frequency ω_o. Thus we have red shift parameter, the relative change of the wavelength $\lambda = 2\pi c/\omega$,

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} = \frac{dt_o}{dt_e} - 1 = \frac{a(ct_o)}{a(ct_e)} - 1 \tag{358}$$

where the difference of Eq. (357) and its analogy written for the times $t_e \to t_e + dt_e$, $t_o \to t_o + dt_o$ was used in the last equation. By assuming that the scale factor $a(\tau)$ changes slowly in time we find

$$\begin{aligned} z &= \frac{a_o}{a_o + (ct_e - ct_o)\dot{a}_o + \frac{1}{2}(ct_e - ct_o)^2\ddot{a}_o + \cdots} - 1 \\ &= \frac{\dot{a}_o}{a_o}(ct_o - ct_e) + \left[\left(\frac{\dot{a}_o}{a_o}\right)^2 - \frac{\ddot{a}_o}{2a_o}\right](ct_o - ct_e)^2 + \cdots \end{aligned} \tag{359}$$

where $a_o = a(ct_o)$. The deceleration parameter,

$$q = -\frac{a\ddot{a}}{\dot{a}^2} \tag{360}$$

measures the rate of change of the Hubble-constant in time and its experimental value is about 1. Hubble's law, the linearity of the red shift in the distance is obtained by

$$z \approx \frac{\dot{a}_o}{a_o}(ct_o - ct_e) \approx \frac{\dot{a}_o}{a_o}\ell = H\ell. \tag{361}$$

As a simple application let us now consider the absolute luminosity $\mathcal{L}$ of a galaxy and the measured flux (energy per unit time and unit area) $\mathcal{F}$. The quantity

$$d_L = \sqrt{\frac{\mathcal{L}}{4\pi\mathcal{F}}} \tag{362}$$

can be interpreted as the luminosity distance of the galaxy. But due to the expansion of the universe the distance of the galaxy at the time of the emission of the observed signal, r_e was different than r_o, the distance at the time of observation. Does the distance d_L agree one of them? Not. Energy conservation requires

$$\mathcal{F} = \frac{\mathcal{L}}{4\pi a^2(ct_o)r_e^2} = \frac{\mathcal{L}}{4\pi a^2(ct_o)r_e^2(1+z)^2}, \tag{363}$$

where the factor $dA/4\pi a^2(ct_o)r_e^2$ is the solid angle of a surface dA observed, one factor of $1+z$ arises from the red shift of the observed photons and the other from the ratio of the time intervals dt_e/dt_o and we find

$$d_L = a(ct_o)r_e(1+z). \tag{364}$$

E. Particle horizon

Can we receive signals from any part of the universe? It is true that the universe was smaller in the past but the speed of light appeared smaller, too. This is the kind of dilemma which led to Eq. (362) and requires more careful consideration. We now use the form (327) of the metric and look for the spatial region form which signals can be received in the case of flat geometry, $k = 0$. The metric is now conformally flat, we can replace the overall conformal factor $a^2(\eta)$ by one from the point of view of this question and the problem is reduced to flat Minkowski geometry. Obviously if the integral

$$\int_\epsilon^\tau \frac{d\tau'}{a(\tau')} \tag{365}$$

remains finite when $\epsilon \to 0$ then we can not receive signals from the whole universe and there is a particle horizon. The short time solution of the Einstein equations give $a(\tau) = \mathcal{O}\left(\tau^{2/3}\right)$ for $\tau \approx 0$ even for a dust dominated universe and the particle horizon appears. For $k = -1$ the curvature term becomes negligible for short time and the same result is recovered. The situation is more complicated for a closed universe and it turns out that the horizon disappears when the universe reaches its maximal size for a dust domination but remains present for all times in the radiation dominated case.

The presence of horizons raises serious problems for the description of the evolution of the universe. The reason is that the difference between regions of the early universe which are separated by horizon can not relax during the evolution and any inhomogeneity which appeared at such scale in the very early universe should be observable today. The cosmic microwave background which is supposed to originate from a rather early period of the universe is found to be homogeneous and isotropic to such a large extent which is not possible to understand unless one assumes that either the universe is created is an unusually homogeneous state or the Robertson-Walker metric is not valid for the early phase. The inflationary universe model which involves a rapid expansion in the early phase leads to an enlargement of the horizon and offers a possibility to reconcile the entropy estimates of the early universe with the present homogeneity of the microwave background radiation.

But the latest, more sensitive measurements of the anisotropy and inhomogeneities of the microwave background radiation lead to another serious problem, these fluctuations are apparently too large to be explained by the inflationary standard model. Further problems arise from the more precisely determined slowing down of the expansion of the universe. That phenomenon might be explained by assuming that the majority of the matter in our universe is participating in gravitational interaction only and pulls back the expanding, visible matter. But such a rescue operation of the Einstein equation seems to be out of proportion and the slight modification of the Einstein-Hilbert action appears to be a more economical and better justifiable way of establishing consistency.

F. Evolution of the universe

We shall briefly review the salient feature of the evolutionary big bang model. It is clearly unreasonable to expect that classical physics, in particularly general relativity as we know it today is capable to trace the evolution just from the beginning, from $t = 0$, $a = 0$. The basic physical constants can be put together to form a length scale,

$$\ell_P = \sqrt{\frac{G\hbar}{c^3}} \approx 10^{-33} \text{ cm}, \tag{366}$$

called Planck length. At distances smaller than this scale the gravitational interaction should be stronger than the quantum effects. In general, one should not extrapolate the physical laws across such a wast regime of scale what separate the length scale 10^{-16}cm, the space resolution of the the experiments today, from the Planck length because the basic structure of the physical laws changed in the last century several time by the discovery of new interactions or particles as the observational length scale was reduced from the mm scale to the proton size, 10^{-13}cm. It is a real surprise that the structure of quantum mechanics found to be valid in this range when all other laws in physics proved to be limited to longer length scales.

The first question one might raise after leaving the **Planck-era** ($t \approx 10^{-43}$s, $T \approx 10^{31}$K, $\rho \approx 10^{92}$ gm/cm^3) is whether thermodynamical equilibrium was reached by the universe. The inflation, the rapid expansion of the universe can provide such a thermalization by the enlargement of the horizons. The background microwave radiation is the indication that thermal equilibrium was reached by a hot universe. According to Eqs. (340) and (341) the early universe was dominated by radiation.

Another interesting issue of this period is the generation of matter-anti matter domains. Our galaxy tend to have matter rather than anti-matter, the baryon number (+1 for a nucleon and -1 for an anti-nucleon) its total baryon number, a conserved quantity, is positive. If the universe was created by positive net baryon number then this is not surprising. But it is more natural to imagine that the universe started its existence with vanishing conserved charges and the matter excess observed around us is compensated by an anti-matter excess somewhere else in the universe. What was the mechanism which created these matter-anti matter domains? There is no generally accepted and satisfactory answer.

The following milestones should be mentioned for the subsequent evolution:

1. At the beginning after the Planck-era the non-gravitational interactions are supposed to be unified by the **Grand Unified Model**. At about $t \approx 10^{-36}$s the symmetry of the Grand Unified Model is broken spontaneously, the strong and the electro-weak interactions separate into Quantum Chromodyanmics and the unified electro-weak theory, respectively and the matter-anti matter island are supposed to be formed.

2. At about $t \approx 10^{-12}$s the symmetry of the **unified electro-weak theory** is broken spontaneously and electromagnetic and weak forces separate.

3. At $t \approx 10^{-5}$s quarks which were freely propagating became confined and the present day **hadrons are formed**.

4. At $t \approx 1$s the universe consists of mainly neutrinos, photons, electrons, protons and neutrons and their anti-particles were in thermal equilibrium up to now but the interaction with neutrinos becomes weak to maintain equilibrium from now on. As a result **neutrinos decouple** and follow a passive red shift in the rest of the time. Soon after the **proton-neutron conversion is frozen** out ($T = 10^{10}$K corresponds approximately to 1MeV, the neutron-proton mass difference), too.

5. At $t \approx 4$s the **electron-positron equilibrium is lost** because $T = T \approx 5 \cdot 10^9$K belongs to 0.5MeV, the mass of the electron. The annihilation process eliminate positrons and heat up the photons slightly.

6. At $10s < t < 10min$ the thermal energy reaches the nuclear binding scale and **nucleosynthesis** starts by producing ^{4}He nuclei. The strong Coulomb barrier and the lack of other stable elements with $Z < 8$ leave helium the only nuclei produced in mass, until all neutrons left over after step 4. are bound into helium. At the end of the helium dominated era which lasts few minutes 25% of the mass is in the form of ^{4}He, the rest is essential distributed over ^{2}H, ^{3}He and ^{7}Li. The reproduction of the fraction taken in the form of ^{4}He is a convincing success of the hot big bang nucleosynthesis model.

7. At $t \approx 4{\cdot}10^5$ year the **matter reaches equilibrium with radiation** and starts to dominate.

8. At $t \approx 10^6$ year the thermal energy reaches the ionization energy of the hydrogen atom and we enter in the recombination era when stable, **neutral atoms**, starting with the hydrogen are formed. Most of the universe becoming electrically neutral the photons decouple. Their subsequent expansion and cooling leads to the actual temperature 2.7K, observed with high accuracy in the **cosmic microwave background** in the last decades. This is the second decisive victory of the hot big bang model for the universe.

The decoupling of matter and radiation signals the end of the **quantum-driven evolution** era. The resulting loss of radiation pressure leads to gravitational instabilities for masses $M > 10^5 M_{sun}$ where $M_{sun} \approx 2 \cdot 10^{33}$g is the solar mass and **galaxies** start to be formed. About $t \approx 10^3 - 10^7$years hadronic matter started to dominate the total energy of the universe whose present age is currently believed to be $14 \cdot 10^9$years.

VIII. OPEN QUESTIONS

General Relativity has appeared a century ago as a radically new vision of interactions. In addition the experimental texts of gravity are quite challenging. Hence there is an unusually large set of open question left for future studies. Only the most general issues are mentioned here, without any attempt to be exhaustive.

1. *What is the action of General Relativity?* The physical phenomenons and laws depend on the scale of their observations. By using the units $c = \hbar = 1$ one may use the length as a single scale. All of our theories are effective, they cover given length scale windows. There are always terms in the Lagrangian which represent the physics beyond the scale window but they are negligible in the range of application of the theory in question. However these terms become more important as we reach the edge of the scale window. The lesson of this feature is that any experimentally established Lagrangian remains open to corrections, demanded by later observations.

   Gravity has the widest window of applicability among the four fundamental interactions with the inconvenience that the scales where it is supposed to be strong are extremely large (size of the Universe) or small (Schwarzschild radius, Plank's scale), far away form us. There are however experimentally confirmed phenomenons requiring some modifications of the traditional Einstein-Hilbert action:

   - *Galactical scale:* The rotation of and the velocity distribution within a galaxy indicates stronger gravitational force than expected. The currently followed strategy to incorporate such changes is to introduce dark matter, a matter which enters only into gravitational interactions. The elimination of the dynamical degrees of freedom of dark matter, in a manner similar to the application of the renormalization group method, generates new terms to the gravitational action. It is more natural and promising to address the changes of the action directly, without restricting the attention to a particular, experimentally inaccessible sector of physics.

   - *Cosmological scale:* The observed expansion of the Universe suggests that matter density is about the critical value, corresponding to the flat solution of the Robertson-Walker model. The known matter can not make up the critical density and the deficit in the cosmological constant is called dark energy contribution. Whatever is its physical realization, the dark energy is another extension of the traditional action.

- *Consistency:* The different modifications of the action must be consistent. The value of the Hubble constant, inferred from observing some nearby galaxies, $H = 73.5 \pm 0.5$km/sec/Mpc, and extracting from the inhomogeneity of the cosmic microwave background at Galaxy scale, $H = 67.4 \pm 1.4$km/sec/Mpc, are inconsistent and requires a more thorough approach to changing the gravitational action.

2. *Is gravity quantized?* The four fundamental interactions seem to belong to the class of gauge theories and the other three are quantized. While there is neither experimental nor theoretical evidence that gravity needs quantization one tends to apply the same scheme for all fundamental interactions and inquire about the possibility of quantizing gravity. However the application of the usual quantization procedure creates apparently insurmountable difficulties.

   - The properties exists only after observation according to Quantum Mechanics which assumes the existence of the macroscopic world, in particular the availability of space and time. How to approach the problem of generating the space-time coordinates by observations, more precisely by the interaction of the space-time coordinates with a sufficiently large environment, to generate the classical space-time, the carrier of the quantization procedure of the other three fundamental interactions? What kind of mathematical basis should we use for the emergence of the space-time coordinates?
   - The operator mixing, the presence of higher than second order products of the coordinate and the momentum operator in the Hamiltonian generates short time divergences in Quantum Mechanics. Quantum field theories contain ultraviolet divergences due to the unboundedness of the three-momentum. The lesson is that quantum mechanics and the continuity of the space-time are incompatible. This problem is handled by the introduction of a cutoff, reflecting our ignorance of physics at infinitesimal length scales. This ignorance is build in into the theories by assuming no dynamics at shorter length scale than the cutoff. Nevertheless there is a historic preference of renormalizable theories where the cutoff can be sent to zero in such a manner that an appropriately chosen cutoff-dependence of the parameters of theory renders the predictions of the theory convergent. The renormalizability allows the hiding of the cutoff scale, an irrelevant parameter from the point of view of physical phenomenons taking place at finite scales, in our equations. This is a matter of convenience rather than necessity.

Nevertheless the question is there: Is gravity renormalizable? There are indications of a non-perturbative renormalization in case of a slightly extended Einstein-Hilbert action. Another possibility is to use string theory if one is ready to go beyond four dimensions and the field theory framework to reach this goal.

3. *Impact of gravity on thermodynamics:* The thermodynamical laws are modified at short and long distances by gravitational interactions. The absence of gravitational screening leaves Newton's gravitational force long range and creates difficulties in deriving the thermodynamical laws in Statistical Mechanics. The one-way causal structure of event horizons introduces an information loss for processes taking place close to the horizon. Entropy is related to the missing information hence event horizons generate a particular contribution to the thermodynamical potentials. This latter question has thoroughly been studied analytically but there is no experimental support in either issue.

Appendix A: Uniformly accelerating observer

If inertial and gravitational forces are identical then a uniformly accelerating observer should experience a homogeneous gravitational potential and it is instructive to look more closely in this case. Consider the first motion in the spatial x direction where the world line $x^\mu(s)$,

$$x^\mu = (t, x) = \frac{1}{a}(\sinh a\tau, \rho - 1 + \cosh a\tau), \tag{A1}$$

s being the invariant length, gives rise the four-velocity $u^\mu(s) = \frac{d}{ds}x^\mu(s) = \dot{x}^\mu(s)$ which satisfies $u^2 = 1$ and the four acceleration $a^\mu(s) = \ddot{x}^\mu(s)$ with constant invariant length, $a^2 = -a_0^2$. The four-velocity, written as $u^\mu = (\cosh f(s), \sinh f(s), 0, 0)$, gives the desired acceleration with the choice $f(s) = as$. An integration of the velocity produces the world line

$$x^\mu = \frac{1}{a}(\sinh as, \cosh as, 0, 0), \tag{A2}$$

a hyperbole with positive value of the x coordinate. It is advantageous to introduce a coordinate system $(t, x) \to (\eta, \rho)$, where the hyperbolas are labeled by the spatial coordinate ρ and η is proportional to the proper time,

$$x^\mu = \rho(\sinh \eta, \cosh \eta, 0, 0), \tag{A3}$$

where the proper time and the acceleration of the world lines are $s = \eta/\rho$ and $a = 1/\rho$, respectively and the invariant distance can be written as

$$ds^2 = \rho^2 d\eta^2 - d\rho^2. \tag{A4}$$

The following features of the Rindler geometry are noteworthy:

1. The Rindler space covers the part $x \geq 0$ of the whole Minkowski geometry,
2. The geometry, described by the metric (A4) is flat, it is a reparametrization of the Minkowski space-time.
3. The acceleration along a given world line is constant but different world lines display different acceleration hence the gravitational field is static but spatially inhomogeneous.
4. There are gravitational effects (inertial forces) even in flat space-time.
5. There is no contradiction with the Equivalence Principle because the accelerating world lines fill up a part of the space-time only and no statement is made about the flatness of the rest.
6. Gravitational forces seem to generate space-time dependent metric tensor.
7. Gravitational forces seem to produce singularity whose more precise nature, namely whether it is a coordinate singularity due to the wrong choice of coordinates or it is a real singularity, reflected by physical quantities, remains to be clarified.
8. Gravitational forces seem to make a part of the space-time inaccessible, they can generate a horizon.

Appendix B: Continuous groups

The continuous groups, $\{\omega(\alpha)\}$, are equipped with a continuous topology and the group multiplication law,

$$\omega(\alpha)\omega(\beta) = \omega(F(\alpha, \beta)) \tag{B1}$$

where α and β are n-dimensional vectors for an n-dimensional group and the function $F(\alpha, \beta)$ describes the multiplication law. For instance the translations in space-time make up the group $\{\omega(\alpha)\}$, α being a four vector $\alpha = x^\mu$, with $F(\alpha, \beta) = \alpha + \beta$. The widely followed convention, adopted here as well, is the choose the unit element at $\alpha = 0$, $\omega(0) = \mathbb{1}$.

The classification of the possible continuous symmetry groups is made simple by Ado's theorem asserting that any finite dimensional Lie-algebra is identical with a subspace of the generators of the matrix group $GL(N)$ (GL=General Linear group), consisting of non-singular $N \times N$ real matrices, for sufficiently large N. Thus any continuous group is locally identical with a subgroup of $GL(N)$ for certain N. In order to cover all continuous groups it is sufficient to study the matrix groups. The important matrix groups are called classical matrix groups and are shown in Tables I and II. The rest, the exceptional groups have not yet found application in physics.

The infinitesimal group elements are in the vicinity of the identity,

$$\omega = \mathbb{1} + \sum_{n=1}^{n} \epsilon^a \tau^a + \mathcal{O}\left(\epsilon^2\right), \tag{B2}$$

with $\tau^a = \frac{\partial\omega(0)}{\partial\alpha^a}$, called generators. The name originates from the possibility of enlarging this structure over the whole matrix group by the exponential map, defined by the help of the equation

$$\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = \lim_{n\to\infty} e^{n\ln(1+\frac{a}{n})} = \lim_{n\to\infty} e^{n(\frac{a}{n}+\mathcal{O}(n^{-2}))} = e^a, \tag{B3}$$

valid for any finite number a. The choice $a = \sum_a \alpha^a\tau^a$ yields

$$e^{\sum_a \alpha^a\tau^a} = \lim_{n\to\infty}\left(1+\sum_a \frac{\alpha^a}{n}\tau^a\right)^n, \tag{B4}$$

showing that the repeating of an appropriately chosen infinitesimal transformation many times generates the finite group element. One can be shown that any group element can be obtained in such a form in a connected group. One can also be shown that the commutator of generators is also a generator,

$$[\tau^a, \tau^b] = \sum_c f^{a,b,c}\tau^c. \tag{B5}$$

The real numbers $f^{a,b,c}$ are called structure constants and they uniquely determine the multiplication of the infinitesimal group elements.

Appendix C: Classical Field theory

The goal of this Appendix is a brief introduction to variational principle and conservation laws in classical field theory. The reason to go beyond the usual, differential equation based definition of the dynamics is have equations of motion which preserve their form under arbitrary transformation of the space-time coordinates.

TABLE I: Real classical matrix groups.

Symbol	Name	Definition	Dimension	Generators
$GL(N)$	general linear group	$\det A \neq 0^a$	N^2	$\{\tau : \text{real } N \times N \text{ matrices}\}$
$SL(N)$	special linear group	$\det A = 1$	$N^2 - 1$	$\mathrm{tr}\tau = 0^b$
$O(N)$	orthogonal group	$A^{tr} A = 1\!\!1^c$	$\frac{1}{2}N(N-1)$	$\tau^{tr} = -\tau$
$SO(N)$	special orthogonal group	$A^{tr} A = 1\!\!1$, $\det A = 1$	$\frac{1}{2}N(N-1)$	$\tau^{tr} = -\tau$, $\mathrm{tr}\tau = 0$

[a]The matrix A is supposed to be an element of the group in question.
[b]$\det(1\!\!1 + \epsilon\tau) = 1 + \epsilon \mathrm{tr}\tau + \mathcal{O}\left(\epsilon^2\right)$
[c]$\det A^{tr} A = (\det A)^2 = 1$ and $\det A = \pm 1$.

TABLE II: Complex classical matrix groups.

Symbol	Name	Definition	Dimension	Generators
$GL(N,C)$	complex general linear group	$\det A \neq 0$	$2N^2$	$\{\tau : \text{complex } N \times N \text{ matrices}\}$
$SL(N,C)$	complex special linear group	$\det A = 1$	$2N^2 - 2$	$\mathrm{tr}\tau = 0$
$U(N)$	unitary group	$A^\dagger A = 1\!\!1^a$	N^2	$\tau^\dagger = -\tau$
$SU(N)$	special unitary group	$A^\dagger A = 1\!\!1$, $\det A = 1$	$N^2 - 1$	$\tau^\dagger = -\tau$, $\mathrm{tr}\tau = 0$

[a]$\det A^\dagger A = (\det A)^* \det A = |\det A|^2 = 1$

1. Variational principle

Field theory is a dynamical system containing degrees of freedom, denoted by $\phi(\mathbf{x})$, at each space point $\mathbf{x}$. The coordinate $\phi(\mathbf{x})$ can be a single real number (real scalar field) or consist n-components (n-component field). Our goal is to provide an equation satisfied by the trajectory $\phi_{cl}(t, \mathbf{x})$. The index cl is supposed to remind us that this trajectory is the solution of a classical equation of motion. The problem of identifying $\phi_{cl}(t, \mathbf{x})$ will be outlined in three steps.

a. Single point on the real axis

Problem: identification of a point on the real axis, $x_{cl} \in \mathbb{R}$, in a manner which is independent of the reparametrization of the real axis.

Solution: Find a function with vanishing derivative at x_{cl} only:

$$\frac{df(x)}{dx}_{|x=x_{cl}} = 0 \tag{C1}$$

To check the reparametrization invariance of this equation we introduce new coordinate y by the function $x = x(y)$ and find

$$\frac{df(x(y))}{dy}_{|y=y_{cl}} = \underbrace{\frac{df(x)}{dx}_{|x=x_{cl}}}_{0} \frac{dx(y)}{dy}_{|y=y_{cl}} = 0 \tag{C2}$$

Variational principle: There is simple way of rewriting Eq. (C1). Let us perform an infinitesimal variation of the coordinate $x \to x + \delta x$, and write

$$\begin{aligned} f(x_{cl} + \delta x) &= f(x_{cl}) + \delta f(x_{cl}) \\ &= f(x_{cl}) + \delta x \underbrace{f'(x_{cl})}_{0} + \frac{\delta x^2}{2} f''(x_{cl}) + \mathcal{O}\left(\delta x^3\right) \end{aligned} \tag{C3}$$

The variation principle, equivalent of Eq. (C1) is

$$\delta f(x_{cl}) = \mathcal{O}\left(\delta x^2\right), \tag{C4}$$

stating that x_{cl} is characterized by the property that an infinitesimal variation around it, $x_{cl} \to x_{cl} + \delta x$, induces an $\mathcal{O}\left(\delta x^2\right)$ change in the value of $f(x_{cl})$.

b. *Non-relativistic point particle*

Problem: identification of a trajectory in a coordinate choice independent manner.

Variational principle: Let us identify a trajectory $x_{cl}(t)$ by specifying the coordinate at the initial and final time, $x_{cl}(t_i) = x_i$, $x_{cl}(t_f) = x_f$ (by assuming that the equation of motion is of second order in time derivatives) and consider a variation of the trajectory $x(t)$: $x(t) \to x(t) + \delta x(t)$ which leaves the initial and final conditions invariant (ie. does not modify the solution). Our function $f(x)$ of the previous section becomes a functional, called action

$$S[x(\cdot)] = \int_{t_i}^{t_f} dt L(x(t), \dot{x}(t)) \tag{C5}$$

involving the Lagrangian $L(x(t), \dot{x}(t))$. (The symbol $x(\cdot)$ in the argument of the action functional is supposed to remind us that the variable of the functional is a function. It is better to put a dot in the place of the independent variable of the function $x(t)$ otherwise the notation $S[x(t)]$ can be mistaken with an embedded function $S(x(t))$.) The variation of the action is

$$\begin{aligned}
\delta S[x(\cdot)] &= \int_{t_i}^{t_f} dtL\left(x(t)+\delta x(t), \dot{x}(t)+\frac{d}{dt}\delta x(t)\right) - \int_{t_i}^{t_f} dtL(x(t),\dot{x}(t)) \\
&= \int_{t_i}^{t_f} dt\bigg[L(x(t),\dot{x}(t)) + \delta x(t)\frac{\partial L(x(t),\dot{x}(t))}{\partial x} + \frac{d}{dt}\delta x(t)\frac{\partial L(x(t),\dot{x}(t))}{\partial \dot{x}} + \mathcal{O}\left(\delta x(t)^2\right) \\
&\quad - \int_{t_i}^{t_f} dtL(x(t),\dot{x}(t))\bigg] \\
&= \int_{t_i}^{t_f} dt\delta x(t)\left[\frac{\partial L(x(t),\dot{x}(t))}{\partial x} - \frac{d}{dt}\frac{\partial L(x(t),\dot{x}(t))}{\partial \dot{x}}\right] + \underbrace{\delta x(t)}_{0}\frac{\partial L(x(t),\dot{x}(t))}{\partial \dot{x}}\bigg|_{t_f}^{t_i} + \mathcal{O}\left(\delta x(t)^2\right)
\end{aligned} \tag{C6}$$

The variational principle amounts to the suppression of the integral in the last line for an arbitrary variation, yielding the Euler-Lagrange equation:

$$\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} = 0 \tag{C7}$$

The generalization of the previous steps for a n-dimensional particle gives

$$\frac{\partial L(\mathbf{x},\dot{\mathbf{x}})}{\partial \mathbf{x}} - \frac{d}{dt}\frac{\partial L(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{\mathbf{x}}} = 0. \tag{C8}$$

It is easy to check that the Lagrangian

$$L = T - U = \frac{m}{2}\dot{\mathbf{x}}^2 - U(\mathbf{x}) \tag{C9}$$

leads to the usual Newton equation

$$m\ddot{\mathbf{x}} = -\nabla U(\mathbf{x}). \tag{C10}$$

It is advantageous to introduce the generalized momentum:

$$p = \frac{\partial L(x,\dot{x})}{\partial \dot{x}} \tag{C11}$$

which allows to write the Euler-Lagrange equation as

$$\dot{p} = \frac{\partial L(x,\dot{x})}{\partial x} \tag{C12}$$

The coordinate not appearing in the Lagrangian in an explicit manner is called cyclic coordinate,

$$\frac{\partial L(x,\dot{x})}{\partial x_{cycl}} = 0. \tag{C13}$$

Noether's therem, discussed below folllows from observation that each cyclic coordinate generates a conserved quantity. In fact, the generalized momentum of a cyclic coordinate, p_{cycl}, is conserved according to Eqs. (C11) and (C13).

c. Scalar field

Problem: identification of the equation of motion for an n-component field, $\phi_a(x)$, $a = 1, \ldots, n$. (Notation: $x = (t, \mathbf{x})$.)

Variational principle: let us consider a variation of the trajectory $\phi(x)$:

$$\phi(x) \to \phi(x) + \delta\phi(x), \qquad \delta\phi(t_i, \mathbf{x}) = \delta\phi(t_f, \mathbf{x}) = 0. \tag{C14}$$

The variation of the action

$$S[\phi(\cdot)] = \int_V dt d^3x L(\phi, \partial\phi) \tag{C15}$$

is

$$\begin{aligned}
\delta S &= \int_V dt d^3x \left(\frac{\partial L(\phi, \partial\phi)}{\partial \phi_a} \delta\phi_a + \frac{\partial L(\phi, \partial\phi)}{\partial \partial_\mu \phi_a} \delta\partial_\mu \phi_a \right) + \mathcal{O}\left(\delta^2\phi\right) \\
&= \int_V dt d^3x \left(\frac{\partial L(\phi, \partial\phi)}{\partial \phi_a} \delta\phi_a + \frac{\partial L(\phi, \partial\phi)}{\partial \partial_\mu \phi_a} \partial_\mu \delta\phi_a \right) + \mathcal{O}\left(\delta^2\phi\right) \\
&= \int_{\partial V} ds^\mu \delta\phi_a \frac{\partial L(\phi, \partial\phi)}{\partial \partial_\mu \phi_a} + \int_V dt d^3x \delta\phi_a \left(\frac{\partial L(\phi, \partial\phi)}{\partial \phi_a} - \partial_\mu \frac{\partial L(\phi, \partial\phi)}{\partial \partial_\mu \phi_a} \right) + \mathcal{O}\left(\delta^2\phi\right)
\end{aligned} \tag{C16}$$

The first term for $\mu = 0$,

$$\int_{\partial V} ds^0 \delta\phi_a \frac{\partial L(\phi, \partial\phi)}{\partial \partial_0 \phi_a} = \int_{t=t_f} d^3x \underbrace{\delta\phi_a}_{0} \frac{\partial L(\phi, \partial\phi)}{\partial \partial_0 \phi_a} - \int_{t=t_i} d^3x \underbrace{\delta\phi_a}_{0} \frac{\partial L(\phi, \partial\phi)}{\partial \partial_0 \phi_a} = 0 \tag{C17}$$

is vanishing because there is no variation at the initial and final time. When $\mu = j$ then

$$\int_{\partial V} ds^j \delta\phi_a \frac{\partial L(\phi, \partial\phi)}{\partial \partial_j \phi_a} = \int_{x_j=\infty} ds^j \delta\phi_a \underbrace{\frac{\partial L(\phi, \partial\phi)}{\partial \partial_j \phi_a}}_{0} - \int_{x_j=-\infty} ds^j \delta\phi_a \underbrace{\frac{\partial L(\phi, \partial\phi)}{\partial \partial_j \phi_a}}_{0} = 0 \tag{C18}$$

and it is still vanishing because we are interested in the dynamics of localized systems and the interactions are supposed to be short ranged. Therefore, $\phi = 0$ at the spatial infinites and the Lagrangian is vanishing. The suppression of the second term gives the Euler-Lagrange equation

$$\frac{\partial L(\phi, \partial\phi)}{\partial \phi_a} - \partial_\mu \frac{\partial L(\phi, \partial\phi)}{\partial \partial_\mu \phi_a} = 0. \tag{C19}$$

The generalized momentum of a particle, $\partial L / \partial_t x$, can be generalized in field theory to the current, $j^\mu_\phi = \partial L / \partial \partial \phi_\mu$, associated to the field ϕ. A field variable is cyclic when it does not appear in the Lagrangian, $\partial L / \partial \phi_{cycl} = 0$. The Noether theorem for field theory stems from the conservation of the current, associated to a cyclic field variable, $\partial_\mu j^\mu_{\phi_{cycl}} = 0$.

Example: Scalar field ($\hbar = c = 1$):

$$L = \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\phi^2 - U(\phi) \quad \Longrightarrow \quad (\partial_\mu\partial^\mu + m^2) = -U'(\phi) \tag{C20}$$

2. Noether theorem

The reparametrization invariance of the Euler-Lagrange equation shows that there is a conserved current for each continuous symmetry.

Symmetry: A transformation of the space-time coordinates $x^\mu \to x'^\mu$, and the field $\phi_a(x) \to \phi'_a(x)$ preserves the equation of motion. Since the equation of motion is obtained by varying the action, the action should be preserved by the symmetry transformations. A slight generalization is that the action can in fact be changed by a surface term which does not influence its variation, the equation of motion at finite space-time points. Therefore, the symmetry transformations satisfy the condition

$$L(\phi, \partial\phi) \to L(\phi', \partial'\phi') + \partial'_\mu\Lambda^\mu \tag{C21}$$

with a certain local vector function $\Lambda^\mu(\phi(x), \partial\phi(x), x)$.

Continuous symmetry: There are infinitesimal symmetry transformations in an arbitrary small neighborhood of the identity, $x^\mu \to x^\mu + \delta x^\mu$, $\phi_a(x) \to \phi_a(x) + \delta\phi_a(x)$. Examples: Rotations, translations in the space-time, and $\phi(x) \to e^{i\alpha}\phi(x)$ for a complex field.

Conserved current: $\partial_\mu j^\mu = 0$, conserved charge: $Q(t)$:

$$\partial_0 Q(t) = \partial_0 \int_V d^3x j^0 = -\int_V d^3x \partial v j = -\int_{\partial V} d\mathbf{s} \cdot \mathbf{j} \tag{C22}$$

It is useful to distinguish external and internal spaces, corresponding to the space-time and the values of the field variable. Eg.

$$\phi_a(x): \underbrace{\mathbb{R}^4}_{\text{external space}} \to \underbrace{\mathbb{R}^n}_{\text{internal space}} . \tag{C23}$$

Internal and external symmetry transformations act on the internal or external space, respectively.

a. Point particle

The main points of the construction of the Noether current for internal symmetries can be best understood in the framework of a particle.

The intuitive idea behind Noether's theorem is to consider an infinitesimal symmetry transformation which is almost the identity and exists only for continuous symmetry as a variation around the solution of the Equler-Lagrange equation. Since the action is stationary for arbitrary variations it remains stationary for such a special variation, too. Thus the Euler-Lagrange equation which assumes the same form in any coordinate system reamins satisfied for the time dependent parameter of the transformation. The symmetry transformation should leave the Lagrangian invariance hence its parameter should be a cyclic coordinate hence its generalize momentum is conserved.

The more detailed proof is slightly more involved because of the more general way symmetry transformation may act. Let us start with the definition of the symmetry as a transformation of the time and the coordinate, $t \to t'$, $\mathbf{x}(t) \to \mathbf{x}'(t')$, which leaves the equation of motion unchanged. A sufficient condition for a transformation be symmetry is that it preserves the Lagrangian. But this is not necessary since the variational equations remains unchanged when a total time derivative is added to the Lagrangian it contributes by a boundary term which is irrelevant from the point of view of variations, performed at intermediate time. Thus a transformation is symmetry if the Lagrangian changes by a total time derivative,

$$L(\mathbf{x}, \dot{\mathbf{x}}) = L(\mathbf{x}', \dot{\mathbf{x}'}) + \dot{\Lambda}(t', \mathbf{x}'). \tag{C24}$$

The symmetry transformation consist of a group and the elements of a continuous group can be parametrized by continuous, real numbers. The usual convention is to assign 0 to the identity transformation hence the infinitesimal transformation in the vicinity of the identity can therefore be written in the form $\mathbf{x} \to \mathbf{x} + \epsilon \mathbf{f}(t, \mathbf{x})$, $t \to t + \epsilon f(t, \mathbf{x})$ with infinitesimal ϵ.

Let us consider only two kinds of symmetries for the sake of simplicity:

Change of the coordinates: $\mathbf{f} \neq 0$, $f = 0$: The symmetry transformation $\mathbf{x} \to \mathbf{x}' = \mathbf{x} + \epsilon \mathbf{f}$ with constant ϵ and the total time derivative can be neglected in (C24) as long as only the variational equations are sought. Hence the symmetry can be expressed by taking the derivative with respect to ϵ of (C24) without the total derivative term,

$$\begin{aligned} 0 &= \partial_\epsilon L(\mathbf{x} + \epsilon \mathbf{f}, \dot{\mathbf{x}} + \epsilon \partial_t \mathbf{f} + \epsilon (\dot{\mathbf{x}} \partial) \mathbf{f}) \\ &= \frac{\partial L}{\partial \mathbf{x}} \mathbf{f} + \frac{\partial L}{\partial \dot{\mathbf{x}}} (\partial_t \mathbf{f} + \dot{\mathbf{x}} \partial \mathbf{f}). \end{aligned} \tag{C25}$$

The parameter of the variation $\delta \mathbf{x} = \epsilon(t) \mathbf{f}$ is made time dependent, $\epsilon \to \epsilon(t)$ and its Lagrangian is

$$\begin{aligned} \tilde{L}(\epsilon, \dot{\epsilon}) &= L(\mathbf{x} + \epsilon \mathbf{f}, \dot{\mathbf{x}} + \epsilon \partial_t \mathbf{f} + \epsilon (\dot{\mathbf{x}} \partial) \mathbf{f} + \dot{\epsilon} \mathbf{f}) + \mathcal{O}\left(\epsilon^2\right) \\ &= \epsilon \left(\frac{\partial L}{\partial \mathbf{x}} \mathbf{f} + \frac{\partial L}{\partial \dot{\mathbf{x}}} \partial_t \mathbf{f} \right) + \frac{\partial L}{\partial \dot{\mathbf{x}}} [\epsilon (\dot{\mathbf{x}} \partial) \mathbf{f} + \dot{\epsilon} \mathbf{f}] + \mathcal{O}\left(\epsilon^2\right), \end{aligned} \tag{C26}$$

The variation principle, stated now as $\delta S[\epsilon] = \mathcal{O}\left(\epsilon^2\right)$, yields the Euler-Lagrange equation

$$\frac{\partial \tilde{L}(\epsilon, \dot{\epsilon})}{\partial \epsilon} = \frac{d}{dt}\frac{\partial \tilde{L}(\epsilon, \dot{\epsilon})}{\partial \dot{\epsilon}}. \tag{C27}$$

Note that the symmetry makes the symmetry transformation parameter a cyclic variable and its generalized momentum,

$$p_\epsilon = \frac{\partial \tilde{L}(\epsilon, \dot{\epsilon})}{\partial \dot{\epsilon}} = \frac{\partial L}{\partial \dot{\mathbf{x}}}\mathbf{f}, \tag{C28}$$

conserved.

Examples:

1. Translation symmetry, $\mathbf{f} = \mathbf{n}$, $\mathbf{n}^2 = 1$, of the Lagrangian $L = \frac{m}{2}\dot{\mathbf{x}}^2 - U(\mathbf{T}\mathbf{x})$ with $T = \mathbb{1} - \mathbf{n} \otimes \mathbf{n}$, leads to the conservation of the momentum $p_\epsilon = m\dot{\mathbf{x}}\mathbf{n}$.

2. Rotational symmetry, $\mathbf{f} = \mathbf{n} \times \mathbf{x}$, $\mathbf{n}^2 = 1$ of the Lagrangian $L = \frac{m}{2}\dot{\mathbf{x}}^2 - U(|\mathbf{x}|)$ implies the conservation of the angular momentum, $p_\epsilon = m\dot{\mathbf{x}}(\mathbf{n} \times \mathbf{x}) = \mathbf{n}(\mathbf{x} \times m\dot{\mathbf{x}}) = \mathbf{n}\mathbf{L}$.

Change of the time: The argument is different in this case. We use shifted time, $t \to t' = t + \epsilon(t)$, and the trajectory $\mathbf{x}(t) = \mathbf{x}(t' - \epsilon(t)) \approx \mathbf{x}(t' - \epsilon(t')) \approx \mathbf{x}(t') - \epsilon(t')\dot{\mathbf{x}}(t')$ to rewrite the action,

$$S[\mathbf{x}] = \int_{t_i+\epsilon(t_i)}^{t_f+\epsilon(t_f)} \frac{dt'}{1 + \dot{\epsilon}(t')} L(\mathbf{x}(t' - \epsilon(t')), \dot{\mathbf{x}}(t' - \epsilon(t'))) + \mathcal{O}\left(\epsilon^2\right). \tag{C29}$$

whose $\mathcal{O}\left(\epsilon\right)$ part is

$$\begin{aligned} 0 &= -\int_{t_i}^{t_f} dt \left(\epsilon\dot{\mathbf{x}}\frac{\partial L}{\partial \mathbf{x}} + \frac{d}{dt}\epsilon\dot{\mathbf{x}}\frac{\partial L}{\partial \dot{\mathbf{x}}} + \dot{\epsilon}L\right) + \epsilon L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\Big|_{t_i}^{t_f} \\ &= -\int_{t_i}^{t_f} dt \left[\epsilon\dot{\mathbf{x}}\left(\frac{\partial L}{\partial \mathbf{x}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}}\right) + \dot{\epsilon}L\right] + \epsilon\left(L - \dot{\mathbf{x}}\frac{\partial L}{\partial \dot{\mathbf{x}}}\right)_{t_i}^{t_f}. \end{aligned} \tag{C30}$$

The integral is vanishing in the last line because trajectory solves the original equation of motion and by setting a time-independent transformation parameter, $\epsilon(t) = \epsilon$, one finds that the Hamiltonian,

$$H = \frac{\partial L}{\partial \dot{\mathbf{x}}}\dot{\mathbf{x}} - L, \tag{C31}$$

is conserved.

b. Internal symmetries

An internal symmetry transformation of field theory acts on the internal space only. We shall consider linearly realized internal symmetries for simplicity,

$$\delta x^\mu = 0, \quad \delta\phi_a(x) = \epsilon\tau_{ab}\phi_b(x). \tag{C32}$$

where τ is called generator, c.f. section B. This transformation is a symmetry,

$$L(\phi, \partial\phi) = L(\phi + \epsilon\tau\phi, \partial\phi + \epsilon\tau\partial\phi) + \mathcal{O}\left(\epsilon^2\right). \tag{C33}$$

Let us introduce new "coordinates", ie. new field variable, $\Phi(\phi)$, in such a manner that $\Phi^1(x) = \epsilon(x)$ where $\phi(x) = \phi_{cl}(x) + \epsilon(x)\tau\phi_{cl}(x)$, $\phi_{cl}(x)$ being the solution of the equations of movement. The linearized Lagrangian for $\epsilon(x)$ is

$$\begin{aligned} L(\epsilon, \partial\epsilon) &= L(\phi_{cl} + \epsilon\tau\phi, \partial\phi_{cl} + \partial\epsilon\tau\phi + \epsilon\tau\partial\phi) \\ &= \frac{\partial L(\phi_{cl}, \partial\phi_{cl})}{\partial\phi}\epsilon\tau + \frac{\partial L(\phi_{cl}, \partial\phi_{cl})}{\partial\partial_\mu\phi}[\partial_\mu\epsilon\tau\phi + \epsilon\tau\partial_\mu\phi] + \mathcal{O}\left(\epsilon^2\right). \end{aligned} \tag{C34}$$

The symmetry, Eq. (C33), indicates that ϵ is a cyclic coordinate and the equation of motion

$$\frac{\partial L(\epsilon, \partial\epsilon)}{\partial\epsilon} - \partial_\mu \frac{\partial L(\epsilon, \partial\epsilon)}{\partial\partial_\mu\epsilon} = 0, \tag{C35}$$

shows that the current,

$$J^\mu = \frac{\partial L(\epsilon, \partial\epsilon)}{\partial\partial_\mu\epsilon} = \frac{\partial L(\phi_{cl}, \partial\phi_{cl})}{\partial\partial_\mu\phi}\tau\phi \tag{C36}$$

defined up to a multiplicative constant as the generalized momentum of ϵ, is conserved. Notice that (i) we have an independent conserved current corresponding to each independent direction in the internal symmetry group and (ii) the conserved current is well defined up to a multiplicative constant only.

Examples:

1. Real scalar field: ϕ_a, $a = 1, \cdots, n$, the symmetry group is $G = O(n)$,

$$\begin{aligned} L &= \frac{1}{2}(\partial\phi)^2 - V(\phi) \\ \delta\phi &= \epsilon^a\tau^a\phi, \quad \tau^a \in o(n) \\ J^a_\mu &= \partial_\mu\phi\tau^a\phi \end{aligned} \tag{C37}$$

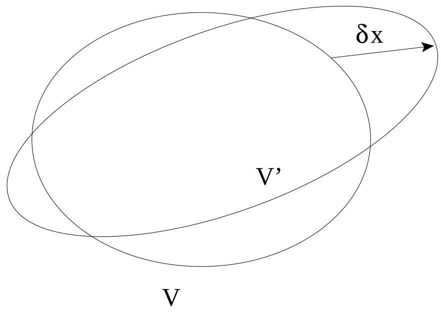


FIG. 11: Deformation of the space-time region.

2. Complex scalar field: ϕ_a, $a = 1, \cdots, n$, $G = U(n)$

$$\begin{aligned} L &= \partial\phi^\dagger\partial\phi - V(\phi) \\ \delta\phi &= i\epsilon^a\tau^a\phi, \quad \tau^a \in u(n) \\ J^a_\mu &= i\partial_\mu\phi^\dagger\tau^a\phi - i\phi^\dagger\tau^a\partial_\mu\phi \end{aligned} \tag{C38}$$

3. Electromagnetic current: ϕ, $G = U(1) = SO(2)$, $\phi = \phi_1 + i\phi_2 = \binom{\phi_1}{\phi_2}$, $i = \binom{0,-1}{1,\ 0}$

$$\begin{aligned} L &= \frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 + \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - m^2(\phi_1^2 + \phi_2^2) - V(\phi_1^2 + \phi_2^2) \\ J_\mu &= -^{(\partial_\mu\phi_1,\partial_\mu\phi_2)}\binom{0,-1}{1,\ 0}\binom{\phi_1}{\phi_2} = -\partial_\mu\phi_2\phi_1 + 2\partial_\mu\phi_1\phi_2 = i(\partial_\mu\phi^\dagger\phi - \phi^\dagger\partial_\mu\phi) = -i\phi^\dagger\overleftrightarrow{\partial}_\mu\phi \end{aligned} \tag{C39}$$

c. External Symmetries

The non-relativistic external symmetry group, the Galilean group consists of translations of the space-time, rotations of the space and boosts and is a $4 + 3 + 3 = 10$ dimensional continuous group. The relativistic Poincaré group has the same dimension and contains translations and Lorentz transformations of the space-time. We shall consider the conserved currents related to the translation invariance only for the sake of simplicity.

We can rewrite the action in terms of the infinitesimally changed coordinates, $x'^\mu = x^\mu + \epsilon^\mu$, $S = S'$. The invariance of S' under the transformation $x \to x' + \epsilon\ \phi(x) \to \phi'(x) = \phi(x) + \delta\phi(x)$, $\delta\phi(x) = -\epsilon^\mu\partial_\mu\phi(x)$, renders the $\mathcal{O}(\epsilon)$ part of S' vanishing,

$$\begin{aligned} 0 &= \int_V \delta L(\phi(x), \partial\phi(x)) + \int_{V'-V} dx L(\phi(x), \partial\phi(x)) \\ &= \int_V \delta L(\phi(x), \partial\phi(x)) + \int_{\partial V} dS_\nu\epsilon^\nu L(\phi(x), \partial\phi(x)) \\ &= -\int_V dx\epsilon^\nu\partial_\nu\phi\left(\frac{\partial L(\phi,\partial\phi)}{\partial\phi} - \partial_\mu\frac{\partial L(\phi,\partial\phi)}{\partial\partial_\mu\phi}\right) \\ &\quad \int_{\partial V} dS_\mu\left[-\epsilon^\nu\partial_\nu\phi\frac{\partial L(\phi,\partial\phi)}{\partial\partial_\mu\phi} + \epsilon^\mu L(\phi,\partial\phi)\right] \end{aligned} \tag{C40}$$

The field configuration satisfies the equation of motion hence the first term on the right hand side of the last equation is vanishing. This holds for arbitrary translation ϵ^ν and space-time region, V, rendering the the energy-momentum tensor,

$$T^{\mu\nu} = \frac{\partial L}{\partial\partial_\nu\phi}\partial^\mu\phi - g^{\mu\nu}L \tag{C41}$$

conserved, $\partial_\mu T^{\mu\nu} = 0$. The "charge" of the translation ϵ^ν,

$$P^\mu = \int d^3x T^{0\mu}, \tag{C42}$$

defines the energy-momentum vector. The energy-momentum tensor can be parametrized by

$$T^{\mu\nu} = \begin{pmatrix} \epsilon & c\mathbf{p} \\ \frac{1}{c}\mathbf{s} & \sigma \end{pmatrix} \tag{C43}$$

where

$$\begin{aligned} \epsilon &= \text{energy density} \\ \mathbf{p} &= \text{momentum density} \\ \mathbf{S} &= \text{energy flux density} \\ \sigma^{jk} &= \text{momentum flux } p^k \text{ in the direction } j \end{aligned} \tag{C44}$$

(c is restored).

The Lorentz symmetry leads to six conserved currents, 3 of which give the angular momentum and other three are the generators of the Lorentz boosts. The conservation of angular momentum can be used to prove that the energy-momentum tensor is symmetric, $T^{\mu\nu} = T^{\nu\mu}$ for bosonic field theories.

Appendix D: Parallel transport along a path

The expression of parallel transport, the solution to Eq. (48) is worked out in this Appendix. The solution can formally be written as

$$W_\gamma(y,x) = P\left[e^{-\int_x^y d\gamma^\mu A_\mu(\gamma)}\right] = P\left[e^{-\int_0^1 ds \frac{d\gamma^\mu(s)}{ds}A_\mu(\gamma(s))}\right], \tag{D1}$$

by means of the path ordered product of non-commuting objects defined along the path γ, defined as

$$P[A(s_A)B(s_B)] = \Theta(s_A - s_B)A(s_A)B(s_B) + \Theta(s_B - s_A)B(s_B)A(s_A). \tag{D2}$$

To see that we have the correct solution let us write first the integral in the exponent in Eq. (D1) as

$$\int_0^1 ds \frac{d\gamma^\mu(s)}{ds} A_\mu(\gamma(s)) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^N \frac{d\gamma^\mu(s_j)}{ds} A_\mu(\gamma(s_j)) \tag{D3}$$

where $s_j = j/N$. The exponential function of an operator is defined by its Taylor-series,

$$\begin{aligned} e^{-\frac{1}{N}\sum_{j=1}^N \frac{d\gamma^\mu(s_j)}{ds} A_\mu(\gamma(s_j))} &= 1 - \frac{1}{N}\sum_{j=1}^N \frac{d\gamma^\mu(s_j)}{ds} A_\mu(\gamma(s_j)) \\ &+ \frac{1}{N^2}\sum_{j_1,j_2=1}^N \frac{d\gamma^\mu(s_{j_1})}{ds}\frac{d\gamma^\mu(s_{j_2})}{ds} A_\mu(\gamma(s_{j_1}))A_\mu(\gamma(s_{j_2})) + \cdots, \end{aligned} \tag{D4}$$

and the path ordering applies term-by-term,

$$\begin{aligned} W_\gamma(y,x) &= \lim_{N\to\infty}\Bigg[1 - \frac{1}{N}\sum_{j=1}^N \frac{d\gamma^\mu(s_j)}{ds} A_\mu(\gamma(s_j)) \\ &+ \frac{1}{N^2}\sum_{j_1,j_2=1}^N \frac{d\gamma^\mu(s_{j_1})}{ds}\frac{d\gamma^\mu(s_{j_2})}{ds} P[A_\mu(\gamma(s_{j_1}))A_\mu(\gamma(s_{j_2}))] + \cdots\Bigg]. \end{aligned} \tag{D5}$$

We would have

$$W_\gamma(y,x) = \lim_{N\to\infty}\prod_{j=1}^N e^{-\frac{1}{N}\frac{d\gamma^\mu(s_j)}{ds}A_\mu(\gamma(s_j))} \tag{D6}$$

according to the well known rule $e^a e^b = e^{a+b}$, valid for numbers, without paying attention to the non-commutativity of the objects occurring in the product. But the path ordering places the contributions corresponding to higher j more to the left in the products and we find

$$W_\gamma(y,x) = \lim_{N\to\infty} e^{-\frac{1}{N}\frac{d\gamma^\mu(s_N)}{ds}A_\mu(\gamma(s_N))}\cdots e^{-\frac{1}{N}\frac{d\gamma^\mu(s_1)}{d\tau}A_\mu(\gamma(s_1))} \tag{D7}$$

by repeating the same resummation as in Eq. (D6). The path ordering succeeded in factorizing the dependence on the N-th division point to the integral at the very left of the product. The final step is the calculation of the partial derivative of W along the path,

$$\begin{aligned} \frac{1}{N}\frac{d\gamma^\mu}{ds}\partial_{y^\mu}W_\gamma(y,x) &= W_\gamma(y,x) - W_\gamma\left(y - \frac{1}{N}\frac{d\gamma^\mu(s_N)}{ds}, x\right) \\ &= \left[e^{-\frac{1}{N}\frac{d\gamma^\mu(s_N)}{ds}A_\mu(\gamma(s_N))} - \mathbb{1}\right] e^{-\frac{1}{N}\frac{d\gamma^\mu(s_{N-1})}{ds}A_\mu(\gamma(s_{N-1}))}\cdots e^{-\frac{1}{N}\frac{d\gamma^\mu(s_1)}{d\tau}A_\mu(\gamma(s_1))} \\ &\approx -\frac{1}{N}\frac{d\gamma^\mu(s_N)}{ds}A_\mu(\gamma(s_N))W_\gamma\left(y - \frac{1}{N}\frac{d\gamma^\mu(s_N)}{ds}, x\right) \\ &\approx -\frac{1}{N}\frac{d\gamma^\mu(1)}{ds}A_\mu(\gamma(1))W_\gamma(y,x), \end{aligned} \tag{D8}$$

which yields Eq. (48).

Appendix E: Gauge theory of the Poincaré group

The brief derivation of the Euler-Lagrange equations is presented here for the gauge theory formalism based on the Poincaré-group. The external space is the space-time as usual. The gauge group is chosen to be the symmetry of the local dynamics expressed in a coordinate system specified by the Equivalence Principle. The gravitation interaction is absence and the Poincaré symmetry of the Special Relativity is recovered in this coordinate system. Therefore the internal space is chosen to be the fundamental representation of the Poincaré group, a four dimensional real vector space equipped with the Lorentzian metric tensor

$$\eta_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{E1}$$

Notice that this metric is homogeneous, space-time independent as required by the Equivalence Principle. Gauge transformations act as

$$\xi^a(x) \to \xi^a(x) + \Lambda^a{}_b(x)\xi^b(x) + \zeta^a(x) \tag{E2}$$

where $\xi^a(x)$ is the internal space coordinate corresponding to the space-time point x, $\Lambda^a{}_b$ denotes a Lorentz transformation matrix and ζ^a stands for translations. The gauge group, the Poincaré group, is the direct product of translations and Lorentz transformations, $P = T \times L$.

1. Covariant derivatives

We shall introduce the covariant derivative in the following steps. Fist we construct it for scalars $s(x)$, vectors $v^a(x)$ and tensors $t^{ab\cdots}(x)$ of the internal Lorentz symmetry. Next we extend it for vectors or tensors defined by the space-time, $v^\mu(x)$, $t^{\mu\nu\cdots}(x)$.

a. Vierbein

Let us start by recalling that the unique feature of gravity as gauge theory is that the internal space is the tangent space of the base manifold. Therefore there is a unique correspondence between infinitesimal translations in the internal and the external spaces and the dependence on x and ξ can be traded locally. A function $f(y)$ defined in the vicinity of x gives rise the function

$$f_x(\xi) = f(x + \Delta x) \tag{E3}$$

defined in the vicinity of the origin of the Lorentz space where

$$\xi^a = e^a_\mu \Delta x^\mu. \tag{E4}$$

The matrix e^a_μ relating the directions in the two spaces,

$$e^a_\mu = \frac{\partial \xi^a}{\partial x^\mu}, \tag{E5}$$

is called vierbein. The relation

$$e^a_\mu \partial_a s = \frac{\partial s}{\partial x^\mu} = \partial_\mu s \tag{E6}$$

follows in an obvious manner and allows to represent the functions over the space-time locally as functions over the tangent space, according to the intuitive role played by the coordinates in the Equivalence Principle.

The inverse of the transformation (E5) is

$$e^\mu_a = \frac{\partial x^\mu}{\partial \xi^a}, \tag{E7}$$

and the metric tensor is given by

$$g^{\mu\nu}(x) = e^\mu_a(x)\eta^{ab}e^\nu_b(x). \tag{E8}$$

Thus we have the identities

$$e^\mu_a e^a_\nu = \delta^\mu_\nu, \qquad e^\mu_a e^b_\mu = \delta^b_a, \tag{E9}$$

and any vector or tensor can be represented as a Lorentz or world vector or tensor as in Eq. (E6), eg.

$$v^\mu = \frac{\partial x^\mu}{\partial \xi^a} v^a = e^\mu_a v^a, \qquad v^a = \frac{\partial \xi^a}{\partial x^\mu} v^\mu = e^a_\mu v^\mu. \tag{E10}$$

Note that the invariant integral measure can be written as

$$\sqrt{-\det g_{\mu\nu}}\, d^4x = \det(e^a_\mu) d^4x. \tag{E11}$$

b. Holonomic and anholonomic vector fields

The Equivalence Principle assures that the gravitation interaction can locally by eliminated by a suitable choice of the coordinate system which is realised by the internal space. This special coordinate system can be extended in a finite regions in the absence of gravitational field only.

What is the condition on four world vector fields $e_a^\mu(x)$ with $a = 0, 1, 2, 3$ for the existence of a coordinate systems with these coordinate axes?

The solution of the differential equation

$$\frac{\partial x^\mu(\xi)}{\partial \xi} = e^\mu(x) \tag{E12}$$

can be written as

$$x^\mu(\xi) = e^{\xi \frac{\partial}{\partial \xi}} x^\mu(0). \tag{E13}$$

We are now looking for the coordinates, defined by the coordinate axes

$$\frac{\partial x^\mu(\xi)}{\partial \xi^a} = e_a^\mu(x). \tag{E14}$$

Let us suppose that the world vectors $e_a^\mu(x)$ are independent and the derivatives $\partial_a = \frac{\partial}{\partial \xi^a}$ commute,

$$\begin{aligned} [\partial_a, \partial_b] &= [e_a^\mu \partial_\mu, e_b^\nu \partial_\nu] \\ &= (e_a^\mu \partial_\mu e_b^\nu - e_b^\mu \partial_\mu e_a^\nu)\partial_\nu = 0. \end{aligned} \tag{E15}$$

Then the coordinates can locally be introduced for example by

$$x^\mu(\xi) = e^{\xi^0 \frac{\partial}{\partial \xi^0}} e^{\xi^1 \frac{\partial}{\partial \xi^1}} e^{\xi^2 \frac{\partial}{\partial \xi^2}} e^{\xi^3 \frac{\partial}{\partial \xi^3}} x^\mu(\xi)_{\xi=0}. \tag{E16}$$

In fact, by varying ξ around zero we cover all four dimensions and

$$\begin{aligned} \partial_a x^\mu(0) &= \partial_a e^{\xi^0 \frac{\partial}{\partial \xi^0}} e^{\xi^1 \frac{\partial}{\partial \xi^1}} e^{\xi^2 \frac{\partial}{\partial \xi^2}} e^{\xi^3 \frac{\partial}{\partial \xi^3}} x^\mu(\xi)_{\xi=0} \\ &= \partial_a \left(1 + \xi^a \frac{\partial}{\partial \xi^a}\right) x^\mu(\xi)_{\xi=0} \\ &= e_a^\mu(x(0)). \end{aligned} \tag{E17}$$

When the vectors e_a^μ are independent only but the commutators are non-vanishing then ξ^a is already an admissible coordinate system but the coordinate axes are different than the vectors e_a^μ.

Conversely, let us suppose that the the solution of the system of differential equations lead to admissible coordinates. The invertibility of the coordinate transformation requires the independence of the world vectors e_a^μ and the commutativity of the derivatives, ie. the symmetry of partial derivatives of continuously derivable functions is trivial in the coordinate system ξ^a.

The vectors e_a^μ are called holonomic if they are the directional vectors of a local coordinate system. Therefore the vierbein defines anholonomic vectors in the presence of gravitational field.

c. Local Lorentz transformations

The covariant derivative involves a generator valued vector field. The gauge group is a direct product, $P = T \times L$, therefore there will be separate gauge fields for translations and Lorentz transformations. Due to continuity requirements the proper Lorentz group is generated by these generators only, discrete inversions will be left out.

Local Lorentz transformations generate space-dependent orientation for the orthogonal Lorentz coordinate basis and lead to the covariant derivative

$$D_\mu^{(L)} = \frac{\partial}{\partial x^\mu} + \omega_\mu \tag{E18}$$

where the affine connection, $\omega_{ab\mu} = \eta_{ac}(\omega_\mu)^c{}_b = \eta_{ac}\omega^c{}_{b\mu}$, $\omega_{ab\mu} = -\omega_{ba\mu}$ is a generator of the special Lorentz group in the fundamental representation (E2). The connection acts in the internal space only, ie. on Lorentz tensors $T_{a,b,\cdots}$, eg.

$$D_\mu^{(L)} v^b = \partial_\mu v^b + \omega^b{}_{c\mu} v^c \tag{E19}$$

and considers the world tensors $T_{\mu\nu\cdots}$ as scalars,

$$D_\mu^{(L)} v^\mu = \partial_\mu v^\mu. \tag{E20}$$

In order to preserve the Lorentzian scalar product by parallel transport,

$$D_\mu^{(L)}(v_b u^b) = \partial_\mu(v_b u^b) \tag{E21}$$

we define

$$D_\mu^{(L)} v_b = \partial_\mu v_b - v_c \omega^c{}_{b\mu}. \tag{E22}$$

Notice the vanishing of the covariant derivative of the internal space metric,

$$\begin{aligned} D_\mu^{(i)} \eta^{ab} &= \omega^a{}_{c\mu}\eta^{cb} + \omega^b{}_{c\mu}\eta^{ac} \\ &= \omega^{ab}{}_\mu + \omega^{ba}{}_\mu \\ &= 0. \end{aligned} \tag{E23}$$

d. Local translations

Local translations generate space-time dependent shift of the Lorentz coordinate system which in turn induces a shift in the space-time due to the fact that the internal space is the tangent space of the space-time. The corresponding covariant derivative is

$$D^{(T)}_\mu s = \delta^a_\mu \partial_a s + t^a_\mu \partial_a s = e^a_\mu \partial_a s \tag{E24}$$

for a scalar function $s(x)$ with $t^a_\mu(x)$ as gauge field. The first term on the right hand side implements the infinitesimal shift $\Delta x^\mu D^{(T)}_\mu$ in the internal space which corresponds to an infinitesimal shit $\Delta x^\mu \partial_\mu$ in the external space and the second terms compensates for the difference of the position of the origin of the internal coordinate system at the space-time points x and $x + \Delta x$. Thus the vierbein, satisfying Eq. (E6) is

$$e^a_\mu = \delta^a_\mu + t^a_\mu. \tag{E25}$$

e. Full local Poincare group

The full internal symmetry is incorporated into the covariant derivative $D^{(i)}_\mu$ acts as $D^{(L)}_\mu$,

$$\begin{aligned} D^{(i)}_\mu s &= D^{(T)}_\mu s = D^{(L)}_\mu s = \partial_\mu s \\ D^{(i)}_\mu v^b &= \delta^a_\mu \partial_a v^b + t^a_\mu \partial_a v^b + \omega^b{}_{c\mu} v_c = \partial_\mu v^b + \omega^b{}_{c\mu} v_c = D^{(L)}_\mu v^b \\ D^{(i)}_\mu v_b &= \delta^a_\mu \partial_a v_b + t^a_\mu \partial_a v_b - v_c \omega^c{}_{b\mu} = \partial_\mu v_b - v_c \omega^c{}_{b\mu} = D^{(L)}_\mu v_b \end{aligned} \tag{E26}$$

on scalars and Lorentz tensors and treating world vectors and tensors as scalars because ∂_μ can be considered as the covariant derivative for local translations.

It is an unusual feature that one can introduce derivatives in the internal space directions, ∂_a. The corresponding covariant derivative is

$$D^{(i)}_a = e^\mu_a D^{(i)}_\mu = e^\mu_a (\partial_\mu + \omega_\mu). \tag{E27}$$

Notice the natural relation

$$D^{(i)}_a v = \partial_a v + e^\mu_a t^b_\mu \partial_b v = \partial_a v + e^\mu_a (e^b_\mu - \delta^b_\mu) \partial_b v = 2\partial_a v - e^\mu_a \delta^b_\mu \partial_b v = \partial_a v. \tag{E28}$$

f. World vectors and tensors

We could stop at this stage and start to work out the gauge theory for the Poincaré group by means of the covariant derivative $D^{(i)}$. The drawback would be to use the internal Minkowski space vectors v^a or tensors $t^{ab\cdots}$ only. This is obviously an artificial constraint because the coordinates x^μ lead naturally to world vectors v^μ or tensors $t^{\mu\nu\cdots}$. In order to be construct covariant equations for these vectors or tensors we need the covariant derivative for the $GL(4)$ gauge theory, controlling the effects of coordinate transformations induced by the application of the local Poincaré group in the internal space. The usual affine connection Γ_μ, introduced in differential geometry, realizes this covariant derivative

$$D_\mu^{(e)} = \partial_\mu + \Gamma_\mu \tag{E29}$$

by acting on world scalars, vectors and tensors, eg.

$$\begin{aligned} D_\mu^{(e)} s &= \partial_\mu s \\ D_\mu^{(e)} v^\nu &= \partial_\mu v^\nu + \Gamma^\nu{}_{\rho\mu} v^\rho \\ D_\mu^{(e)} v_\nu &= \partial_\mu v_\nu - v_\rho \Gamma^\nu{}_{\rho\mu}. \end{aligned} \tag{E30}$$

The most general covariant derivatives D compensates in all spaces and indices, eg. its action a Lorentz and world vector $v^{a\mu}$ is

$$(D_\nu v)^{a\mu} = [\partial_\nu v + (\omega_\nu + \Gamma_\nu) v]^{a\mu}, \tag{E31}$$

etc.

2. Field strength tensors

The field strength tensor,

$$\begin{aligned} F_{\mu\nu} &= [D_\mu, D_\nu] = [\partial_\mu + \Gamma_\mu + \omega_\mu, \partial_\nu + \Gamma_\nu + \omega_\nu] \\ &= \partial_\mu \omega_\nu - \partial_\nu \omega_\mu + [\omega_\mu, \omega_\nu] + \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + [\Gamma_\mu, \Gamma_\nu], \end{aligned} \tag{E32}$$

is called curvature, it satisfies the Bianchi identity,

$$\begin{aligned} 0 &= [D_\mu, [D_\nu, D_\rho]] + [D_\nu, [D_\rho, D_\mu]] + [D_\rho, [D_\mu, D_\nu]] \\ &= [D_\mu, F_{\nu\rho}] + [D_\nu, F_{\rho\mu}] + [D_\rho, F_{\mu\nu}] \\ &= D_\mu F_{\nu\rho} + D_\nu F_{\rho\mu} + D_\rho F_{\mu\nu}. \end{aligned} \tag{E33}$$

The field strength tensor acting on the internal space,

$$F_{\mu\nu}^{(i)} = [D_\mu^{(i)}, D_\nu^{(i)}] = [\partial_\mu + \omega_\mu, \partial_\nu + \omega_\nu] = \partial_\mu \omega_\nu - \partial_\nu \omega_\mu + [\omega_\mu, \omega_\nu], \tag{E34}$$

reads as

$$F^{(i)a}{}_{b\mu\nu} = (F_{\mu\nu}^{(i)})^a{}_b = \partial_\mu \omega^a{}_{b\nu} - \partial_\nu \omega^a{}_{b\mu} + \omega^a{}_{c\mu} \omega^c{}_{b\nu} - \omega^a{}_{c\nu} \omega^c{}_{b\mu} \tag{E35}$$

when all indices are shown. One can introduce

$$F^{(i)}_{cd\mu\nu} = \eta_{ce} F^{(i)e}{}_{d\mu\nu} = -F^{(i)}_{cd\nu\mu} = -F^{(i)}_{dc\mu\nu}. \tag{E36}$$

The field strength tensor for the external space reads as

$$F^{(e)}_{\mu\nu} = [D^{(e)}_\mu, D^{(e)}_\nu] = [\partial_\mu + \Gamma_\mu, \partial_\nu + \Gamma_\nu] = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + [\Gamma_\mu, \Gamma_\nu] \tag{E37}$$

or

$$F^{(e)\rho}{}_{\sigma\mu\nu} = (F^{(e)}_{\mu\nu})^\rho{}_\sigma = \partial_\mu \Gamma^\rho{}_{\sigma\nu} - \partial_\nu \Gamma^\rho{}_{\sigma\mu} + \Gamma^\rho{}_{\kappa\mu}\Gamma^\kappa{}_{\sigma\nu} - \Gamma^\rho{}_{\kappa\nu}\Gamma^\kappa{}_{\sigma\mu}. \tag{E38}$$

The field strength tensor corresponding to internal directions is

$$\begin{aligned}
F_{ab} &= [D_a, D_b] = [e^\mu_a D_\mu, e^\nu_b D_\nu] \\
&= e^\mu_a e^\nu_b [D_\mu, D_\nu] + e^\mu_a [D_\mu, e^\nu_b] D_\nu + e^\nu_b [e^\mu_a, D_\nu] D_\mu + [e^\mu_a, e^\nu_b] D_\nu D_\mu \\
&= e^\mu_a e^\nu_b F_{\mu\nu} + e^\mu_a [D_\mu, e^\nu_b] D_\nu + e^\nu_b [e^\mu_a, D_\nu] D_\mu \\
&= e^\mu_a e^\nu_b F_{\mu\nu} + e^\mu_a D_\mu e^\nu_b D_\nu - e^\mu_a e^\nu_b D_\mu D_\nu + e^\nu_b e^\mu_a D_\nu D_\mu - e^\nu_b D_\nu e^\mu_a D_\mu \\
&= e^\mu_a e^\nu_b F_{\mu\nu} + e^\mu_a (D_\mu e^\nu_b) D_\nu - e^\nu_b (D_\nu e^\mu_a) D_\mu \\
&= e^\mu_a e^\nu_b F_{\mu\nu} + S^\mu_{ab} D_\mu
\end{aligned} \tag{E39}$$

with

$$\begin{aligned}
S^\mu_{ab} &= e^\nu_a D_\nu e^\mu_b - e^\nu_b D_\nu e^\mu_a \\
&= D_a e^\mu_b - D_b e^\mu_a \\
&= e^\nu_a (\partial_\nu e^\mu_b - e^\mu_c \omega^c{}_{b\nu} + \Gamma^\mu{}_{\rho\nu} e^\rho_b) - e^\nu_b (\partial_\nu e^\mu_a - e^\mu_c \omega^c{}_{a\nu} + \Gamma^\mu{}_{\rho\nu} e^\rho_a)
\end{aligned} \tag{E40}$$

being the torsion tensor. The translational part of D_a defines the field strength tensor

$$t^\mu{}_{ab} = \partial_a e^\mu_b - \partial_b e^\mu_a, \tag{E41}$$

cf. Eq.(E15), because

$$\begin{aligned}
[\partial_a, \partial_b] &= [e^\mu_a \partial_\mu, e^\nu_b \partial_\nu] \\
&= e^\mu_a [\partial_\mu, e^\nu_b] \partial_\nu + e^\nu_b [e^\mu_a, \partial_\nu] \partial_\mu + [e^\mu_a, e^\nu_b] \partial_\nu \partial_\mu \\
&= e^\mu_a [\partial_\mu, e^\nu_b] \partial_\nu + e^\nu_b [e^\mu_a, \partial_\nu] \partial_\mu \\
&= e^\mu_a \partial_\mu e^\nu_b \partial_\nu - e^\mu_a e^\nu_b \partial_\mu \partial_\nu + e^\nu_b e^\mu_a \partial_\nu \partial_\mu - e^\nu_b \partial_\nu e^\mu_a \partial_\mu \\
&= e^\mu_a (\partial_\mu e^\nu_b) \partial_\nu - e^\nu_b (\partial_\nu e^\mu_a) \partial_\mu \\
&= t^\mu{}_{ab} \partial_\mu.
\end{aligned} \tag{E42}$$

3. Variational equations

Let us suppose that the action can be written as

$$S = S_g[e,\omega] + S_m[\phi,e,\omega] \tag{E43}$$

where $S_m[\phi,e,\omega]$ controls the dynamics of the matter field, denoted generically by ϕ on a given geometry, specified by the fields e and ω. For a scalar field we may have

$$S_m[\phi,e,\omega] = \int dxE\left[\frac{1}{2}D^{(e)}_\mu\phi D^{(e)\mu}\phi - V(\phi)\right] \tag{E44}$$

where

$$E = \mathrm{det e^a_\mu} \tag{E45}$$

and for a fermion

$$S_m[\bar\psi,\psi,e,\omega]\int dxE\bar\psi i\gamma^a \underbrace{e^\mu_a(\partial_\mu + \omega^{ab}_\mu\tau_{ab})}_{D_a}\psi \tag{E46}$$

with $\tau_{ab} = -\frac{1}{4}[\gamma_a,\gamma_b]$. The gravitational action is chosen to be of the Einstein-Hilbert type,

$$S_g[e,\omega] = -\frac{1}{16\pi G}\int dxER. \tag{E47}$$

a. *Variation of* ω

The scalar curvature can be written as

$$\begin{aligned} R &= \eta^{bc}F^a_{\ bac} \\ &= \eta^{bd}e^\mu_a e^\nu_d(\partial_\mu\omega^a_{\ b\nu} - \partial_\nu\omega^a_{\ b\mu} + \omega^a_{\ c\mu}\omega^c_{\ b\nu} - \omega^a_{\ c\nu}\omega^c_{\ b\mu}) \\ &= \eta^{bd}(e^\mu_a e^\nu_d - e^\nu_a e^\mu_d)(\partial_\mu\omega^a_{\ b\nu} + \omega^a_{\ c\mu}\omega^c_{\ b\nu}) \\ &= \eta^{bd}T^{\mu\nu}_{ad}(\partial_\mu\omega^a_{\ b\nu} + \omega^a_{\ c\mu}\omega^c_{\ b\nu}) \end{aligned} \tag{E48}$$

where

$$T^{\mu\nu}_{ab} = e^\mu_a e^\nu_b - e^\nu_a e^\mu_b. \tag{E49}$$

The variation of the Lorentz connection gives the equation

$$\begin{aligned} 0 = -16\pi G\frac{\delta S}{\delta\omega^a_{\ b\mu}} &= E\eta^{cd}T^{\mu\nu}_{ad}\omega^b_{\ c\nu} + E\eta^{bd}T^{\nu\mu}_{cd}\omega^c_{\ a\nu} - \partial_\nu(E\eta^{bd}T^{\nu\mu}_{ad}) \\ &= ET^{\mu\nu c}_a\omega^b_{\ c\nu} + ET^{\nu\mu b}_c\omega^c_{\ a\nu} - \partial_\nu(ET^{\nu\mu b}_a) \\ &= -E\omega^b_{\ c\nu}T^{\nu\mu c}_a + ET^{\nu\mu b}_c\omega^c_{\ a\nu} - \partial_\nu(ET^{\nu\mu b}_a) \\ &= -D^{(i)}_\nu(ET^{\nu\mu b}_a) \end{aligned} \tag{E50}$$

The relation

$$\partial_\mu E = \partial_\mu e^{\mathrm{tr}\ln e} = e^{\mathrm{tr}\ln e}\mathrm{tr}e^{-1}\partial_\mu e = E e_a^\nu \partial_\nu e_\mu^a \tag{E51}$$

allows us to write it in the form

$$\begin{aligned} 0 &= E T_a^{\nu\mu b} e_c^\rho \partial_\rho e_\nu^c + E\partial_\nu T_a^{\nu\mu b} + E\omega^b{}_{c\nu} T_a^{\nu\mu c} - E T_c^{\nu\mu b}\omega^c{}_{a\nu} \\ &= E[T_a^{\nu\mu b} e_c^\rho \partial_\rho e_\nu^c + \partial_\nu T_a^{\nu\mu b} + \omega^b{}_{c\nu} T_a^{\nu\mu c} - T_c^{\nu\mu b}\omega^c{}_{a\nu}] \\ &= E[T_a^{\nu\mu b} e_c^\rho \partial_\rho e_\nu^c + D_\nu^{(i)} T_a^{\nu\mu b}] \end{aligned} \tag{E52}$$

or

$$-T_a^{\nu\mu b} e_c^\rho \partial_\rho e_\nu^c = \partial_\nu T_a^{\nu\mu b} + \omega^b{}_{c\nu} T_a^{\nu\mu c} - T_c^{\nu\mu b}\omega^c{}_{a\nu} \tag{E53}$$

which in turn gives

$$\begin{aligned} T_{ab}^{\nu\mu}\partial_\rho e_c^\rho e_\nu^c &= -T_{ab}^{\nu\mu} e_c^\rho \partial_\rho e_\nu^c \\ &= \partial_\nu T_{ab}^{\nu\mu} + \omega_{bc\nu} T_a^{\nu\mu c} - T_{cb}^{\nu\mu}\omega^c{}_{a\nu} \\ &= \partial_\nu T_{ab}^{\nu\mu} - \omega_{cb\nu} T_a^{\nu\mu c} - T_{cb}^{\nu\mu}\omega^c{}_{a\nu} \\ &= \partial_\nu T_{ab}^{\nu\mu} - \omega^c{}_{b\nu} T_{ac}^{\nu\mu} - T_{cb}^{\nu\mu}\omega^c{}_{a\nu} \\ &= D_\nu T_{ab}^{\nu\mu}. \end{aligned} \tag{E54}$$

In order to solve the equation

$$T_{ab}^{\nu\mu}\partial_\rho e_c^\rho e_\nu^c = \partial_\nu T_{ab}^{\nu\mu} - \omega^c{}_{b\nu} T_{ac}^{\nu\mu} - T_{cb}^{\nu\mu}\omega^c{}_{a\nu} \tag{E55}$$

for the Lorentz connection we write

$$\begin{aligned} (e_a^\nu e_b^\mu - e_a^\mu e_b^\nu)\partial_\rho e_c^\rho e_\nu^c &= \partial_\nu(e_a^\nu e_b^\mu - e_a^\mu e_b^\nu) - \omega^c{}_{b\nu}(e_a^\nu e_c^\mu - e_a^\mu e_c^\nu) - (e_c^\nu e_b^\mu - e_c^\mu e_b^\nu)\omega^c{}_{a\nu} \\ e_b^\mu \partial_\rho e_a^\rho - e_a^\mu \partial_\rho e_b^\rho &= \partial_\nu(e_a^\nu e_b^\mu - e_a^\mu e_b^\nu) + \omega^c{}_{bc} e_a^\mu - \omega^c{}_{ac} e_b^\mu + 2\omega^\mu{}_{ab} \end{aligned} \tag{E56}$$

and we find

$$2\omega^\mu{}_{ab} + \omega^c{}_{bc} e_a^\mu - \omega^c{}_{ac} e_b^\mu = t^\mu{}_{ab} \tag{E57}$$

where

$$t^\mu{}_{ab} = \partial_\nu(e_a^\nu e_b^\mu - e_a^\mu e_b^\nu) - e_b^\mu \partial_\rho e_a^\rho + e_a^\mu \partial_\rho e_b^\rho \tag{E58}$$

It will be more useful to have Lorentz indices only for the connection, therefore we multiply Eq. (E57) by e^d_μ

$$2\omega^d{}_{ab} + \omega^c{}_{bc}\delta^d_a - \omega^c{}_{ac}\delta^d_b = t^d{}_{ab}. \tag{E59}$$

To find the second and third terms on the left hand side we contract the indices a and d and find

$$\omega^c{}_{bc} = \frac{1}{3}t^c{}_{cb} \tag{E60}$$

by recalling the antisymmetry of the spacial Lorentz group generators, $\omega_{abc} = -\omega_{bac}$. This result leads us to the solution

$$\omega^d{}_{ab} = \frac{1}{6}t^c{}_{ca}\delta^d_b - \frac{1}{6}t^c{}_{cb}\delta^d_a + \frac{1}{2}t^d{}_{ab} \tag{E61}$$

where

$$\begin{aligned}
t^c{}_{ab} &= e^c_\mu\partial_\nu(e^\nu_a e^\mu_b - e^\mu_a e^\nu_b) - \delta^c_b\partial_\rho e^\rho_a + \delta^c_a\partial_\rho e^\rho_b \\
&= e^c_\mu\partial_\nu e^\nu_a e^\mu_b - e^c_\mu\partial_\nu e^\mu_a e^\nu_b + e^c_\mu e^\nu_a\partial_\nu e^\mu_b - e^c_\mu e^\mu_a\partial_\nu e^\nu_b - \delta^c_b\partial_\rho e^\rho_a + \delta^c_a\partial_\rho e^\rho_b \\
&= \delta^c_b\partial_\nu e^\nu_a - e^c_\mu e^\nu_b\partial_\nu e^\mu_a + e^c_\mu e^\nu_a\partial_\nu e^\mu_b - \delta^c_a\partial_\nu e^\nu_b - \delta^c_b\partial_\rho e^\rho_a + \delta^c_a\partial_\rho e^\rho_b \\
&= e^c_\mu e^\nu_a\partial_\nu e^\mu_b - e^c_\mu e^\nu_b\partial_\nu e^\mu_a \\
&= e^c_\mu(\partial_a e^\mu_b - \partial_b e^\mu_a) \\
&= e^c_\mu t^\mu{}_{ab}.
\end{aligned} \tag{E62}$$

is given in terms of the translational field strength tensor introduced in Eq. (E41). Its contracted expression,

$$\begin{aligned}
t^b{}_{ba} &= e^b_\mu(\partial_b e^\mu_a - \partial_a e^\mu_b) \\
&= \partial_\mu e^\mu_a - e^b_\mu e^\nu_a\partial_\nu e^\mu_b,
\end{aligned} \tag{E63}$$

inserted in Eq. (E61) gives

$$\begin{aligned}
\omega^c{}_{ab} &= \frac{1}{6}t^d{}_{da}\delta^c_b - \frac{1}{6}t^d{}_{db}\delta^c_a + \frac{1}{2}t^c{}_{ab} \\
&= \frac{1}{6}(\partial_\mu e^\mu_a - e^d_\mu e^\nu_a\partial_\nu e^\mu_d)\delta^c_b - \frac{1}{6}(\partial_\mu e^\mu_b - e^d_\mu e^\nu_b\partial_\nu e^\mu_d)\delta^c_a + \frac{1}{2}e^c_\mu e^\nu_a\partial_\nu e^\mu_b - \frac{1}{2}e^c_\mu e^\nu_b\partial_\nu e^\mu_a
\end{aligned} \tag{E64}$$

and

$$\begin{aligned}
\omega^{c}{}_{a\mu} &= \frac{1}{6}e^b_\mu\partial_\rho e^\rho_a\delta^c_b - \frac{1}{6}e^b_\mu e^d_\rho e^\nu_a\partial_\nu e^\rho_d\delta^c_b - \frac{1}{6}e^b_\mu\partial_\rho e^\rho_b\delta^c_a + \frac{1}{6}e^b_\mu e^d_\rho e^\nu_b\partial_\nu e^\rho_d\delta^c_a + \frac{1}{2}e^b_\mu e^c_\rho e^\nu_a\partial_\nu e^\rho_b - \frac{1}{2}e^b_\mu e^c_\rho e^\nu_b\partial_\nu e^\rho_a \\
&= \frac{1}{6}e^c_\mu\partial_\rho e^\rho_a - \frac{1}{6}e^c_\mu e^d_\rho e^\nu_a\partial_\nu e^\rho_d - \frac{1}{6}\delta^c_a e^b_\mu\partial_\rho e^\rho_b + \frac{1}{6}\delta^c_a e^d_\rho\partial_\mu e^\rho_d + \frac{1}{2}e^b_\mu e^c_\rho e^\nu_a\partial_\nu e^\rho_b - \frac{1}{2}e^c_\rho\partial_\mu e^\rho_a \\
&= \frac{1}{6}e^c_\mu e^b_\rho\partial_b e^\rho_a - \frac{1}{6}e^c_\mu e^d_\rho\partial_a e^\rho_d - \frac{1}{6}\delta^c_a e^b_\mu e^d_\rho\partial_d e^\rho_b + \frac{1}{6}\delta^c_a e^d_\rho e^b_\mu\partial_b e^\rho_d + \frac{1}{2}e^b_\mu e^c_\rho\partial_a e^\rho_b - \frac{1}{2}e^c_\rho e^b_\mu\partial_b e^\rho_a \\
&= \frac{1}{6}e^c_\mu e^b_\rho t^\rho_{ba} + \frac{1}{6}\delta^c_a e^b_\mu e^d_\rho t^\rho_{bd} + \frac{1}{2}e^b_\mu e^c_\rho t^\rho_{ab} \\
&= \frac{1}{6}(e^c_\mu\delta^d_a e^b_\rho + \delta^c_a e^b_\mu e^d_\rho - 3e^c_\rho\delta^d_a e^b_\mu)t^\rho_{bd}.
\end{aligned} \tag{E65}$$

Notice that the affine connection of the external space covariant derivative does not appear in this equation because the scalar curvature (E48) is expressed in terms of the field strength tensor of ω_μ.

b. Variation of e

We consider vierbein components e^a_μ as independent variables and the variation of

$$e^\nu_a e^a_\mu = \delta^\nu_\mu \tag{E66}$$

gives

$$\delta e^\nu_a e^a_\mu + e^\nu_a \delta e^a_\mu = 0 \tag{E67}$$

and

$$\delta e^\nu_b = -e^\mu_b e^\nu_a \delta e^a_\mu. \tag{E68}$$

The variational equation for the vierbein which appears in the tensor $T^{\mu\nu}_{ab}$, given by Eq. (E49) and in the determinant E is

$$0 = -16\pi G\frac{\delta S}{\delta e^b_\rho} = 16\pi G\frac{\delta S}{\delta e^\kappa_c}e^\kappa_b e^\rho_c = 16\pi G\left(E\frac{\delta R}{\delta T^{\mu\nu}_{ad}}\frac{\delta T^{\mu\nu}_{ad}}{\delta e^\kappa_c} + \frac{\delta E}{\delta e^\kappa_c}R\right)e^\kappa_b e^\rho_c \tag{E69}$$

what we can write by means of Eqs. (E48) and

$$\delta E = \delta e^{\mathrm{tr}\ln e} = e^{\mathrm{tr}\ln e}\mathrm{tr}e^{-1}\delta e = E e^\mu_a \delta e^a_\mu = -E e^a_\mu \delta e^\mu_a \tag{E70}$$

as

$$\begin{aligned}
0 &= \eta^{bd}(\partial_\mu\omega^{a}{}_{b\nu} + \omega^{a}{}_{e\mu}\omega^{e}{}_{b\nu})(\delta^\mu_\kappa\delta^c_a e^\nu_d + e^\mu_a\delta^\nu_\kappa\delta^c_d - \delta^\nu_\kappa\delta^c_a e^\mu_d - e^\nu_a\delta^\mu_\kappa\delta^c_d) - e^c_\kappa R \\
&= (\partial_\mu\omega^{ad}{}_\nu + \omega^{a}{}_{e\mu}\omega^{ed}{}_\nu)(\delta^\mu_\kappa\delta^c_a e^\nu_d + e^\mu_a\delta^\nu_\kappa\delta^c_d - \delta^\nu_\kappa\delta^c_a e^\mu_d - e^\nu_a\delta^\mu_\kappa\delta^c_d) - e^c_\kappa R \\
&= (\partial_\kappa\omega^{cd}{}_\nu + \omega^{c}{}_{e\kappa}\omega^{ed}{}_\nu)e^\nu_d + (\partial_\mu\omega^{ac}{}_\kappa + \omega^{a}{}_{e\mu}\omega^{ec}{}_\kappa)e^\mu_a - (\partial_\mu\omega^{cd}{}_\kappa + \omega^{c}{}_{e\mu}\omega^{ed}{}_\kappa)e^\mu_d \\
&\quad -(\partial_\kappa\omega^{ac}{}_\nu + \omega^{a}{}_{e\kappa}\omega^{ec}{}_\nu)e^\nu_a - e^c_\kappa R
\end{aligned} \tag{E71}$$

We contract the last equation with e_b^κ and write

$$\begin{aligned}
0 &= (\partial_\kappa \omega^{cd}{}_\nu + \omega^c{}_{e\kappa}\omega^{ed}{}_\nu)e_d^\nu e_b^\kappa + (\partial_\mu \omega^{ac}{}_\kappa + \omega^a{}_{e\mu}\omega^{ec}{}_\kappa)e_a^\mu e_b^\kappa \\
&\quad -(\partial_\mu \omega^{cd}{}_\kappa + \omega^c{}_{e\mu}\omega^{ed}{}_\kappa)e_d^\mu e_b^\kappa - (\partial_\kappa \omega^{ac}{}_\nu + \omega^a{}_{e\kappa}\omega^{ec}{}_\nu)e_a^\nu e_b^\kappa - \delta_b^c R \\
&= 2(e_d^\nu e_b^\kappa - e_d^\kappa e_b^\nu)(\partial_\kappa \omega^{cd}{}_\nu + \omega^c{}_{e\kappa}\omega^{ed}{}_\nu) - \delta_b^c R. \qquad \text{(E72)}
\end{aligned}$$

The Ricci tensor

$$\begin{aligned}
R_{bc} &= e_c^\nu e_a^\mu F^a{}_{b\mu\nu} \\
&= e_c^\nu e_a^\mu (\partial_\mu \omega^a{}_{b\nu} - \partial_\nu \omega^a{}_{b\mu} + \omega^a{}_{c\mu}\omega^c{}_{b\nu} - \omega^a{}_{c\nu}\omega^c{}_{b\mu}) \\
&= (e_c^\nu e_a^\mu - e_c^\mu e_a^\nu)(\partial_\mu \omega^a{}_{b\nu} + \omega^a{}_{c\mu}\omega^c{}_{b\nu}) \\
&= T_{ca}^{\nu\mu}(\partial_\mu \omega^a{}_{b\nu} + \omega^a{}_{c\mu}\omega^c{}_{b\nu}) \qquad \text{(E73)}
\end{aligned}$$

gives finally the Einstein equation

$$R_{ab} - \frac{1}{2}\eta_{ab} R = 0 \qquad \text{(E74)}$$

c. *Affine connection for world vectors*

The affine connection Γ_μ appearing in the covariant derivative $D^{(e)}$ has not been included in the action (E47) and it will be determined by a non-dynamical principle. The parallel transport of a vector v^μ during a displacement δx^μ, expressed by the equation

$$\delta x^\nu D_\nu^{(e)} v^\mu = 0 \qquad \text{(E75)}$$

must be equivalent with the similar equation expressing the parallel transport of the vector v^a,

$$\delta x^\nu D_\nu^{(i)} v^a = 0. \qquad \text{(E76)}$$

The covariant condition for the equivalence of the two parallel transports, valid for arbitrary vector field v^a, is

$$D_\nu^{(e)} e_a^\mu v^a = e_a^\mu D_\nu^{(i)} v^a. \qquad \text{(E77)}$$

The combination $e_a^\mu v^a$ is an internal space scalar,

$$D_\nu^{(e)} e_a^\mu v^a = D_\nu e_a^\mu v^a \qquad \text{(E78)}$$

thus we have

$$D_\nu e_a^\mu v^a = e_a^\mu D_\nu v^a, \qquad \text{(E79)}$$

or

$$D_\nu e_a^\mu = 0 \tag{E80}$$

which leads to metric admissibility,

$$D_\nu e_a^\mu \eta^{ab} e_b^\kappa = D_\nu g^{\mu\kappa} = 0 \tag{E81}$$

The affine connection, Γ, can easily be obtained by solving Eq. (E80),

$$D_\nu e_a^\mu = D_\nu^{(i)} e_a^\mu + \Gamma^\mu{}_{\rho\nu} e_a^\rho = 0 \tag{E82}$$

with the result

$$\begin{aligned}\Gamma^\mu{}_{\rho\nu} &= -e_\rho^a D_\nu e_a^\rho \\ &= e_a^\mu D_\nu e_\rho^a \\ &= e_a^\mu \partial_\nu e_\rho^a + e_a^\mu \omega_{\nu b}^a e_\rho^b \neq \Gamma_{\nu\rho}^\mu.\end{aligned} \tag{E83}$$