Theoretical Gravitation and Cosmology

Jacopo Tissino

2021-11-23

1 General Relativity basics

Wednesday 2021-11-10, compiled 2021-11-23

Course given by Andrea Maselli. There's lots of stuff to say, and the course cannot possibly be comprehensive.

The exam can be two things: there will be a list of "old" papers, literature by now, to study in more detail. We can prepare a presentation on this, to be shared among each other.

The alternative is to study more "book-like" topics: say, the TOV solution.

Three standard reference books are

1. Bernard Schultz, "A first course in GR";
2. Clifford and Will, "Theory and experiments of gravitational physics", which is more experimentally oriented;
3. Ferrari, Gualtieri and Pani, "GR and its applications", which is very complete.

The equivalence principles

There is a plural in the title. Principia impose constraints which modify the theory, not the other way around.

The thing to be careful with is whether we are testing *principia* or a specific theory. For example, the Pound-Rebka redshift experiment was testing the equivalence principle.

Alternative theories, such as scalar-tensor theories which were pioneered by Brams and Dicke, could also be written satisfying the same equivalence principle.

We can then classify alternative theories of gravity according to the principles they satisfy.

Newton Equivalence Principle This is also dubbed "NEP".

"The quantity that I mean hereafter by the name of mass…"

He distinguishes "mass" and "weight", and his formulation is not a principle yet.

> In the Newtonian limit, the inertial and gravitational masses are all proportional.

Newton's theory says that a density $\rho(x(t), t)$ will source a gravitational potential

$$\nabla^2 \phi(x(t), t) = 4\pi G \rho \,, \tag{1.1}$$

and particles will move according to the law

$$m\ddot{x}(t) = -m\nabla\phi(x(t), t) \,. \tag{1.2}$$

The density ρ and the mass m are conceptually different: $\rho = \rho_{\text{act}}$ is the *active* gravitational density, the mass $m = m_{\text{in}}$ multiplying $\ddot{x}$ is the inertial mass, while the $m = m_{\text{pass}}$ multiplying $\nabla\phi$ is the passive gravitational mass.

We can show that the active and passive gravitational masses are proportional by exploiting the third of Newton's laws: consider two objects 0 and 1, the attraction forces exerted by 0 on 1 and respectively by 1 on 0 will be

$$f_1 = m_1^{\text{in}} a_1 = \frac{G m_0^{\text{act}} m_1^{\text{pass}}}{r^2} \tag{1.3a}$$

$$f_0 = m_0^{\text{in}} a_0 = \frac{G m_1^{\text{act}} m_0^{\text{pass}}}{r^2} \,, \tag{1.3b}$$

but the third law imposes $f_0 = f_1$, which means

$$\frac{m_0^{\text{pass}}}{m_0^{\text{act}}} = \frac{m_1^{\text{pass}}}{m_1^{\text{act}}} \,. \tag{1.4}$$

Therefore, these are proportional. Setting the proportionality constant to 1 is just a matter of the units with which we measure them. Therefore, we can say $m_{\text{act}} = m_{\text{pass}} = m_{\text{grav}}$.

If we have one more body, 2, the force exerted by 0 will be

$$f_2 = \frac{G m_0^{\text{act}} m_2^{\text{pass}}}{r^2} \,, \tag{1.5}$$

so the accelerations of bodies 1 and 2 will be $a_1 = f_1 / m_1^{\text{in}}$ and $a_2 = f_2 / m_2^{\text{in}}$.

If these are equal, then we get $m_1^{\text{pass}} / m_1^{\text{in}} = m_2^{\text{pass}} / m_2^{\text{in}}$.

A way to quantify discrepancies from this principle is

$$\eta = \frac{\frac{m_1^{\text{grav}}}{m_1^{\text{in}}} - \frac{m_2^{\text{grav}}}{m_2^{\text{in}}}}{\frac{m_1^{\text{grav}}}{m_1^{\text{in}}} + \frac{m_2^{\text{grav}}}{m_2^{\text{in}}}} \,. \tag{1.6}$$

We know from experiments that $|\eta| < 10^{-13}$. In the Newtonian language, this is purely a coincidence.

Consider a nonrelativistic object moving in an external gravitational field, so that

$$m_{\text{in}} \ddot{x} = m_{\text{grav}} g + \underbrace{\sum_k f_k}_{F} \,, \tag{1.7}$$

where f_k are all other forces affecting this body. We are assuming that g is a constant.

We can make a transformation $y(t) = x(t) - gt^2/2$, $t' = t$. The equation of motion will be

$$m_{\text{in}}\ddot{y} = (m_{\text{grav}} - m_{\text{in}})g + F\,. \tag{1.8}$$

If the equivalence principle holds, the equation of motion becomes $m_{\text{in}}\ddot{y} = F$. In this reference system, gravity disappears completely.

We can see hints of the locality of this procedure: we are neglecting any derivatives of g, which will never be truly constant. This is the very meaning of a "local inertial observer".

How does this work in GR? (we are actually doing this in the simplest possible, SR terms)

Weak Equivalence Principle

> "The trajectory of an uncharged test particle in a point of spacetime is independent of that particle's structure and composition."

What "test particle" means is that it has negligible self-gravity, and that it is small enough that it does not couple to the inhomogeneities of the gravitational field.

The first statement can be quantified by the compactness parameter: GM/c^2R. For the Sun, this is of the order of 10^{-6}.

The second statement can be quantified by saying that the size of the body should be small compared to the length scale of the curvature of spacetime.

These two constraints are logically independent. In GR, the term "mass" is quite difficult to deal with. Therefore, instead of discussing them we only talk of trajectories.

"Test experiment" means that its effects are "weak".

Einstein Equivalence Principle There are two alternative formulations.

> "The outcome of any non-gravitational test experiment is not affected locally, at any point in spacetime, by the presence of the gravitational field."

> "The outcome of any non-gravitational test experiment is independent of the position of the lab in space and of the velocity of the free-falling apparatus."

The second formulation more explicitly requires invariance under the Poincaré group.

In the WEP we are only talking about mechanical laws, while here we extend to all the laws of physics.

This means that we can always find a reference system which cancels gravity.

This is establishing the connection between local frames in a gravitational field, and frames in the absence of gravity.

This sounds a bit circular without a careful definition of a "gravitational experiment"...

Strong equivalence principle

> "The outcome of any experiment is not affected at any point in spacetime by the presence of the gravitational field."

There is no proof that SEP implies GR, but GR surely satisfies it.

The EEP is what tells us that the theory must be a metric one, since it includes SR, which has the Minkowski tensor.

There must be (at least one) metric tensor, so that it can be made equal to the Minkowski tensor at each point in spacetime, up to a conformal transformation:

$$[g_1, g_2 \dots] \to [\phi_1(P)\eta, \phi_2(p)\eta, \dots] \,. \tag{1.9}$$

If this is the case, we will be able to rescale $\phi_i = C_i \phi(P)$, but then we can rescale our units so that $C_i = 1$, and the metric as $\overline{g} = \phi^{-1} g$, which finally yields the transformation $g \to \eta$.

In the end, then, there must be a reference frame such that

$$g_{\mu\nu}(P) = \eta_{\mu\nu} + \sum_\alpha \mathcal{O}\left(\left| x^\alpha - x^\alpha(P) \right|^2 \right) . \tag{1.10}$$

In a Local Lorentz Frame there is a family of preferred curves for $g_{\mu\nu}$, geodesics, which are straight lines: free-falling trajectories are straight lines for free-falling observers.

There are purely metric theories, with a single metric tensor, but also ones in which other fields mediate the interaction.

The presence of the LLF means that there is a frame such that, around a point, $\mathrm{d}s^2 = \eta_{\mu\nu} \, \mathrm{d}y^\mu \, \mathrm{d}y^\nu$.

What are the EoM for a free-falling particle? We know that they read $\frac{\mathrm{d}^2 y}{\mathrm{d}\tau^2} = 0$ in the LLF.

Suppose we then move to a different frame, $x^\alpha = x^\alpha(y)$: then the interval will transform as

$$\mathrm{d}s^2 = \underbrace{\eta_{\mu\nu} \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial y^\nu}{\partial x^\beta}}_{g_{\alpha\beta}} \mathrm{d}x^\alpha \, \mathrm{d}x^\beta \ . \tag{1.11}$$

The EoM then reads

$$0 = \frac{\mathrm{d}}{\mathrm{d}\tau} \left(\frac{\partial y^\alpha}{\partial x^\beta} \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} \right) = \frac{\mathrm{d}^2 x^\beta}{\mathrm{d}\tau^2} \frac{\partial y^\beta}{\partial x^\beta} + \frac{\partial^2 y^\alpha}{\partial x^\beta \partial x^\gamma} \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} \frac{\mathrm{d}x^\gamma}{\mathrm{d}\tau} \tag{1.12a}$$

$$= \frac{\mathrm{d}^2 x^\beta}{\mathrm{d}\tau^2} \underbrace{\frac{\partial y^\beta}{\partial x^\beta} \frac{\partial x^\rho}{\partial x^\beta}}_{\delta_\beta{}^\rho} + \underbrace{\left(\frac{\partial x^\rho}{\partial y^\alpha} \frac{\partial^2 y^\alpha}{\partial x^\beta \partial x^\gamma} \right)}_{\Gamma^\rho_{\beta\gamma}} \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} \frac{\mathrm{d}x^\gamma}{\mathrm{d}\tau} \tag{1.12b}$$

$$= \frac{\mathrm{d}^2 x^\rho}{\mathrm{d}\tau^2} + \Gamma^\rho_{\beta\gamma} \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} \frac{\mathrm{d}x^\gamma}{\mathrm{d}\tau} \,. \tag{1.12c}$$

This is the **Geodesic equation**. The crucial thing is that Γ vanishes when the spacetime is flat.

The WEP implies the NEP. On the other hand, the NEP does not imply the WEP. A theory in which the equations of motion contain weird stuff and not just m_I / m_G satisfies the NEP and not the WEP.

Differential geometry

Monday 2021-11-15

A point in $\mathbb{R}^n$ is described by the n-tuple $(x_1, \dots, x_n)$. These are a dense set (and also a complete one).

Open sets in $\mathbb{R}^n$ are defined as follows: $S \subseteq \mathbb{R}^n$ is open if for all $x \in S$ we can make a ball $B_r(x)$ with $r > 0$ such that $B_x \subseteq S$, where a ball is a set

$$B_r(x) = \left\{ y \in \mathbb{R}^n \middle| |x - y| < r \right\} . \tag{1.13}$$

A topological space is a collection of open sets which is closed under arbitrary intersections and finite unions.

A map ρ between two sets M and N takes a point $x \in M$ to a point $y = f(x) \in N$.

If we consider a set $S \subseteq M$ we can look at the set of the images of S under f, denoted as $T = \rho(S) \subseteq N$. We can also define the inverse image, $S = \rho^{-1}(T)$.

Composition of maps between different sets can also be defined: suppose we have three spaces M, N, Z, with some maps $f \colon M \to N$ and $g \colon N \to Z$; their composition will be $g \circ f \colon M \to Z$.

A map of M is **into** N if all points of M are mapped to N; a map of M is **onto** N if all points of N have inverse maps to M.

A **continuous** map f between topological spaces M and N is one for which, given any point $x \in M$ such that $y = f(x) \in N$, there is an open set of N containing $f(x)$, which is the image of an open set of M.

This is less stringent than the ϵ-δ definition of continuity because there is no size requirement…??

A differentiable map is defined as follows: a function $f(x)$, where $x \in S$ and $S \subseteq \mathbb{R}^n$ is open, is of order $\mathcal{C}^k$ if all partial derivatives of order up to k exist and are continuous.

A **manifold** M is a collection of points such that each of them has an open neighborhood which has a continuous one-to-one map with an open set of $\mathbb{R}^n$. The number n is called the **dimension** of the manifold.

He talks about the locality of this definition as being about the fact that M is a subset of a larger space…

A coordinate system or *chart* is a pair (M, ρ), where M is an open set while ρ is a map.

Suppose we have two coordinate systems (U, f) and (V, g), where U and V overlap at least at a point (so, in a region as well).

Consider the region $f(U \cap V)$: in order to move from it to $g(U \cap V)$ we need to apply the map $g \circ f^{-1}$. This is, therefore, a **change of coordinates**, which we will be able to write as $\vec{y} = \vec{y}(\vec{x})$.

A $\mathcal{C}^k$ differentiable **manifold** is one for which

1. each point of M belongs at least to an open set (and a corresponding chart);
2. each chart is $\mathcal{C}^k$ related to other charts it overlaps with.

These changes of coordinates can be written as $y^i = F^i(x^j)$, for which we can define a Jacobian determinant:

$$J = \left| \frac{\partial F^i}{\partial x^j} \right| . \tag{1.14}$$

If $J \neq 0$ at a point P, there is an onto, one-to-one map in some neighborhood of P.

A **curve** is a path in M such that we associate a number, called the parameter, to each point in it, and this provides a map between it and a path in $\mathbb{R}^n$, called the image of the curve.

This is typically denoted as $\gamma \colon s \in [a, b] \to [x^1(s), \dots, x^\mu(s)]$.

The quantities

$$\frac{\mathrm{d}x^i}{\mathrm{d}s} = \left[\begin{array}{ccc} \frac{\mathrm{d}x^1}{\mathrm{d}s} & \dots & \frac{\mathrm{d}x^n}{\mathrm{d}s} \end{array} \right] \tag{1.15}$$

are the *components of the tangent vector*.

Suppose we want to make a coordinate change, $x'(x)$: how does our tangent vector change? By the chain rule, it will read

$$\frac{\mathrm{d}x'^i}{\mathrm{d}s} = \frac{\partial x'^i}{\partial x^j} \frac{\mathrm{d}x^j}{\mathrm{d}s} . \tag{1.16}$$

This kind of vector is called a **contravariant vector**, since it transforms with the Jacobian matrix.

We want to show that directional derivatives form a vector space at a point P. Consider a curve with its parameter λ and a differentiable function $\phi(x_1, \dots, x_m)$. The derivative

$$\frac{\mathrm{d}\phi}{\mathrm{d}\lambda} = \frac{\partial \phi}{\partial x^i} \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} . \tag{1.17}$$

With this in mind, we can define the **directional derivative operator**

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} = \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\partial}{\partial x^i} . \tag{1.18}$$

However, we already know that the $\mathrm{d}x^i \big/ \mathrm{d}\lambda$ are the *components of the tangent vector*.

If we have two curves, $x^i(\lambda)$ and $x^i(s)$, we can define a directional derivative for each of them. We can then sum these, by defining the sum as

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} + \frac{\mathrm{d}}{\mathrm{d}s} = \left(\frac{\mathrm{d}x^i}{\mathrm{d}\lambda} + \frac{\mathrm{d}x^i}{\mathrm{d}s} \right) \frac{\partial}{\partial x^i} . \tag{1.19}$$

This vector will still be tangent to the curve, and we will be able to find a third parameter μ such that

$$\frac{\mathrm{d}}{\mathrm{d}\mu} = \frac{\mathrm{d}x^i}{\mathrm{d}\mu} \frac{\partial}{\partial x^i} . \tag{1.20}$$

We can also scale the directional derivative operator:

$$a\frac{\mathrm{d}}{\mathrm{d}\lambda} = \underbrace{\left(a\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\right)}_{\mathrm{d}x^i/\mathrm{d}\sigma}\frac{\partial}{\partial x^i}\,. \tag{1.21}$$

This proves (?) that the space of directional derivatives is a vector space. This space is denoted as T_P, where $\mathrm{d}/\mathrm{d}\lambda$ is a vector.

The crucial idea here is that vectors at different points belong to different vector spaces, and cannot be directly compared.

Coordinate lines are the ones for which all but one of the coordinates we are using remain constant. What this means is that directional derivatives along these read

$$\frac{\mathrm{d}}{\mathrm{d}x^i} = \frac{\mathrm{d}x^j}{\mathrm{d}x^i}\frac{\partial}{\partial x^j} = \delta^j_i\frac{\partial}{\partial x^j} = \frac{\partial}{\partial x^i}\,. \tag{1.22}$$

This also means that any directional derivatives can be expressed (uniquely) as linear combinations of these ∂_i, which are therefore a basis for the tangent space at each point.

We have a one-to-one connection between the tangents of curves at a point P and the derivatives along a curve at P. Because of this, we associate the directional derivative $\mathrm{d}/\mathrm{d}\lambda$ to a tangent vector to a curve $x^i(\lambda)$.

We denote the i-th basis vector as $\vec{e}_{(i)} = \frac{\partial}{\partial x^i}$. The number between round brackets enumerates the vectors in the basis, it is not a spatial index.

We can then express any vector in a more conventional notation as $\vec{A} = A^i\vec{e}_{(i)}$. If we change coordinates, the basis will shift like

$$\vec{e}_{k'} = \widetilde{\Lambda}^j_{k'}\vec{e}_j\,, \tag{1.23}$$

where $\widetilde{\Lambda}$ is the inverse of the Jacobian.

One-forms A one-form $\widetilde{q}$ is a real-valued, linear function of vectors:

$$\widetilde{q}(a\vec{V}+\vec{W}) = a\widetilde{q}(\vec{V}) + \widetilde{q}(\vec{W})\,. \tag{1.24}$$

We can sum one-forms by summing their action, and multiply them by scalars by multiplying the result of their application by a scalar. These are then a vector space, which is called the *dual*: T^*_P.

The reason for the name is the dual symmetry: taking $\widetilde{q}(\vec{V}) \simeq \vec{V}(\widetilde{q})$.

The basis for the dual is called the conjugate basis $\widetilde{\omega}$, and it is convenient to select it so that $\widetilde{\omega}^{(i)}(\vec{V}) = V^i$. Alternatively, we can write this as $\widetilde{\omega}^{(i)}(\vec{e}_{(j)}) = \delta^i_j$.

Any 1-form can also be written in terms of the basis covectors, $\widetilde{q} = q_i\widetilde{\omega}^{(i)}$.

We can make the manipulation

$$\widetilde{q}(\vec{V}) = \widetilde{q}(V^i\vec{e}_{(i)}) = V^i\widetilde{q}(\vec{e}_{(i)}) = \widetilde{\omega}^{(i)}(\vec{V})\widetilde{q}(\vec{e}_{(i)}) = q^iV_i\,. \tag{1.25}$$

The conjugate basis is often denoted as $\widetilde{\omega}^{(J)} = \mathrm{d}\widetilde{x}^{(J)}$. Wednesday 2021-11-17

The application of a one-form on a vector, $\widetilde{q}(\vec{V}) = q_J V^J$, can be computed as a *contraction*. A one-form transforms with $\partial x^\mu / \partial x'^\alpha$, a vector transforms with $\partial x'^\mu / \partial x^\alpha$.

Since a basis vector transforms like

$$\vec{e}_{(\alpha')} = \Lambda^\mu{}_{\alpha'} \vec{e}_{(\mu)} \tag{1.26}$$

we have

$$q_J = \widetilde{q}(\vec{e}_{(J)}) = \widetilde{q}(\vec{e}_{(k')} \Lambda^{k'}{}_J) = \Lambda^{k'}{}_J \widetilde{q}(\vec{e}_{(k')}) = \Lambda^{k'}_J q_{k'} \,. \tag{1.27}$$

This shows that the one-form transforms with the inverse of the Jacobian.

An example of a one-form is the gradient of a scalar function ϕ:

$$\frac{\partial \phi}{\partial x^{(J')}} = \frac{\partial \phi}{\partial x^k} \frac{\partial x^k}{\partial x'^{J'}} = \frac{\partial \phi}{\partial x^k} \Lambda^k{}_{J'} \,. \tag{1.28}$$

Tensors A tensor of type (N, R) is a function which takes N one-forms, R vectors, and yields a number. We require that it is multilinear (linear in all its arguments),

One-forms are $(0, 1)$ tensors, vectors are $(1, 0)$ tensors.

The components of the tensor are defined by its application to the basis vectors / covectors of the vector space. For a $(0, 2)$ tensor, we have

$$F_{\alpha\beta} = F(\vec{e}_{(\alpha)}, \vec{e}_{(\beta)}) \,, \tag{1.29}$$

from which we can recover by linearity any application of the tensor: $F(\vec{A}, \vec{B}) = A^\alpha B^\beta F_{\alpha\beta}$.

We can then try to construct a basis for the tensor space such that

$$F = F_{\alpha\beta} \widetilde{\omega}^{(\alpha,\beta)} \,. \tag{1.30}$$

We do this by writing

$$F(\vec{A}, \vec{B}) = \left(F_{\alpha\beta} \widetilde{\omega}^{(\alpha,\beta)} \right) (\vec{A}, \vec{B}) = F_{\alpha\beta} \widetilde{\omega}^{\alpha}(\vec{A}) \widetilde{\omega}^{(\beta)}(\vec{B}) \,, \tag{1.31}$$

therefore we can write the $(0, 2)$ basis as an outer product:

$$\widetilde{\omega}^{(\alpha,\beta)} = \widetilde{\omega}^{(\alpha)} \otimes \widetilde{\omega}^{(\beta)} \,. \tag{1.32}$$

We can iterate this procedure to get a basis for any tensor space. How do tensors transform?

$$F = F_{\alpha'\beta'} \Lambda^{\alpha'}_\gamma \widetilde{\omega}^{(\gamma)} \otimes \widetilde{\omega}^{(\rho)} \Lambda^{\beta'}_\gamma = F_{\gamma\rho} \widetilde{\omega}^{(\gamma)} \otimes \widetilde{\omega}^{(\rho)} \,. \tag{1.33}$$

this looks wrong...

The transformation law is therefore

$$F_{\alpha'\beta'} \Lambda^{\alpha'}{}_\gamma \Lambda^{\beta'}{}_\delta = F_{\gamma\delta} \,, \tag{1.34}$$

or equivalently

$$F_{\alpha'\beta'} = \Lambda^{\rho}{}_{\alpha'}\Lambda^{\sigma}{}_{\beta'}F_{\rho\sigma}\,. \tag{1.35}$$

Tensors of the same kind can be added, and tensors of any order can be multiplied and contracted.

Tensors can be symmetric: in an algebraic sense, F is symmetric if $F(\vec{A},\vec{B}) = F(\vec{B},\vec{A})$. In terms of its indices, this means $F_{\alpha\beta} = F_{\beta\alpha}$.

We can symmetrize a tensor by $F_{(\alpha\beta)} = (F_{\alpha\beta} + F_{\beta\alpha})/2$.

Symmetric n-dimensional tensors have $n(n+1)/2$ independent components.

Tensors can be antisymmetric: in an algebraic sense, F is antisymmetric if $F(\vec{A},\vec{B}) = -F(\vec{B},\vec{A})$. In terms of its indices, this means $F_{\alpha\beta} = -F_{\beta\alpha}$.

We can antisymmetrize a tensor by $F_{[\alpha\beta]} = (F_{\alpha\beta} - F_{\beta\alpha})/2$.

Antisymmetric n-dimensional tensors have $n(n-1)/2$ independent components.

The **metric tensor**! We call its application to two vectors their scalar product, $g(\vec{A},\vec{B}) = \vec{A}\cdot\vec{B}$. It is symmetric, therefore it has 10 free components.

The distance between two points separated by an infinitesimal $\mathrm{d}s$ is

$$\mathrm{d}s^2 = \mathrm{d}s\cdot\mathrm{d}s = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu\ . \tag{1.36}$$

A curve $\gamma\colon [a,b] \to \mathcal{M}$ can be measured: its length will be

$$s = \int_a^b \mathrm{d}s = \int_a^b \mathrm{d}\lambda\,\underbrace{\sqrt{g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}}_{\mathrm{d}s/\mathrm{d}\lambda}\,. \tag{1.37}$$

This gives the length of the **path**, which is invariant with respect to the parametrization.

Thanks to the metric tensor we can lower or raise indices: the map $\omega_V\colon U \to g(U,V)$ is a one-form by the properties we require of the metric, so we can say it is the "dual" form of V, typically just denoted as $V_\mu = g_{\mu\nu}V^\nu$.

In order to do the inverse, we need the inverse of the metric tensor, $g^{\mu\nu}$, which satisfies $g^{\mu\nu}g_{\nu\rho} = \delta^\mu_\rho$.

Consider the metric tensor

$$\eta_{\alpha\beta} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\,. \tag{1.38}$$

Let us rotate the coordinate system:

$$x^0 = x^{0'} \tag{1.39}$$
$$x^1 = r\cos\theta \tag{1.40}$$
$$x^2 = r\sin\theta\,, \tag{1.41}$$

where r and θ are polar coordinates for the unprimed system.

How does the metric transform?

$$
\begin{aligned}
g_{0'0'} &= \Lambda^{\mu}_{0'}\Lambda^{\nu}_{0'}\eta_{\mu\nu} = \eta_{00} = -1 && (1.42)\\
g_{0'i'} &= \Lambda^{\mu}_{0'}\Lambda^{\nu}_{i'}\eta_{\mu\nu} = 0 && (1.43)\\
g_{1'1'} &= \Lambda^{\mu}_{1'}\Lambda^{\nu}_{1'}\eta_{\mu\nu} = \cos^2\theta + \sin^2\theta = 1 && (1.44)\\
g_{2'2'} &= \Lambda^{\mu}_{2'}\Lambda^{\nu}_{2'}\eta_{\mu\nu} = r^2(\cos^2\theta + \sin^2\theta) = r^2\,. && (1.45)
\end{aligned}
$$

So, we find the line element in polar coordinates

$$
g_{\alpha'\beta'} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & r^2 \end{bmatrix}. \tag{1.46}
$$

Covariant differentiation

Monday 2021-11-22

Consider a vector $\vec{V} = V^{\alpha}\vec{e}_{\alpha}$.

How do we compute $\partial\vec{V} \big/ \partial x^{\mu}$? It will read

$$
\frac{\partial\vec{V}}{\partial x^{\beta}} = \frac{\partial V^{\alpha}}{\partial x^{\beta}}\vec{e}_{(\alpha)} + V^{\alpha}\frac{\partial\vec{e}_{(\alpha)}}{\partial x^{\beta}}\,. \tag{1.47}
$$

What we need to argue now is that the second term is a vector just like the first one: we need to use the equivalence principle. We can move to a reference in which $g_{\mu\nu} = \eta_{\mu\nu}$, such that $\vec{e}_{(\alpha)}$ also become constant.

The new and old will be related through transformation matrices:

$$
\vec{e}_{(\alpha')} = \Lambda^{\mu}_{\alpha'}\vec{e}_{\mu}\,, \tag{1.48}
$$

so we will get

$$
\frac{\partial}{\partial x^{\beta}}\vec{e}_{(\alpha)} = \frac{\partial}{\partial x^{\beta}}\left[\Lambda^{\mu'}_{\alpha}\right]\vec{e}_{(\mu')}\,. \tag{1.49}
$$

This object can be expressed as a linear combination of the basis vectors in the new basis, and is therefore a vector. We call it

$$
\frac{\partial\vec{e}_{(\alpha)}}{\partial x^{\beta}} = \Gamma^{\rho}_{\alpha\beta}\vec{e}_{(\rho)}\,. \tag{1.50}
$$

This is an *affine connection*. With this expression, we write

$$
\frac{\partial\vec{V}}{\partial x^{\beta}} = \frac{\partial V^{\alpha}}{\partial x^{\beta}} + V^{\alpha}\Gamma^{\rho}_{\alpha\beta}\vec{e}_{(\rho)}\,. \tag{1.51}
$$

The formalism is as follows:

$$
V^{\alpha}{}_{,\mu} = \partial_{\mu}V^{\alpha} = \frac{\partial V^{\alpha}}{\partial x^{\mu}} \tag{1.52}
$$

$$V^{\alpha}{}_{;\mu} = \nabla_{\mu} V^{\alpha} = \frac{\partial V^{\alpha}}{\partial x^{\mu}} + V^{\beta} \Gamma^{\alpha}_{\mu\beta}\,. \tag{1.53}$$

The covariant derivative can also be written as

$$\nabla \vec{V} = \left[V^{\alpha}{}_{;\mu} \vec{e}_{(\alpha)} \right] \otimes \widetilde{\omega}^{(\mu)}\,. \tag{1.54}$$

Therefore, $\nabla \vec{V}$ is a $(1,1)$ tensor.

In a Local Inertial Frame, covariant and component derivatives are equal, since $\Gamma^{\mu}_{\nu\rho} = 0$. Also, the covariant derivative of a scalar is equal to its partial derivative.

We can also differentiate a one-form: this will be a $(0,2)$ tensor. Its components will read

$$q_{\alpha;\beta} = \partial_{\beta} q_{\alpha} - \Gamma^{\rho}_{\alpha\beta} q_{\rho}\,. \tag{1.55}$$

When taking the covariant derivative of a tensor with any amount of indices we need to include a Christoffel term for all indices.

Let us write out a mixed tensor's derivative:

$$\nabla_{\beta} T^{\mu}_{\nu} = T^{\mu}_{\nu,\beta} + \Gamma^{\mu}_{\rho\beta} T^{\rho}_{\nu} - \Gamma^{\rho}_{\nu\beta} T^{\mu}_{\rho}\,. \tag{1.56}$$

In GR, we always have $g_{\mu\nu;\rho} = 0$. Why is this? If we move to a LIF, it equals $\eta_{\mu\nu,\rho} = 0$.

The Christoffel symbols are also symmetric: $\Gamma^{\rho}_{\alpha\beta} = \Gamma^{\rho}_{(\alpha\beta)}$. There is an explicit expression for them in terms of the metric:

$$\Gamma^{\rho}_{\alpha\beta} = \frac{1}{2} g^{\rho\sigma} \Big(g_{\sigma\alpha,\beta} + g_{\sigma\beta,\alpha} - g_{\alpha\beta,\sigma} \Big)\,. \tag{1.57}$$

Are the Christoffel symbols the components of a tensor? No. They vanish in the LIF, so if they were a tensor they would always be zero. If one tries to compute their transformation law, it comes out to depend on the mixed second derivatives of the coordinate change.

The fact that in the LIF $g_{\mu\nu,\alpha} = 0$ and $\Gamma^{\rho}_{\alpha\mu} = 0$ is the mathematical representation of the equivalence principle.

Parallel transport It helps us understand the intrinsic geometry of the manifold. We will do it for a vector along a path. We ask that, for each infinitesimal displacement, the displaced vector remains parallel to itself and does not change magnitude.

This means that we cannot, in general, define a *global* constant vector field, or "comb a hairy ball".

Suppose we have a curve parametrized by λ, whose tangent vector is

$$u^{\mu} = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}\lambda}\,. \tag{1.58}$$

We want to transport a vector $\vec{V}$ along it. How do we do it? We can move to a LIF, with coordinates ξ^{α}. If we move $\vec{V}$ by $\mathrm{d}\lambda$ along the curve, we will change it by

$$\frac{\mathrm{d}V^{\mu}}{\mathrm{d}\lambda} = \frac{\mathrm{d}V^{\mu}}{\mathrm{d}\xi^{\alpha}} \frac{\mathrm{d}\xi^{\alpha}}{\mathrm{d}\lambda} = u^{\alpha} \partial_{\alpha} V^{\mu}\,. \tag{1.59}$$

However, $\partial = \nabla$ in a LIF, so this non-tensorial expression is equal in this reference frame to the tensorial expression $u^\alpha \nabla_\alpha V^\mu$, which we set to zero in order to impose that the vector does not change.

This is often also written like $\nabla_{\vec{u}} \vec{V} = 0$, or

$$\nabla_{\vec{u}} V^\alpha = \frac{\mathrm{d}V^\alpha}{\mathrm{d}\lambda} + V^\mu u^\beta \Gamma^\alpha_{\mu\beta} = 0 \,. \tag{1.60}$$

In the presence of a curved spacetime, the components of the vector do change, depending on the curve!

Geodesics are curves which parallel-transport their own tangent vector: $u^\mu \nabla_\mu u^\nu = 0$. This means that

$$u^\beta u^\alpha_{;\beta} = \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \left[\frac{\partial x^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\beta\rho} u^\rho \right] = \frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\beta\rho} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda} \,. \tag{1.61}$$

Any affine transformation $\lambda \to a\lambda + b$ leaves this unchanged.

Can't we do any monotone differentiable transformation?

The curvature tensor

It is the only four-index tensor which can be built only from derivatives of the metric.

A way to do it is to try and build an object which transforms like a tensor from the (derivatives of the) Christoffel symbols.

The Riemann tensor is written as

$$R^\lambda_{\mu\nu\kappa} = -\left(\partial_\alpha \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\kappa} + \Gamma^\lambda_{\kappa\eta} \Gamma^\eta_{\mu\nu} - \Gamma^\lambda_{\nu\eta} \Gamma^\eta_{\mu\kappa} \right) . \tag{1.62}$$

It can be schematically represented as $R \sim \partial\Gamma + \Gamma\Gamma$. In the LIF, $\Gamma\Gamma = 0$ but the first part does not vanish; this allows it to be tensorial.

In that case, it is also written as

$$R^\alpha_{\beta\mu\nu} = \frac{1}{2} g^{\alpha\sigma} \left(g_{\sigma\nu,\beta\mu} - g_{\sigma\mu,\beta\nu} + g_{\beta\mu,\sigma\nu} - g_{\beta\nu,\sigma\mu} \right) . \tag{1.63}$$

Let us write out its symmetries in the $(0,4)$ version:

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta} \,, \tag{1.64}$$

and also

$$R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} = 0 \,. \tag{1.65}$$

These reduce the number of independent components: it would have $4^4 = 256$, but the symmetries reduce that number to 20.

If we parallel-transport a vector along a loop of coordinate directions 1 and 2 it changes by

$$\delta V^\alpha = R^\alpha_{\beta 12} V^\beta \,. \tag{1.66}$$

So, we can see that the spacetime is globally flat iff $R^{\alpha}_{\beta\mu\nu} = 0$.
It can also be defined as

$$\left[\nabla_{\mu}, \nabla_{\nu}\right] V^{\alpha} = R^{\alpha}_{\beta\mu\nu} V^{\beta} . \tag{1.67}$$

We will also need the Bianchi identities:

$$R_{\alpha\beta[\mu\nu;\lambda]} = 0 . \tag{1.68}$$

Exercise: Einstein's carousel.
The transformation law to and from flat spacetime is according to the matrix

$$L^{\mu}_{\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\omega t) & -\sin(\omega t) & 0 \\ 0 & \sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \tag{1.69}$$

so that $\xi^{\mu} = L^{\mu}_{\nu} x^{\nu}$, and $x^{\mu} = \widetilde{L}^{\mu}_{\nu} \xi^{\nu}$, where $\widetilde{L}$ is the inverse of L, which can be obtained by mapping $t \to -t$ in it.
Now, we need to compute the derivatives of this transformation in the expression for the Christoffel symbols:

$$\Gamma^{\mu}_{\alpha\beta} = \frac{\partial^2 \xi^{\rho}}{\partial x^{\alpha} \partial x^{\beta}} \frac{\partial x^{\mu}}{\partial \xi^{\rho}} . \tag{1.70}$$

The transformation matrix $\widetilde{L}^{\mu}_{\nu}$ depends on ξ only through its component $\xi^0 = t$:

$$\frac{\partial x^{\mu}}{\partial \xi^{\rho}} = \frac{\partial}{\partial \xi^{\rho}} \left(\widetilde{L}^{\mu}_{\nu} \xi^{\nu} \right) \tag{1.71}$$

$$= \widetilde{L}^{\mu}_{\nu} \delta^{\nu}_{\rho} + \frac{\partial \widetilde{L}^{\mu}_{\nu}}{\partial \xi^{\rho}} \xi^{\nu} \tag{1.72}$$

$$= \widetilde{L}^{\mu}_{\rho} + \dot{\widetilde{L}}^{\mu}_{\nu} \xi^{\nu} \delta^{t}_{\rho} , \tag{1.73}$$

and the reciprocal transformation is quite similar:

$$\frac{\partial \xi^{\rho}}{\partial x^{\beta}} = \frac{\partial}{\partial x^{\beta}} \left(L^{\rho}_{\sigma} x^{\sigma} \right) \tag{1.74}$$

$$= L^{\rho}_{\beta} + \dot{L}^{\rho}_{\sigma} \xi^{\sigma} \delta^{t}_{\beta} . \tag{1.75}$$

Now, however, we need to take a second derivative of this term:

$$\frac{\partial^2 \xi^{\rho}}{\partial x^{\alpha} \partial x^{\beta}} = \frac{\partial}{\partial x^{\alpha}} \left(L^{\rho}_{\beta} + \dot{L}^{\rho}_{\sigma} \xi^{\sigma} \delta^{t}_{\beta} \right) \tag{1.76}$$

$$= \dot{L}^{\rho}_{\beta} \delta^{t}_{\alpha} + \dot{L}^{\rho}_{\sigma} \delta^{\sigma}_{\alpha} \delta^{t}_{\beta} + \ddot{L}^{\rho}_{\sigma} \delta^{t}_{\alpha} \xi^{\sigma} \delta^{t}_{\beta} \tag{1.77}$$

$$= 2 \dot{L}^{\rho}_{(\alpha} \delta^{t}_{\beta)} + \ddot{L}^{\rho}_{\sigma} \xi^{\sigma} \delta^{t}_{\alpha} \delta^{t}_{\beta} , \tag{1.78}$$

which we can combine into the full Christoffel symbol expression, which is (thankfully) manifestly symmetric in $\alpha\beta$:

Terms like $\dot{L}^\rho \delta^t_\rho$ vanish!

$$\Gamma^\mu_{\alpha\beta} = \left(\dot{L}^\rho_\beta \delta^t_\alpha + \dot{L}^\rho_\alpha \delta^t_\beta + \ddot{L}^\rho_\sigma \xi^\sigma \delta^t_\alpha \delta^t_\beta\right)\left(\widetilde{L}^\mu_\rho + \dot{\widetilde{L}}^\mu_\nu \xi^\nu \delta^t_\rho\right) \tag{1.79}$$

$$= \dot{L}^\rho_\beta \delta^t_\alpha \widetilde{L}^\mu_\rho + \dot{L}^\rho_\alpha \delta^t_\beta \widetilde{L}^\mu_\rho + \ddot{L}^\rho_\sigma \widetilde{L}^\mu_\rho \xi^\sigma \delta^t_\alpha \delta^t_\beta\,, \tag{1.80}$$

which is already starting to look like Coriolis + centrifugal terms.
We can write

$$\ddot{L}^\rho_\sigma = -\omega^2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \cos(\omega t) & -\sin(\omega t) & 0 \\ 0 & \sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = -\omega^2\left(L^\rho_\sigma - \delta^\rho_t \delta^t_\sigma - \delta^\rho_z \delta^z_\sigma\right), \tag{1.81}$$

using which we can rewrite the centrifugal term like

$$\ddot{L}^\rho_\sigma \widetilde{L}^\mu_\rho \xi^\sigma \delta^t_\alpha \delta^t_\beta = -\omega^2\left(\delta^\mu_\sigma \xi^\sigma \delta^t_\alpha \delta^t_\beta - \delta^\rho_t \delta^t_\sigma \widetilde{L}^\mu_\rho \xi^\sigma \delta^t_\alpha \delta^t_\beta - \delta^\rho_z \delta^z_\sigma \widetilde{L}^\mu_\rho \xi^\sigma \delta^t_\alpha \delta^t_\beta\right) \tag{1.82}$$

$$= \omega^2 \xi^\mu \delta^t_\alpha \delta^t_\beta = \omega^2 L^\mu_\nu x^\nu \delta^t_\alpha \delta^t_\beta\,. \tag{1.83}$$

Let us now move to the Coriolis term: we will get identical results if we swap α and β, and we know that one of them must be t, so let us set $\alpha = t$: we get

$$\left(\Gamma^\mu_{t\beta}\right)_{\text{Coriolis}} = \dot{L}^\rho_\beta L^\mu_\rho \tag{1.84}$$

$$= \omega \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -\sin(\omega t) & -\cos(\omega t) & 0 \\ 0 & -\cos(\omega t) & \sin(\omega t) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}^\rho_\beta \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\omega t) & \sin(\omega t) & 0 \\ 0 & -\sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^\mu_\rho \tag{1.85}$$

$$= \omega \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}^\mu_\beta \tag{1.86}$$

$$= \omega \epsilon^{\mu z}_{t\beta}\,, \tag{1.87}$$

where the ϵ is the four-dimensional Levi-Civita symbol. This also shows that the symmetric $\beta = t$ term equals $\omega \epsilon^{\mu z}_{\alpha t}$.
We can thus compactly write

$$\Gamma^\mu_{\alpha\beta} = \omega\left(\epsilon^{\mu z}_{\alpha t} \delta^t_\beta + \epsilon^{\mu z}_{t\beta} \delta^t_\alpha\right) + \omega^2 L^\mu_\nu x^\nu \delta^t_\alpha \delta^t_\beta\,. \tag{1.88}$$

Now, let us write the geodesic equation:

$$\ddot{x}^\mu = -\Gamma^\mu_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta \tag{1.89}$$

$$= -\left[\omega\left(\epsilon^{\mu z}_{\alpha t}\delta^t_\beta + \epsilon^{\mu z}_{t\beta}\delta^t_\alpha\right) + \omega^2 L^\mu_\nu x^\nu \delta^t_\alpha \delta^t_\beta\right]\dot{x}^\alpha \dot{x}^\beta \,. \tag{1.90}$$

The sum includes several terms: for example, the $\alpha = \beta = t$ one reads

$$-\omega^2 L^\mu_\nu x^\nu (\dot{x}^t)^2 \,. \tag{1.91}$$

Let us look at this more concretely: the four-velocity of a timelike trajectory is $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$; which in terms of the three-velocity $\vec{v}$ reads $u^\mu = [\gamma, \gamma\vec{v}]$, where $\gamma = 1/\sqrt{1-|\vec{v}|^2}$.
In our case, $\dot{x}^\mu = u^\mu$.
Let us then look at the $\mu = i$, spatial components of this equation:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\gamma v^i\right) = -\omega\left[\epsilon^{iz}_{jt}u^j + \epsilon^{iz}_{tj}u^j\right]u^t - \omega^2 \xi^i (u^t)^2 \tag{1.92}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(\gamma v^x) = -2\omega u^y u^t - \omega^2 \xi^x (u^t)^2 \tag{1.93}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(\gamma v^y) = +2\omega u^x u^t - \omega^2 \xi^y (u^t)^2 \,, \tag{1.94}$$