General Relativity

Maso[*]

November 4, 2023

Contents

Lecture 1

lun 02 ott 2023 14:30

Exam: three easy questions seen during the course, three questions somewhat new, two harder questions that require critical thinking.

1 Introduction

General relativity replaces the Newtonian description of gravity. The force of Newtonian gravity is proportional to the reciprocal of the distance squared between two masses. Changing the distance changes the force immediately, but this is not consistent with Special Relativity.

The problems with Newtonian gravity can be solved using field theoretical methods to construct theories in agreement with Special Relativity. Though, Einstein took a different route. He identified something often overlooked in what was already known: the equivalence principle. Einstein wanted to make sure that Maxwell's equations were in agreement with the principle of relativity. Newtonian mechanics was made possible by the equivalence principle: in its basic form, it states that gravitational mass — that is the mass that appears in Newton's law of gravitation — is equal to inertial mass — that is the mass that appears in Newton's second law of motion. A priori, these two masses can be different, but the equivalence principle postulates their equality. One can study the implications of the principle. Two thought experiments (gedankenexperimente) can help understand such implications.

[*] `https://github.com/M-a-s-o/notes`

First thought experiment. Consider a mass m inside a box with two propellers at the bottom. The box is accelerated upwards at a constant acceleration g equal to the sea-level Earth's acceleration. In the reference frame of the box, there is a downward force pushing the mass towards and then against the floor of the box. This behaviour is the exact same experienced on the surface of the Earth: holding a mass and letting it go, it experiences a downward force towards the floor. In short: it looks like gravity is present.

One can locally mimic gravity with an apparent force. The principle works locally: because of tidal forces — the force of gravity varies with distance — one can distinguish rocket-powered acceleration from a mass' gravitational field.

Second thought experiment. Consider a free falling box towards Earth. A mass inside the box experiences only gravity. In the frame of the box, the mass is floating and one may not distinguish the situation from the one where gravity is absent in the first place. In short: it looks like gravity is absent.

An example of a free falling experiment is the International Space Station (ISS): its altitude from the Earth's surface is $400\,\text{km}$, so the gravity is about 90% the one on the surface, but the feeling is that of weightlessness: the ISS is constantly falling, though it has enough lateral velocity that the Earth below moves away faster than the ISS can fall.

Einstein's equivalence principle. The equivalence principle is true for electromagnetism and all of physics. For a small enough box, one may not tell whether gravity is acting or not. As adding the constancy of the speed of light to Galilean relativity brings Special Relativity, then adding the equivalence principle to classical physics gives General Relativity.

Consequences. Since the principle applies to electromagnetism, then it also applies to light. Consider a laser shining a beam of light from left to right across a box. If the box is propelled upwards, the laser hits the right wall lower (relative to the floor) than it was shot from the left side. From the reference frame of the box, the laser is bending downwards. By the equivalence principle, the same should apply to a box immersed in a gravitational field. This prediction has been experimentally verified by gravitational lensing.

Curvature. Space has always been thought of as Euclidean space. However, if light curves, then the definition of straight line requires more caution. On a sphere, the sum of internal angles of a triangle is no longer π, but greater. After observing light curving, the geometry of space can no longer be Euclidean: since light travels in a curved trajectory, space itself is curved. On a sphere, the minimum distance between two points is given by the arc of the great circle passing between the two points. A straight line is then defined as the line that minimizes distance.

Since light is described differently by the equivalence principle, then massive objects need new equations also. These need to also explain the motion of planets: the precession of the perihelion of Mercury was not explained by Newtonian gravity.

A free particle in space-time maximizes its proper time which is proportional to the relativistic action. Objects that are only subject to gravity, either massive or massless, — that is, objects in free fall — follow geodesics, trajectories that maximize proper time. The trajectory of objects in free fall is described purely by geometrical ideas: there is no force of gravity.

Curvature is described in a way that matter bends space-time and space-time tells matter how to move.

Lecture 2

gio 05 ott 2023 14:30

2 Special relativity

Special relativity is a theory obtain from uniting Galilean principle of relativity and the idea that the speed of light is the same in every frame of reference. From the postulates one can derive time dilation, length contraction, relativity of simultaneity, etc. Lorentz transformations are used to go from one frame of reference to another. The Lorentz transformation Λ for a boost in the x direction is given by

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}, \quad \beta = \frac{v}{c}, \quad \gamma = (1-\beta^2)^{-\frac{1}{2}}$$

In this course, the natural unit $c = 1$ will be used. One can express Lorentz transformations also in terms of rapidity λ by setting

$$\gamma = \cosh\lambda\,, \quad \beta\gamma = \sinh\lambda$$

The Lorentz transformation above becomes

$$\Lambda = \begin{bmatrix} \cosh\lambda & -\sinh\lambda & 0 & 0 \\ -\sinh\lambda & \cosh\lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

which is a hyperbolic rotation: points in space-time are moved along hyperbolae. Ordinary three-dimensional rotation matrices are orthogonal matrices $R^\top R = RR^\top = I$ and leave the norm of $\mathbb{R}^3$ unchanged. Points are moved along circles. Similarly, Lorentz transformations preserve the Minkowski metric $\tau^2 = t^2 - |\vec{x}|^2$. This invariant is the proper time. For timelike vectors, it measures the time between two events happening in the same place. For spacelike vectors, it measures the spatial distance between two simultaneous events.

Points in Minkowski space are called events to highlight the fact that time is also a coordinate. The proper time of a trajectory is the time measured by an observer moving along that trajectory. Proper time is a physically meaningful quantity independent of frame of reference. The proper time between two timelike events is the shortest time one can hope to measure between those two events.

The reference frame of a moving object can be superimposed on one's own reference frame — using a Minkowski diagram — by tilting the object frame's axes by the same angle towards a bisector of the quadrants: the points on the axes are following hyperbolae. In four dimensional space-time, light rays define a light cone and Lorentz transformations define hyperboloids.

Length of a curve. Proper time is the length of a straight path in Minkowski space, but one can generalized the idea of length to more complicated paths. In Euclidean space, the length of a curve γ is given by the integral

$$\int_\gamma \mathrm{d}l = \int_\gamma \sqrt{\mathrm{d}x^i\,\mathrm{d}x^i} = \int_\gamma \mathrm{d}\lambda\,\sqrt{\partial_\lambda x^i\,\partial_\lambda x^i}$$

In Minkowski space, the time of a trajectory measured by an observer moving along such trajectory, the proper time, is

$$\tau(\gamma) = \int_\gamma \mathrm{d}\tau = \int_\gamma \sqrt{(\mathrm{d}t)^2 - \mathrm{d}x^i\,\mathrm{d}x^i} = \int_\gamma \mathrm{d}t\,\sqrt{1 - \partial_t x^i\,\partial_t x^i} = \int_\gamma \mathrm{d}t\,\sqrt{1-v^2} = \int_\gamma \frac{\mathrm{d}t}{\gamma} = \int_\gamma \mathrm{d}\tau$$

The trajectory and observer need not be inertial. In Euclidean space, a path between two points can be arbitrarily long, so the their distance is the shortest path. By analogy, because of the minus sign, the length of a path between two points in space-time, that is proper time, can be arbitrarily short — by going close to the speed of light —, so the distance between the two points is the longest path. Therefore, proper time is maximized by straight paths in space-time which correspond to inertial observers.

Proper time is proportional to the action of a relativistic free particle

$$S = -m \int \mathrm{d}\tau = -m \int \mathrm{d}t\,\sqrt{1-v^2} \sim -m \int \mathrm{d}t\,\left[1 - \frac{1}{2}v^2 + o(v^2)\right] \approx \int \mathrm{d}t\,\left[-m + \frac{1}{2}mv^2\right]$$

Taking the non-relativistic limit, one obtains the kinetic energy of a free particle. By varying the action, one can obtain the equations of motion $\ddot{x}^j = 0$ which is a straight line.

Four-momentum. Energy and momentum can be combined into a four-vector: $P^\mu = (E, \vec{p})$. The invariant associated with the norm of the vector is mass $m^2 = P_\mu P^\mu$. For an object at rest, $\vec{p} = 0$, one obtains $E = mc^2$. To get the explicit expression for the four-momentum in another reference frame, one can perform a boost from the rest frame:

$$\begin{bmatrix} E' \\ p'_x \\ p'_y \\ p'_z \end{bmatrix} = \begin{bmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma m \\ -\gamma m v \\ 0 \\ 0 \end{bmatrix}$$

By boosting in a direction, the object goes the other way. More generally, mixing the boost with a rotation, it follows $E = \gamma m$ and $\vec{p} = \gamma m \vec{v}$. In the limit of low speeds, one recovers the classical relations

$$E = \frac{m}{\sqrt{1 - v^2}} = m \left[1 + \frac{1}{2} v^2 + o(v^2) \right] \approx m + \frac{1}{2} m v^2$$

The rest energy in this expression has a different sign from the one found in the action. The rest energy is a costane potential, so going from the energy to the Lagrangian one has to add a minus sign to the potential.

2.1 Covariant notation

Covariant notation lets one treat time in the same manner as space. The norm of a four-vector is a quadratic form. For the proper time, one has

$$\tau^2 = - \begin{bmatrix} t & x & y & z \end{bmatrix} \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = -x^\top \eta x$$

The metric follows a spacelike notation typical of General Relativity, opposite to the one in QFT. In matrix notation, Lorentz transformations are given by

$$x' = \Lambda x$$

Proper time is an invariant and, calculated in two reference frames, one gets

$$\tau^2 = -x'^\top \eta x' = -(\Lambda x)^\top \eta (\Lambda x) = -x^\top \Lambda^\top \eta \Lambda x \equiv -x^\top \eta x \implies \eta = \Lambda^\top \eta \Lambda$$

Since the equality must hold for all vectors, then the implication must be follow. This means that the quadratic form is invariant under Lorentz transformations — a change of basis — and so is the inner product associated with the quadratic form $x_1^\top \eta x_2$. In Euclidean space, the equation above is $R^\top R = I$, so changes of basis are given by orthogonal matrices.

Lorentz transformations are elements of the Lorentz group $\mathrm{O}(3,1)$. The most general transformation depends on six parameters: three for boosts and three for rotations.

The standard inner product between two four-vectors $x_1^\top x_2$ is not invariant. Though, by defining a four-vector that transforms as

$$y' = \eta \Lambda \eta y$$

one can obtain an invariant inner product

$$x'^\top y' = x^\top \Lambda^\top \eta \Lambda \eta y = x^\top \eta^2 y = x^\top y$$

Vectors transforming as the position vector x are called contravariant and are represented as column vectors, while the vectors defined above are their dual, are called covariant and are represented as row vectors. From a contravariant vector x, one can derive a covariant one by applying the metric ηx and vice versa. From this, the inner product between contravariant and covariant vectors is invariant. In fact, the inner product $x_1^\top \eta x_2^\top$ can be viewed as the inner product between a contravariant vector x_1 and a covariant vector x_2.

The nature of a vector under Lorentz transformations is an important property and is made apparent with notation. Upper indices denote contravariant components x^μ, while lower indices denote covariant components y_μ. Greek indices include all four coordinates, while Latin indices only spatial coordinates. The inner product in Minkowski space becomes $x^\mu y_\mu$ using Einstein's summation convention: upper and lower indices appearing once are summed. Free indices should appear on both sides of an equation in the same position. If these are not the cases, the equation written has not the same form in every reference frame.

Lecture 3

lun 09 ott 2023 14:30

One can derivate the tensorial structure of the metric tensor. Knowing that the position vector is contravariant, then following the previous rules, one has

$$\tau^2 = -x^\top \eta x = -x^\mu \eta_{\mu\nu} x^\nu$$

In the matrix representation of a two-index tensor, the left index gives the μ-th row and the right index gives the ν-th column. The formula is also in agreement with the formula of a covariant vector

$$y = \eta x \implies y_\mu = \eta_{\mu\nu} x^\nu$$

The indices of the identity matrices are

$$x = Ix \implies x^\mu = \delta^\mu{}_\nu x^\nu$$

where $\delta^\mu{}_\nu$ is the Krockener delta. The inverse of the metric tensor is defined as

$$\eta^{-1}\eta = I \implies (\eta^{-1})^{\mu\rho}\eta_{\rho\nu} = \delta^\mu{}_\nu$$

It also happens that $\eta^{-1} = \eta$, but this is only true for Special Relativity. The Lorentz transformation has the following indices

$$x' = \Lambda x \implies x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

A Lorentz transformation is not symmetric in general. The defining equation of the Lorentz transformation is

$$\eta = \Lambda^\top \eta \Lambda \implies \eta_{\mu\nu} = (\Lambda^\top)_\mu{}^\rho \eta_{\rho\sigma} \Lambda^\sigma{}_\nu = \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu$$

3 Non-linear coordinate changes

One wants to generalize coordinate transformations beyond Lorentz's to be able to hop to non-inertial frames of reference.

The infinitesimal arc length, the line element, is given by

$$\mathrm{d}s^2 = \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu$$

It has the opposite sign of proper time and as such it is used to measure spatial lengths. A non-linear coordinate is given by a general function of the coordinates of the starting frame

$$x'^\mu = x'^\mu(x^0, x^1, x^2, x^3) = x'^\mu(x)$$

The infinitesimal transformation is given by the Jacobian

$$\mathrm{d}x'^\mu = \partial_\nu x'^\mu\,\mathrm{d}x^\nu = J^\mu{}_\nu\,\mathrm{d}x^\nu \iff \mathrm{d}x^\mu = \partial_{\nu'} x^\mu\,\mathrm{d}x'^\nu = (J^{-1})^\mu{}_\nu\,\mathrm{d}x'^\nu$$

Inserting the above in the line element gives

$$\mathrm{d}s^2 = \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = \eta_{\mu\nu}\,\partial_{\rho'}x^\mu\,\mathrm{d}x'^\rho\,\partial_{\sigma'}x^\nu\,\mathrm{d}x'^\sigma = (\eta_{\mu\nu}\,\partial_{\rho'}x^\mu\,\partial_{\sigma'}x^\nu)\,\mathrm{d}x'^\rho\,\mathrm{d}x'^\sigma$$

The factors in the expression above can be freely interchanged because sums are performed, it is only when using index-free notation that one needs to be careful about position. One can define the metric

$$g'_{\mu\nu} \equiv \eta_{\mu\nu}\,\partial_{\rho'}x^\mu\,\partial_{\sigma'}x^\nu$$

For Lorentz transformations, the Jacobian matrix is the Lorentz matrix ${\Lambda^\mu}_\nu$ and the metric reduces to the Minkowski metric $g'_{\mu\nu} = \eta_{\mu\nu}$. The metric g is not constant in space-time and across reference frames. Therefore, the line element depends on space-time points. A point-dependent quadratic form is called a metric. In general, one cannot assume the existence of a coordinate system in which the metric corresponds to the Minkowski metric. One has to deal with line elements

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu$$

In this course, the metric $g_{\mu\nu}$ is always non-degenerate because so is $\eta_{\mu\nu}$ and the Jacobians (for a non-degenerate coordinate change).

The metric has signature $(1, 3)$ and is called Lorentzian metric: it has only one negative eigenvalue in every point of space-time. Sometimes it is useful to only consider some coordinates and the metric does not have the negative eigenvalue: such metrics are called Riemannian.

With the understanding that line elements are the same in all reference frames, under a general change of coordinates, one gets

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = g_{\mu\nu}\,\partial_{\rho'}x^\mu\,\mathrm{d}x'^\rho\,\partial_{\sigma'}x^\nu\,\mathrm{d}x'^\sigma = (g_{\mu\nu}\,\partial_{\rho'}x^\mu\,\partial_{\sigma'}x^\nu)\,\mathrm{d}x'^\rho\,\mathrm{d}x'^\sigma = g'_{\rho\sigma}\,\mathrm{d}x'^\rho\,\mathrm{d}x'^\sigma$$

so that the metric transforms as

$$g'_{\mu\nu} = \partial_{\mu'}x^\rho\,\partial_{\nu'}x^\sigma\,g_{\rho\sigma} = {(J^{-1})^\rho}_\mu {(J^{-1})^\sigma}_\nu g_{\rho\sigma} = {[(J^{-1})^\top]_\mu}^\rho g_{\rho\sigma} {(J^{-1})^\sigma}_\nu = [(J^{-1})^\top g J^{-1}]_{\mu\nu}$$

This property can be added to the definition of a metric, but in this section it follows from previous definitions. In matrix notation, it is easy to see a resemblance with the defining equation for Lorentz matrices.

Example of an accelerated observer. Given the position vector, one can define the four-velocity

$$v^\mu \equiv \mathrm{d}_\tau x^\mu$$

In Special Relativity, the invariant associated with its norm is the speed of light

$$v^\mu v_\mu = \eta_{\mu\nu}v^\mu v^\nu = \eta_{\mu\nu}\,\mathrm{d}_\tau x^\mu\,\mathrm{d}_\tau x^\nu = -(\mathrm{d}_\tau\tau)^2 = -1$$

In General Relativity, the norm is

$$v^\mu v_\mu = g_{\mu\nu}v^\mu v^\nu = -1$$

Since the speed of light is the upper limit on velocity, it is clear that velocity does not increase linearly with time. With a constant acceleration, one can expect velocity to asymptotically approach the speed of light. Consider two velocity four-vector at infinitesimally close successive moments — $v^\mu(\tau)$ and $v^\mu(\tau + \mathrm{d}\tau)$ — of an accelerating particle in the x direction. And consider the reference frame instantaneously at rest with such particle. At the moment τ, the four-velocity is

$$v^\mu = \begin{pmatrix} 1 & \vec{0} \end{pmatrix}^\top$$

The following velocity is slightly different. In the frame at time τ, it is

$$v^\mu(\tau + \mathrm{d}\tau) = \begin{pmatrix} 1 & a\,\mathrm{d}\tau & 0 & 0 \end{pmatrix}^\top$$

The acceleration is defined in this way. One way of getting the velocity at all times is to boost back at a frame not at rest with the particle where $v^\mu(\tau)$ and $v^\mu(\tau + \mathrm{d}\tau)$ are known and find a differential equations for the components. Though, there is another method.

A boost in the x direction has the form

$${\Lambda^\mu}_\nu = \begin{bmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cosh\lambda & -\sinh\lambda & 0 & 0 \\ -\sinh\lambda & \cosh\lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In Relativity, the velocity does not sum linearly, but the rapidity does:

$$\Lambda(\lambda_1)\Lambda(\lambda_2) = \Lambda(\lambda_1 + \lambda_2)$$

To get the formula for the velocity, it is enough to note that $v = \tanh\lambda$ and then use the formula for the addition of the hyperbolic tangent. Let S be the rest frame at time τ and S' the rest frame at time $\tau + \mathrm{d}\tau$. In each rest frame, the velocity is $v^\mu = \begin{pmatrix} 1 & \vec{0} \end{pmatrix}^\top$. The velocity of the particle at time $\tau + \mathrm{d}\tau$ in the frame S is calculated from the one in the frame S' by a Lorentz boost:

$$v^\mu(\tau + \mathrm{d}\tau) = \Lambda_x(a' \,\mathrm{d}\tau) v'^\mu(\tau + \mathrm{d}\tau) = \Lambda_x(-a\,\mathrm{d}\tau) v^\mu(\tau)$$

Using also the fact that $\cosh\lambda = 1 + o(\lambda^2)$ and $\sinh\lambda = \lambda + o(\lambda^3)$ for small enough λ. Iterating the boosts, one gets

$$v^\mu(\tau) = \Lambda_x(-a\,\mathrm{d}\tau) v^\mu(\tau - \mathrm{d}\tau) = \Lambda_x(-a\,\mathrm{d}\tau)\Lambda_x(-a\,\mathrm{d}\tau) v^\mu(\tau - 2\,\mathrm{d}\tau) = \cdots$$

The velocity at an arbitrary time is the composition of many infinitesimal boosts. Let $\tau = N\,\mathrm{d}\tau$ for $N \gg 1$. Then

$$v^\tau = v^{N\,\mathrm{d}\tau} = \Lambda^N(-a\,\mathrm{d}\tau) v^\mu(0) = \Lambda(-Na\,\mathrm{d}\tau) v^\mu(0) = \Lambda(-a\tau) v^\mu(0)$$

The power N becomes a coefficient because rapidities sum. For an object initially at rest, one gets

$$v^\mu(\tau) = \begin{bmatrix} \cosh a\tau & \sinh a\tau & 0 & 0 \\ \sinh a\tau & \cosh a\tau & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \cosh(a\tau) \\ \sinh(a\tau) \\ 0 \\ 0 \end{bmatrix}$$

Integrating both sides in proper time, one finds

$$x^\mu = \frac{1}{a} \begin{pmatrix} \sinh(a\tau) & \cosh(a\tau) & 0 & 0 \end{pmatrix}^\top$$

Its trajectory, its world line is a spacelike hyperbola asymptotic to a light ray (with integrations constants chosen so that the ray goes through the origin) in the limit $\tau \to \infty$. It is approaching the speed of light.

One may want to find the coordinates of the accelerated observer. Its space coordinate ξ should be constant for trajectories of other accelerated objects, thus stationary in the observer's frame: the endpoints of an object at rest with the observer would produce lines of constant ξ. This is achieved by boosting the object by a parameter $\alpha = a\tau$. These lines are hyperbolae. Similarly, the time coordinate η should be constant for simultaneous events in the observer's frame. These lines are perpendicular to v^μ (with respect to the metric $\eta_{\mu\nu}$). Since the position x^μ is perpendicular to the velocity v^μ, then the lines at constant η pass through the origin [r].

These requirements are satisfied by

$$x^\mu = \frac{\mathrm{e}^{a\xi}}{a} \begin{pmatrix} \sinh\eta & \cosh\eta & 0 & 0 \end{pmatrix}^\top$$

Lecture 4

gio 12 ott 2023 14:30

Consider an observer on one hyperbola at $t = 0$ and another one at the origin. The second observer sent a message at light speed that is received by the first observer that can then respond. If the second observer sends another message, then it can never receive the first one: it is possibile to outrun a photon given enough head start. Any light message sent at $t \geq 0$ can never reach the accelerated observer.

The horizon is the locus of points beyond which an observer cannot communicate. An accelerated observer should experience phenomena similar to gravity.

[r]

$$\mathrm{d}s^2 = g'_{\mu\nu}\,\mathrm{d}x'^\mu\,\mathrm{d}x'^\nu = \partial_\mu x'^\rho\,\partial_\nu x'^\sigma g_{\rho\sigma}\,\mathrm{d}x^\rho\,\mathrm{d}x^\sigma = \mathrm{e}^{2a\xi}(-\mathrm{d}\eta^2 + \mathrm{d}\xi^2)$$

[r] and image

$$(J^{-1})^\top g J^{-1} = \mathrm{e}^{2a\xi} \begin{pmatrix} \cosh a\tau & \sinh a\tau \\ \sinh a\tau & \cosh a\tau \end{pmatrix} \eta \begin{pmatrix} \cosh a\tau & \sinh a\tau \\ \sinh a\tau & \cosh a\tau \end{pmatrix} = \mathrm{e}^{2a\xi}\eta$$

It is possibile to arrive at the answer in another way. Consider

$$\begin{aligned}\mathrm{d}x^0 &= \partial_{\mu'} x^0 \,\mathrm{d}x'^\mu = \cosh(a\eta)\mathrm{e}^{a\xi}\,\mathrm{d}\eta + \sinh(a\eta)\mathrm{e}^{a\xi}\,\mathrm{d}\xi \\ \mathrm{d}x^1 &= \sinh(a\eta)\mathrm{e}^{a\xi}\,\mathrm{d}\eta + \cosh(a\eta)\mathrm{e}^{a\xi}\,\mathrm{d}\xi\end{aligned}$$

then

$$\begin{aligned}\mathrm{d}s^2 &= \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = -(\mathrm{d}x^0)^2 + (\mathrm{d}x^i\,\mathrm{d}x^i) \\ &= \mathrm{e}^{2a\xi}\left[-\mathrm{d}\eta^2 + \mathrm{d}\xi^2\right] + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2 = g'_{\mu\nu}\,\mathrm{d}x'^\mu\,\mathrm{d}x'^\nu\end{aligned}$$

The metric is then

$$g'_{\mu\nu} = \begin{pmatrix} -\mathrm{e}^{2a\xi} & & & \\ & \mathrm{e}^{2a\xi} & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

This is also called Rindler metric. Rindler's space is the portion of space-time between the horizons [r].

One can study if given a metric one can always find a coordinate transformation such that the metric is Minkowski's [r].

3.1 Vector fields

The Minkowski metric was generalized to a point-dependent quadratic form. It no longer makes sense to talk about vectors, which have to be generalized to include point-dependence: vector fields. At every point in space-time there's a vector associated with it. One wants to study how a vector field transforms by looking at some invariant.

A partial derivative ∂_μ keeps constant all coordinates, but one in which a limit is taken

$$\partial_1 f = \lim_{\varepsilon\to 0} \frac{f(x^0, x^1+\varepsilon, x^2, x^3) - f(x^0, x^1, x^2, x^3)}{\varepsilon}$$

To take a derivative in an arbitrary direction one take the directional derivative, a linear combination of the partial derivatives along the axes

$$v^\mu\,\partial_\mu f = \lim_{\varepsilon\to 0} \frac{f(x^\mu + \varepsilon v^\mu) - f(x^\mu)}{\varepsilon}$$

The result of the limit does not depend on the coordinate system, but depends on a point and the direction of the vector:

$$v'^\mu\,\partial'_\mu f = v^\nu\,\partial_\nu f$$

The new derivative can be written using the chain rule

$$\partial'_\mu = \partial'_\mu x^\nu\,\partial_\nu \implies v'^\mu\,\partial'_\mu x^\nu\,\partial_\nu f = v^\nu\,\partial_\nu f \implies v^\nu = v'^\nu\,\partial'_\mu x^\nu \implies v'^\mu = \partial_\nu x'^\mu\,v^\nu = J^\mu{}_\nu\,x^\nu$$

In General Relativity, a vector field is a quadruple that transforms in the way above. Sometimes one writes

$$v \equiv v^\mu\,\partial_\mu$$

In quantum mechanics, $v^\mu\,\partial_\mu$ is also the generator of translations. While, the generator of rotations is the angular momentum. The generators can be promoted to finite transformations through an exponential map. An integral course is a line tangent to the vector field at any point.

It is natural to look at the commutator of a vector field

$$[v, w] = [v^\mu\,\partial_\mu, w^\nu\,\partial_\nu]$$

Acting on a test function one gets

$$\begin{aligned}[v^\mu\,\partial_\mu, w^\nu\,\partial_\nu]f &= v^\mu\,\partial_\mu(w^\nu\,\partial_\nu f) - w^\nu\,\partial_\nu(v^\mu\,\partial_\mu f) = v^\mu\,\partial_\mu w^\nu\,\partial_\nu f - w^\nu\,\partial_\nu\,\partial_\mu f \\ &= (v^\nu\,\partial_\nu w^\mu - w^\nu\,\partial_\nu v^\mu)\,\partial_\mu f\end{aligned}$$

from which
$$[v^\mu \, \partial_\mu, w^\nu \, \partial_\nu]^\rho = (v^\nu \, \partial_\nu w^\rho - w^\nu \, \partial_\nu v^\rho)$$
This is a Lie bracket and gives a new vector field. One should check that this expression transforms as a vector field.

In Special Relativity, an invariant is
$$\eta_{\mu\nu} v^\mu v^\nu = \|v\|^2 \equiv v^2$$
In General Relativity, one defines the norm
$$g_{\mu\nu} v^\mu v^\nu \equiv \|v\|^2$$
that is no longer invariant, but is a function: it depends on space-time points and it transforms without being multiplied by anything. [r] In fact
$$g'_{\mu\nu} v'^\mu v'^\nu = \partial'_\mu x^\rho \, \partial'_\nu x^\sigma g_{\rho\sigma} \, \partial_\lambda x'^\mu \, v^\lambda \, \partial_\alpha x'^\nu \, v^\alpha = \delta^\rho{}_\lambda \delta^\sigma{}_\alpha g_{\rho\sigma} v^\lambda v^\alpha = g_{\rho\sigma} v^\rho v^\sigma = v^2$$

3.2 Tensors

One can introduce the dual of a vector field. A (one-)form is an object ω_μ that trasforms as
$$\omega'_\mu = \partial'_\mu x^\nu \, \omega_\nu$$
The product between a form and a vector field is a function
$$\omega'_\mu v'^\mu = \partial'_\mu x^\nu \, \omega_\nu \, \partial_\rho x'^\mu \, v^\rho = \delta^\nu{}_\rho \omega_\nu v^\rho = \omega_\rho v^\rho$$
Combining a form with an infinitesimal displacement one defines
$$\omega_\mu \, \mathrm{d}x^\mu \equiv \omega$$

Lecture 5

lun 16 ott 2023 14:30

Example. An example of form is given by the derivative of a function $\partial_\mu f$. Given a vector and a form, there's a natural pairing between the two. The directional derivative is a function $v^\mu \, \partial_\mu f$.

Example. A vector field is the generalization of contravariant vectors from Special Relativity, while a form is the generalization of covariant vectors. Similarly to Special Relativity, one can construct a form from a vector field
$$\omega_\mu = g_{\mu\nu} v^\nu$$
In fact, its transformation law is
$$\omega'_\mu = g'_{\mu\nu} v'^\nu = \partial'_\mu x^\rho \, \partial'_\nu x^\sigma \, g_{\rho\sigma} \, \partial_\lambda x'^\nu v^\lambda = \partial'_\mu x^\rho \, \delta^\sigma{}_\lambda g_{\rho\sigma} v^\lambda = \partial'_\mu x^\rho g_{\rho\sigma} v^\sigma = \partial'_\mu x^\rho \, \omega_\rho$$
It is typical to denote the form constructed from a vector with the same symbol: $v_\mu = g_{\mu\nu} v^\nu$.

Tensors. In general, one can define a tensor field $T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}$ of type (k, l) as an object that transforms as
$$(T')^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} = \partial_{\rho_1} x'^{\mu_1} \cdots \partial_{\rho_k} x'^{\mu_k} T^{\rho_1 \cdots \rho_k}{}_{\sigma_1 \cdots \sigma_l} \, \partial'_{\nu_1} x^{\sigma_1} \cdots \partial'_{\nu_l} x^{\sigma_l}$$
It has k contravariant components and l covariant components. The word field is often omitted, but it is understood in almost any case.

A function is type $(0, 0)$ tensor. A metric is a type $(0, 2)$ tensor. A vector field is a type $(1, 0)$ tensor. A form is a type $(0, 1)$ tensor.

The Kronecker delta $\delta^\mu{}_\nu$ is a type $(1, 1)$ tensor with the same expression in all coordinate systems. In fact, its transformation is
$$(\delta')^\mu{}_\nu = \partial_\rho x'^\mu \, \partial'_\nu x^\sigma \delta^\rho{}_\sigma = \partial_\rho x'^\mu \, \partial'_\nu x^\rho = \partial'_\nu x'^\mu = \delta^\mu{}_\nu$$

Metric. Given a metric $g_{\mu\nu}$, its inverse $(g^{-1})^{\mu\nu}$ satisfies

$$(g^{-1})^{\mu\rho}g_{\rho\nu} = \delta^\mu{}_\nu$$

The inverse is typically denoted as $g^{\mu\nu}$ and it is clear that the metric is different from $g_{\mu\nu}$. The inverse metric is a type $(2,0)$ tensor.

Derivative. Given a vector field v^ν, its derivative $\partial_\mu v^\nu$ is not a tensor. In fact, its transformation is

$$\partial'_\mu v'^\nu = \partial'_\mu x^\rho\, \partial_\rho(\partial_\sigma x'^\nu\, v^\sigma) = \partial'_\mu x^\rho\, \partial_\sigma x'^\nu\, \partial_\rho v^\sigma + \partial'_\mu x^\rho\, \partial_\rho\partial_\sigma x'^\nu\, v^\sigma$$

For it to be a tensor, one expects only the first addendum to appear. There is no natural notion of derivative.

Anti-symmetrization. Given a form ω_ν, its derivative $\partial_\mu\omega_\nu$ is not a tensor. However, the anti-symmetrization is a type $(0,2)$ tensor

$$\partial_\mu\omega_\nu - \partial_\nu\omega_\mu \equiv 2\partial_{[\mu}\omega_{\nu]}$$

The transformation of the derivative is

$$\begin{aligned}\partial'_\mu\omega'_\nu &= \partial'_\mu x^\rho\, \partial_\rho(\partial'_\nu x^\sigma\, \omega_\sigma) = \partial'_\mu x^\rho\, \partial'_\nu x^\sigma\, \partial_\rho\omega_\sigma + \partial'_\mu x^\rho\, \partial_\rho\partial'_\nu x^\sigma\, \omega_\sigma \\ &= \partial'_\mu x^\rho\, \partial'_\nu x^\sigma\, \partial_\rho\omega_\sigma + \partial'^2_{\mu\nu} x^\sigma\, \omega_\sigma\end{aligned}$$

The transformation of the anti-symmetrization is

$$\begin{aligned}2\partial'_{[\mu}\omega'_{\nu]} &= \partial'_{[\mu} x^\rho\, \partial'_{\nu]} x^\sigma\, \partial_\rho\omega_\sigma + \partial'^2_{[\mu\nu]} x^\sigma\, \omega_\sigma = \partial'_{[\mu} x^\rho\, \partial'_{\nu]} x^\sigma\, \partial_\rho\omega_\sigma + \partial'^2_{\mu\nu} x^\sigma\, \omega_\sigma - \partial'^2_{\nu\mu} x^\sigma\, \omega_\sigma \\ &= \partial'_{[\mu} x^\rho\, \partial'_{\nu]} x^\sigma\, \partial_\rho\omega_\sigma = \partial'_\mu x^\rho\, \partial'_\nu x^\sigma\, \partial_\rho\omega_\sigma - \partial'_\nu x^\rho\, \partial'_\mu x^\sigma\, \partial_\rho\omega_\sigma \\ &= \partial'_\mu x^\rho\, \partial'_\nu x^\sigma\, \partial_\rho\omega_\sigma - \partial'_\nu x^\sigma\, \partial'_\mu x^\rho\, \partial_\sigma\omega_\rho = \partial'_\mu x^\rho\, \partial'_\nu x^\sigma(\partial_\rho\omega_\sigma - \partial_\sigma\omega_\rho) \\ &= 2\, \partial'_\mu x^\rho\, \partial'_\nu x^\sigma\, \partial_{[\rho}\omega_{\sigma]}\end{aligned}$$

At the second line, as long as the transformation between coordinates is a smooth function, then the mixed derivates are equal. At the third line, the indices σ and ρ are interchanged since they are summed over. An example of such a tensor is the field-strength tensor

$$F_{\mu\nu} = 2\, \partial_{[\mu}A_{\nu]}$$

where A_μ is the four-potential and also a form.

If $\omega_{\mu\nu}$ is a type $(0,2)$ tensor and it is anti-symmetric, then the following is a tensor

$$\begin{aligned}\partial_{[\mu}\omega_{\nu\rho]} &= \frac{1}{3!}(\partial_\mu\omega_{\nu\rho} - \partial_\mu\omega_{\rho\nu} + \partial_\nu\omega_{\rho\mu} - \partial_\nu\omega_{\mu\rho} + \partial_\rho\omega_{\mu\nu} - \partial_\rho\omega_{\nu\mu}) \\ &= \frac{1}{3}(\partial_\mu\omega_{\nu\rho} + \partial_\nu\omega_{\rho\mu} + \partial_\rho\omega_{\mu\nu})\end{aligned}$$

The homogeneous Maxwell's equations, written as Bianchi's identity, are an example of such a tensor

$$\partial_{[\mu}F_{\nu\rho]} = 0$$

This tensor is zero in every coordinate system. The aim is to formulate equations that have the same form in every reference frame. The inhomogeneous Maxwell's equations in the vacuum are

$$\partial_\mu F^{\mu\nu} = J^\nu = 0$$

but this is not a tensor in General Relativity: it needs to be modified.

More generally, given an anti-symmetric type $(0,k)$ tensor $A_{\mu_1\cdots\mu_k}$, its anti-symmetrized derivative $\partial_{[\mu}A_{\mu_1\cdots\mu_k]}$ is also a tensor. Anti-symmetric type $(0,k)$ tensors are called k-forms.

3.3 Lie derivatives

The directional derivative is given by $vf \equiv v^\mu \, \partial_\mu f$ and it acts on a function. One can define a similar derivative acting on a vector field. Given a vector field w^ν, the expression $v^\mu \, \partial_\mu w^\nu$ is not a vector field. The directional derivative is defined as

$$v^\mu \, \partial_\mu f = \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon v) - f(x)}{\varepsilon}$$

But for the derivative of a vector field it doesn't make sense to compare the vector field at the transformed point $w^\mu(x + \varepsilon v)$ with the vector field at the given point $w^\mu(x)$. Instead, one wants to transform the vector field at the transformed point, then compare it

$$(L_v w)^\mu = \lim_{\varepsilon \to 0} \frac{J^\mu{}_\nu \, w^\nu(x + \varepsilon v) - w^\mu(x)}{\varepsilon}$$

where J is the Jacobian of the transformation generated by $-\vec{v}$. Let $x'^\mu = x^\mu - \varepsilon v^\mu$, the Jacobian is

$$J^\mu{}_\nu = \partial_\nu x'^\mu = \delta^\mu{}_\nu - \varepsilon \, \partial_\nu v^\mu$$

By expanding in a Taylor series, one gets

$$w^\nu(x + \varepsilon v) \approx w^\nu(x) + \varepsilon v^\mu \, \partial_\mu w^\nu$$

Then one has

$$\begin{aligned}
(L_v w)^\mu &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [(\delta^\mu{}_\nu - \varepsilon \, \partial_\nu v^\mu)(w^\nu + \varepsilon v^\rho \, \partial_\rho w^\nu) - w^\mu](x) \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [w^\mu - \varepsilon \, (\partial_\nu v^\mu) w^\nu + \varepsilon v^\rho \, \partial_\rho w^\mu - w^\mu](x) \\
&= \lim_{\varepsilon \to 0} [-\varepsilon \, (\partial_\nu v^\mu) w^\nu + \varepsilon v^\rho \, \partial_\rho w^\mu](x) = v^\nu \, \partial_\nu w^\mu - w^\nu \, \partial_\nu v^\mu = [v, w]^\mu
\end{aligned}$$

which is a vector field. The negative term came from the Jacobian matrix and has a geometrical meaning: when one applies the map generated by v to bring the point $x + \varepsilon v$ back to x, the vector w^ν might rotate a little.

This idea and method can be applied to any tensor: the derivative of a tensor along a vector field is again a tensor. This type of derivative is called Lie derivative.

A second method to determine the form of the more general Lie derivative imposes the Leibniz identity. Consider a one-form ω_μ and a vector field z^μ. It is already known that $\omega_\mu z^\mu$ is a function. The Lie derivative is

$$\begin{aligned}
L_v(\omega_\mu z^\mu) &= (L_v \omega)_\mu z^\mu + \omega_\mu (L_v z)^\mu \\
v^\nu \, \partial_\nu(\omega_\mu z^\mu) &= (L_v \omega)_\mu z^\mu + \omega_\mu (v^\nu \, \partial_\nu z^\mu - z^\nu \, \partial_\nu v^\mu) \\
v^\nu (\partial_\nu \omega_\mu) z^\mu + v^\nu \omega_\mu \, \partial_\nu z^\mu &= (L_v \omega)_\mu z^\mu + \omega_\mu (v^\nu \, \partial_\nu z^\mu - z^\nu \, \partial_\nu v^\mu) \\
v^\nu (\partial_\nu \omega_\mu) z^\mu &= (L_v \omega)_\mu z^\mu - \omega_\mu z^\nu \, \partial_\nu v^\mu \\
(L_v \omega)_\mu z^\mu &= v^\nu (\partial_\nu \omega_\mu) z^\mu + \omega_\mu z^\nu \, \partial_\nu v^\mu \\
(L_v \omega)_\mu z^\mu &= v^\nu (\partial_\nu \omega_\mu) z^\mu + \omega_\nu z^\mu \, \partial_\mu v^\nu \\
(L_v \omega)_\mu z^\mu &= [v^\nu \partial_\nu \omega_\mu + \omega_\nu \, \partial_\mu v^\nu] z^\mu
\end{aligned}$$

Since z^μ is an arbitrary vector, one has the following one-form

$$(L_v \omega)_\mu = v^\nu \partial_\nu \omega_\mu + \omega_\nu \, \partial_\mu v^\nu$$

The second addendum of this formula has the opposite sign as the one in the Lie derivative of a vector: this is a consequence of the Jacobian. If one evaluates the limit definition, one finds the inverse Jacobian and as such it has $\delta^\mu{}_\nu + \varepsilon \, \partial_\nu v^\mu$. The Lie derivative of a scalar is just the directional derivative

$$L_v f = v^\nu \partial_\nu f$$

One may notice that in these three examples of Lie derivates, the directional derivative is always present.

Using either method, one can obtain the Lie derivative of any tensor. In general, for a tensor field one has

$$\begin{aligned}(L_v T)^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = {} & v^\rho \partial_\rho T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} \\ & - (\partial_\rho v^{\mu_1}) T^{\rho\mu_2\cdots\mu_k}{}_{\nu_1\cdots\nu_l} - (\partial_\rho v^{\mu_2}) T^{\mu_1\rho\cdots\mu_k}{}_{\nu_1\cdots\nu_l} - \cdots - (\partial_\rho v^{\mu_k}) T^{\mu_1\cdots\rho}{}_{\nu_1\cdots\nu_l} \\ & + (\partial_{\nu_1} v^\rho) T^{\mu_1\cdots\mu_k}{}_{\rho\nu_2\cdots\nu_l} + (\partial_{\nu_2} v^\rho) T^{\mu_1\cdots\mu_k}{}_{\nu_1\rho\cdots\nu_l} + \cdots + (\partial_{\nu_l} v^\rho) T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\rho}\end{aligned}$$

In general, the derivatives of tensors are not tensors. Lie derivatives let one differentiate a tensor along a vector field through a linear combination of partial derivatives. There is a way, the covariant derivative, to improve the notion of partial derivative without involving the linear combinations above.

The Lie derivative has also a geometrical meaning: it is the comparison between the tensor at a point and the tensor at another point that has been dragged back to the first point.

Metric tensor. The Lie derivative of the metric is

$$(L_v g)_{\mu\nu} = v^\rho \partial_\rho g_{\mu\nu} + (\partial_\mu v^\rho) g_{\rho\nu} + (\partial_\nu v^\rho) g_{\mu\rho}$$

A vector field v^μ such that the Lie derivative above is zero, is an infinitesimal transformation that leaves invariant the metric and is a symmetry of the metric.

Lecture 6

gio 19 ott 2023 14:30

Not every coordinate system can cover all space-time. An example of this is Rindler space.

3.4 Manifolds

Topological spaces are considered to be equivalent if there is a homeomorphism connecting them. A homeomorphism is a continuous bijective map whose inverse is also continuous. A more advanced concept of being equivalent is the homotopic concept.

A manifold of dimension N is a topological space such that every point has a neighbourhood that is homeomorphic to an open set of $\mathbb{R}^n$. [r]

A smooth manifold M is a manifold with a choice of open sets U_i (called charts) and maps $\phi \to \mathbb{R}^n$ such that

$$\bigcup_i U_i = M$$

and in regions of overlap

$$U_i \cap U_j \neq \emptyset \implies \phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

the transition functions are smooth. This means that two coordinate systems transition well when going when one to another.

The collection of all charts and maps is called atlas.

A smooth map f from a smooth manifold M of dimension N to another smooth manifold M' of dimension N' is a map such that at any point $p \in M$, taking one set U in which p belongs to and the associated map ϕ, and taking one V which $f(p)$ belongs to and the associated map ψ, then $\psi \circ f \circ \phi^{-1}$ is smooth.

A diffeomorphism from a smooth manifold M to another smooth manifold M' is a smooth bijective map whose inverse is also smooth.

Two manifolds can be considered the same once one can find a diffeomorphism between them. An interesting case is considering maps from a space to itself $M = M'$.

Exercise. Consider the two-sphere

$$S^2 = \{x_1^2 + x_2^2 + x_3^2 = 1\} \subset \mathbb{R}^3$$

It can be covered by two charts:

$$U_N = S^2 \setminus \{(0,0,-1)\}\,, \quad U_S = S^2 \setminus \{(0,0,1)\}$$

Every point p of the sphere can be mapped to a point $\phi(p)$ in the tangent plane at one pole using a ray starting from the other pole

$$\phi_S : U_S \to \mathbb{R}^2 \,, \quad \phi_N : U_N \to \mathbb{R}^2$$

It is similar to the Riemann sphere. Find the transition from one set to the other.

Another way of working on the sphere employs spherical coordinates (θ, φ) but it only works on $S^2 \setminus \{(0,0,\pm 1)\}$.

Metric. Since multiple coordinates systems are needed, one needs to study the metric. A metric on a manifold M is a choice of metrics g_i on each chart U_i such that on each intersection $U_i \cap U_j$ the line elements coincide $\mathrm{d}s_i^2 = \mathrm{d}s_i^2$:

$$(g_i)_{\mu\nu} = \frac{\partial x_j^\rho}{\partial x_i^\mu} \frac{\partial x_j^\sigma}{\partial x_i^\nu} (g_j)_{\rho\sigma}$$

Example. Consider again the two-sphere S^2. There is a natural metric induced by the euclidean metric in $\mathbb{R}^3$. The euclidean metric is

$$\mathrm{d}s^2 = \mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2$$

Given the transformation from polar coordinates to Cartesian, one gets

$$x_1 = r \sin\theta \cos\varphi \,, \quad x_2 = r \sin\theta \sin\varphi \,, \quad x_3 = r\cos\theta$$

Calculating each infinitesimal, the line element becomes

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\varphi^2)$$

When interested only in the surface of the sphere, the radius stays constant:

$$\mathrm{d}s^2 = \mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\varphi^2 \implies g_{\mu\nu} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}$$

There are three symmetries that leave the sphere invariant: rotations. This is the round metric. It is two-dimensional without time. In this case, one is interested in just a part of space-time, not all of it. This is a Riemannian metric because it is positive definite. This metric is degenerate for $\theta = 0, \pi$, at the poles. This degeneracy is an artifact of the coordinate system.

One can change to a coordinate system that does not present these artifacts. For the two-sphere, one can utilize the stereographic projection and the coordinates of the projective plane

$$x = \tan\frac{\theta}{2}\cos\varphi \,, \quad y = \tan\frac{\theta}{2}\sin\varphi \implies \mathrm{d}s^2 = 4\frac{\mathrm{d}x^2 + \mathrm{d}y^2}{(1 + x^2 + y^2)^2}$$

This is the Fubini–Study metric. This metric applies to projective spaces. It is well-defined except at the south pole. The metric is

$$g_{\mu\nu} = \frac{4}{(1 + x^2 + y^2)^2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

By changing coordinate systems, the degeneracies vanish. This phenomenon also happens with more complicated metrics.

The formalism built patches multiple $\mathbb{R}^n$ together to describe a general space-time. One can also use arbitrary tensor fields.

Lecture 7

lun 23 ott 2023 14:30

4 Curvature

Curvature in the plane. Consider a curve in the plane. Intuitively, curvature tries to quantify how much to steer to keep going along the curve. At every point along the curve, one considers a circle (called osculating circle) that coincides with such curve up to second order. The curvature at a point p is

$$\kappa_p = \frac{1}{R_p}$$

where R_p is the radius of the osculating circle.

Curvature in space. Considering osculating circles implies that curvature depends on direction and varies between a minimum and a maximum. One can also consider ellipsoids that are characterized by the minimum and maximum curvature. On a cylinder, the straight path going from one base to the other has no curvature, while a path going across has curvature R^{-1}. On a sphere, the curvature is the same everywhere R^{-1}.

The problem with this intuitive definition is that curvature is not intrinsic, but depends on the way it is embedded in $\mathbb{R}^3$ (or whichever space one is working in). A cylinder can be constructed from a flat piece of paper, so the curvature depends on the embedding of the surface in $\mathbb{R}^3$. One seeks a definition of curvature that depends only on the metric of the surface, not on the embedding and the higher-dimensional space.

One can find an intrinsic way to define curvature through Gauss's Theorema Egregium: the product $\max\kappa\,\min\kappa$ only depends on the surface, not the embedding in $\mathbb{R}^3$. In such a way, a cylinder has no curvature and similarly a cone. Conversely, a sphere has intrinsic curvature.

Topology studies shapes that are stretchable while in differential geometry there is no possibility of stretching.

In higher dimensions, the generalization uses the sum and the product of curvatures: the trace and the determinant of a matrix.

4.1 Covariant derivative

This course follows a modern definition of curvature that generalizes well to higher dimensions. The covariant derivate is based on the idea that a curved surface does not keep a translating vector pointing in the same direction. For example, consider a tangent vector on a sphere at its equator. One can translate the vector up by an angle $\frac{\pi}{2}$ to the north pole keeping it always tangent, then down right by an angle $\frac{\pi}{2}$ by keeping it pointing in the same direction, lastly one can get back to the original point keeping the vector always tangent: the vector rotates by a $\frac{\pi}{2}$ angle. If one makes the path infinitesimal, the rotation goes to zero quadratically. The coefficient of the rotation matrix is the curvature which can be defined at every point by taking the infinitesimal limit of some closed paths.

To keep a vector in the same direction, some derivative must be zero. One needs to compute the derivative of a vector. If the derivative along the path is zero, the direction does not change: though the derivative of a vector is not a tensor, so one does not know how to take the derivative. The Lie derivative $(L_v w)^\nu$ cannot be the concept one is looking for. If v is the vector giving the direction and w is the vector one wants to traslate, then

$$(L_v w)^\nu = v^\mu\,\partial_\mu w^\nu - w^\mu\,\partial_\mu v^\nu$$

so the derivative depends on the path and the components of v, but the vector v giving the direction is not a vector field and its derivatives depends on its values in a neighbourhood of the path. Instead, one wants a dependance only on the direction itself. Also, it cannot be right because the metric does not appear and the metric is needed because one wants to keep angles the same.

One must add another term to the derivative. This term can be justified with an analogy. In electromagnetism, the fields are gauge invariant

$$A_\mu \to A_\mu - \partial_\mu\lambda$$

The wave function of a charged particles transforms as $\psi \to \mathrm{e}^{\mathrm{i}\lambda}\psi$, but the derivative operator (such as the one in the hamiltonian) transforms as

$$\partial_i \psi \to \partial_i(\mathrm{e}^{\mathrm{i}\lambda}\psi) = \mathrm{e}^{\mathrm{i}\lambda}(\partial_i\psi + \mathrm{i}\,\partial_i\lambda\,\psi)$$

The second term creates issues when calculating expectation values because the gauge transformation may not cancel out. Everything physical should not depend on the gauge. This is similar to the directional derivative of a vector where there is a second term that creates problems. The solution in quantum mechanics considers

$$D_i\psi = (\partial_i - \mathrm{i}A_i)\psi \to \mathrm{e}^{\mathrm{i}\lambda}(\partial_i\psi + \mathrm{i}\,\partial_i\lambda\,\psi) - \mathrm{i}(A_i + \partial_i\lambda)\mathrm{e}^{\mathrm{i}\lambda}\psi = D_i\psi$$

And the extra term is cancelled. The derivative D_i is the covariant derivative.

In General Relativity, one cannot write $\partial_\mu v^\nu + A_\mu v^\nu$, but considers the covariant derivative

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\mu\rho} v^\rho$$

where one should understand the Christoffel symbol as $\Gamma^\nu_{\mu\rho} \equiv (\Gamma_\mu)^\nu{}_\rho$ in which μ is the index along which the derivation takes place and $\nu\rho$ are matrix indices. One looks for the transformation law of $\Gamma^\nu_{\mu\rho}$ that makes the covariant derivative a tensor. Therefore

$$v^\mu \to v'^\mu = \partial_\nu x'^\mu\, v^\nu\,, \quad \partial_\mu \to \partial'_\mu = \partial'_\mu x^\nu\,\partial_\nu$$

then

$$\begin{aligned}
\nabla'_\mu v'^\nu &= \partial'_\mu v'^\nu + \Gamma'^\rho_{\mu\nu} v'^\rho = \partial'_\mu x^\rho\,\partial_\rho(\partial_\sigma x'^\nu\, v^\sigma) + \Gamma'^\rho_{\mu\nu}\,\partial_\sigma x'^\rho\, v^\sigma \\
&= \partial'_\mu x^\rho(\partial_\sigma x'^\nu\,\partial_\rho v^\sigma + \partial_\rho\partial_\sigma x'^\nu\, v^\sigma) + \Gamma'^\rho_{\mu\nu}\,\partial_\sigma x'^\rho\, v^\sigma \\
&\equiv \partial'_\mu x^\rho\,\partial_\sigma x'^\nu\,\nabla_\rho v^\sigma = \partial'_\mu x^\rho\,\partial_\sigma x'^\nu\,(\partial_\rho v^\sigma + \Gamma^\sigma_{\rho\lambda} v^\lambda)
\end{aligned}$$

[r] It follows

$$\Gamma'^\alpha_{\mu\nu} = \partial'_\mu x^\rho\,\partial_\nu x'^\lambda\,(\partial_\beta x'^\alpha \Gamma^\beta_{\rho\lambda} - \partial_{\rho\lambda} x'^\alpha)$$

[r] This makes $\nabla_\mu v^\nu$ a tensor. The Christoffel symbol is not a tensor, but is called a connection. The covariant derivative gives a sense to a vector in two different points pointing in the same direction. In this way, one can define what it means for a vector to be transported to another point without changing. A vector $v^\mu(x^\mu + \varepsilon z^\mu)$ has been parallel transported from $v^\mu(x^\mu) = v^\mu_0$ if

$$0 = z^\nu \nabla_\nu v^\mu = z^\nu(\partial_\nu x^\mu + \Gamma^\mu_{\nu\rho} v^\rho)(x^\mu) \sim \frac{1}{\varepsilon}[v^\mu(x+\varepsilon z) - v^\mu(x)] + (\Gamma^\mu_{\nu\rho} v^\rho z^\nu)(x)$$

Therefore

$$v^\mu(x+\varepsilon z) \sim (v^\mu - \varepsilon\Gamma^\mu_{\nu\rho} v^\rho z^\nu)(x) = v^\mu_0 - \varepsilon\Gamma^\mu_{\nu\rho}(x) v^\rho z^\nu(x)$$

So it is the initial vector rotated by a bit.

One can take the derivative of a function

$$\nabla_\mu f = \partial_\mu f$$

which is a form. Also one can utilize Leibniz rule

$$\begin{aligned}
\nabla_\mu(v^\nu\omega_\nu) &= (\nabla_\mu v^\nu)\omega_\nu + v^\nu(\nabla_\mu\omega_\nu) \\
\partial_\mu(v^\nu\omega_\nu) = (\partial_\mu v^\nu)\omega_\nu + v^\nu(\partial_\mu\omega_\nu)a &= (\partial_\mu v^\nu \Gamma^\nu_{\mu\rho} v^\rho)\omega_\nu + v^\nu(\nabla_\mu\omega_v) \\
v^\nu(\partial_\mu\omega_\nu) - \Gamma^\nu_{\mu\rho} v^\rho\omega_\nu &= v^\nu(\nabla_\mu\omega_\nu) \\
v^\nu(\partial_\mu\omega_\nu) - \Gamma^\nu_{\mu\nu} v^\nu\omega_\nu &= v^\nu(\nabla_\mu\omega_\nu)
\end{aligned}$$

This implies

$$\nabla_\mu\omega_\nu = \partial_\mu\omega_\nu - \Gamma^\rho_{\mu\nu}\omega_\rho$$

Lecture 8

gio 26 ott 2023 14:30

Comparing this expression with the one for a vector, one notices: the index μ is the direction along which one differentiates, there is a sign difference and the ν index changes position so the final free index ρ can be rightly contracted. This is similar to the Lie derivative.

From this observation, one can generalize the covariant derivative to act on any tensor:

$$\begin{aligned}\nabla_\mu T^{\nu_1\cdots\nu_k}{}_{\rho_1\cdots\rho_l} &= \partial_\mu T^{\nu_1\cdots\nu_k}{}_{\rho_1\cdots\rho_l} \\ &\quad + \Gamma^{\nu_1}_{\mu\sigma} T^{\sigma\nu_2\cdots\nu_k}{}_{\rho_1\cdots\rho_l} + \Gamma^{\nu_2}_{\mu\sigma} T^{\nu_1\sigma\cdots\nu_k}{}_{\rho_1\cdots\rho_l} + \cdots + \Gamma^{\nu_k}_{\mu\sigma} T^{\nu_1\cdots\sigma}{}_{\rho_1\cdots\rho_l} \\ &\quad - \Gamma^{\sigma}_{\mu\rho_1} T^{\nu_1\cdots\nu_k}{}_{\sigma\rho_2\cdots\rho_l} - \Gamma^{\sigma}_{\mu\rho_2} T^{\nu_1\cdots\nu_k}{}_{\rho_1\sigma\cdots\rho_l} - \cdots - \Gamma^{\sigma}_{\mu\rho_l} T^{\nu_1\cdots\nu_k}{}_{\rho_1\cdots\sigma}\end{aligned}$$

Once one knows how an operator acts on a co(ntra)variant index, the generalization to any type of tensor is straightforward.

4.2 Levi–Civita connection

At this point one should study the properties of the Christoffel symbol Γ. This symbol defines the condition for two vectors at different points to be the same. From the definition of the Christoffel symbol, one can find a geometrical intuition. A transported vector should have the same magnitude and the same angle along the trajectory. For this notion of parallel transport, one would like to preserve lengths and angles. For this purpose one should impose a condition to connect Γ to the metric since lengths and angles both are given by the metric. Preserving those two is equivalent to requiring that the inner product defined by the metric is preserved. The length squared and the inner product are defined as

$$v^2 \equiv g_{\mu\nu} v^\mu v^\nu \,, \quad v \cdot w \equiv g_{\mu\nu} v^\mu w^\nu$$

Therefore, the metric must be preserved and the covariant derivative must be zero along the trajectory:

$$0 = \nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\sigma_{\rho\mu} g_{\sigma\nu} - \Gamma^\sigma_{\rho\nu} g_{\mu\sigma}$$

In fact, from

$$\nabla_\rho v^2 = (\nabla_\rho g_{\mu\nu}) v^\mu v^\nu + g_{\mu\nu} (\nabla_\rho v^\mu) v^\nu + g_{\mu\nu} v^\mu \nabla_\rho v^\nu$$

and applying the above, it follows that $z^\rho \nabla_\rho v^2 = 0$ where z^ρ is a direction. A similar argument applies to angles.

The definition of the Christoffel symbol Γ from the covariant derivative of the metric is not unique. One can impose a second property just for convenience, a property of minimality. Recalling that the anti-symmetrized partial derivative of a form $\partial_{[\mu}\omega_{\nu]}$ is a tensor, one imposes

$$\partial_{[\mu}\omega_{\nu]} \equiv \nabla_{[\mu}\omega_{\nu]} = \partial_{[\mu}\omega_{\nu]} - \Gamma^\rho_{[\mu\nu]}\omega_\rho \implies \Gamma^\rho_{[\mu\nu]}\omega_\rho = 0 \,, \quad \forall \omega \implies \Gamma^\rho_{[\mu\nu]} = 0$$

In such a way there is only one definition of anti-symmetrized derivative. This means that the Christoffel symbol is symmetric in the covariant indices

$$\Gamma^\rho_{\mu\nu} = \Gamma^\rho_{\nu\mu}$$

The anti-symmetrized Christoffel symbol is a tensor and is called torsion

$$T^\rho_{\mu\nu} = \Gamma^\rho_{[\mu\nu]} = \frac{1}{2}(\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu})$$

There exists only one Christoffel symbol that is torsionless and preserves the metric. The proof is constructive. From the preservation of the metric:

$$\begin{aligned}0 = \nabla_\rho g_{\mu\nu} &= \partial_\rho g_{\mu\nu} - \Gamma^\sigma_{\rho\mu} g_{\sigma\nu} - \Gamma^\sigma_{\rho\nu} g_{\mu\sigma} = \partial_\rho g_{\mu\nu} - \Gamma_{\nu\rho\mu} - \Gamma_{\mu\rho\nu} \\ &= \partial_\rho g_{\mu\nu} - 2\Gamma_{(\nu|\rho|\mu)} = \partial_\rho g_{\mu\nu} - 2\Gamma_{(\nu\mu)\rho}\end{aligned}$$

Even though Γ is not a tensor, one can define the symbol with mixed indices by using the metric, similar to how one can obtain a one-form from a vector and viceversa. The position (left-right) of the indices is set by convention. It would be less confusing to have $(\Gamma_\rho)^\sigma{}_\mu g_{\sigma\nu} = (\Gamma_\rho)_{\nu\mu}$. In the second row, first equality, the symmetrization[1] is done over $\nu\mu$. The second equality follows from the first row, last equality by the torsionless property of Γ.

Minding the change in indices, it follows

$$\partial_\mu g_{\nu\rho} = \Gamma_{\nu\mu\rho} + \Gamma_{\rho\mu\nu} \implies \partial_{(\mu} g_{\nu)\rho} = \Gamma_{(\nu\mu)\rho} + \Gamma_{\rho(\mu\nu)} = \frac{1}{2}\partial_\rho g_{\mu\nu} + \Gamma_{\rho\mu\nu}$$

where the first addendum comes from the last equality above and for the second addendum one must remember that the Christoffel symbol is symmetric $\Gamma_{\rho\mu\nu} = \Gamma_{\rho\nu\mu}$ and therefore

$$\Gamma_{\rho(\mu\nu)} = \frac{1}{2}(\Gamma_{\rho\mu\nu} + \Gamma_{\rho\nu\mu}) = \Gamma_{\rho\mu\nu}$$

One then obtains

$$\Gamma_{\rho\mu\nu} = \partial_{(\mu} g_{\nu)\rho} - \frac{1}{2}\partial_\rho g_{\mu\nu} = \frac{1}{2}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

which can be manipulated back to obtain the explicit expression of the Christoffel symbol:

$$\Gamma^\sigma_{\mu\nu} = g^{\sigma\rho}\Gamma_{\rho\mu\nu} = \frac{1}{2}g^{\sigma\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

Remember that the metric depends on the space-time coordinates, so one cannot contract the inverse metric with the derivatives (also one would get a Kronecker delta and therefore zero). This torsionless connection is called Levi–Civita connection. The Christoffel symbols are the connection coefficients of the Levi–Civita connection in a coordinate basis.

One may want to consider torsion-full connections and those lead to the same results: the lack of torsion is not an important property, it is just for convenience.

The expression above is the constructive proof that Γ is a connection — so it transforms in a particular way —, that connections exist, and that it is a torsionless connection. Connections constitute an affine space: the differences of the elements of such a space constitute a vector space. So one can obtain a generic connection from the Levi–Civita connection. In fact if Γ_1 and Γ_2 are connections, then $\Gamma_1 - \Gamma_2$ is a type $(1, 2)$ tensor. A torsion-full connection can be obtain from the Levi–Civita connection by adding an arbitrary $(1, 2)$ tensor, which is a new field.

In General Relativity one assumes to be always working in the Levi–Civita connection.

4.3 Riemann tensor

In this discussion, curvature has not been defined for space-time. Consider again the analogy of electromagnetism in quantum mechanics

$$D_i\psi = (\partial_i - \mathrm{i}A_i)\psi$$

One can calculate the second derivative

$$\begin{aligned} D_iD_j\psi &= (\partial_i - \mathrm{i}A_i)(\partial_j - \mathrm{i}A_j)\psi = \partial_i\partial_j\psi - \mathrm{i}A_i\,\partial_j\psi - \mathrm{i}\,\partial_i(A_j\psi) - A_iA_j\psi \\ &= \partial_i\partial_j\psi - \mathrm{i}A_i\,\partial_j\psi - \mathrm{i}(\partial_iA_j)\psi - \mathrm{i}A_j\,\partial_i\psi - A_iA_j\psi \end{aligned}$$

Then

$$D_{[i}D_{j]}\psi = -\mathrm{i}(\partial_{[i}A_{j]})\psi = -2\mathrm{i}F_{ij}\psi\,, \quad D_{[i}D_{j]}\psi = \frac{1}{2}[D_i, D_j]$$

with ij space indices. A similar relation can be obtain in in field theory where the covariant derivative is used to write the lagrangian of a field that couples with the electromagnetic field. The covariant derivative contains the four-potential, while the anti-symmetrization of two covariant derivatives contains the physical field.

Since the covariant derivative in General Relativity is inspired by the above derivative, one can wonder the same about applying the covariant derivative twice. In fact, the anti-symmetrization of the covariant derivative is proportional to the commutator of the two covariant derivatives which is a closed path starting and ending at the same point: the product D_iD_j goes first along

[1] See `https://en.wikipedia.org/wiki/Ricci_calculus#Symmetric_and_antisymmetric_parts`.

the path i then the path j, similarly the product D_jD_i goes first along the path j then the path i; by changing the sign to one of them $D_iD_j - D_jD_i$ one goes along i then j, then goes along i backwards and finally j backwards arriving at the starting point (provided that the vector fields defining the directions of transport commute $[v, w] = 0$ otherwise the path does not close).

Therefore

$$\begin{aligned}\nabla_\mu\nabla_\nu\nu^\rho &= \partial_\mu(\nabla_\nu v^\rho) - \Gamma^\sigma_{\mu\nu}(\nabla_\sigma v^\rho) + \Gamma^\rho_{\mu\sigma}(\nabla_\nu v^\sigma)\\ &= \partial_\mu(\partial_\nu v^\rho + \Gamma^\rho_{\nu\lambda}v^\lambda) - \Gamma^\sigma_{\mu\nu}(\partial_\sigma v^\rho + \Gamma^\rho_{\sigma\lambda}v^\lambda) + \Gamma^\rho_{\mu\sigma}(\partial_\nu v^\sigma + \Gamma^\sigma_{\nu\lambda}v^\lambda)\\ &= \partial_\mu\partial_\nu v^\rho + (\partial_\mu\Gamma^\rho_{\nu\lambda})v^\lambda + \Gamma^\rho_{\nu\lambda}\,\partial_\mu v^\lambda - \Gamma^\sigma_{\mu\nu}\partial_\sigma v^\rho - \Gamma^\sigma_{\mu\nu}\Gamma^\rho_{\sigma\lambda}v^\lambda + \Gamma^\rho_{\mu\sigma}\partial_\nu v^\sigma + \Gamma^\rho_{\mu\sigma}\Gamma^\sigma_{\nu\lambda}v^\lambda\end{aligned}$$

In the first row one can explicitly write the first covariant derivative as a type $(1, 1)$ tensor $T^\rho_\nu = \nabla_\nu v^\rho$ so it is easier to apply the previous rules for covariant derivatives. By anti-symmetrization, the first, fourth and fifth terms disappear, in particular these last two thanks to the torsionless property.

$$\begin{aligned}\nabla_{[\mu}\nabla_{\nu]}v^\rho &= (\partial_{[\mu}\Gamma^\rho_{\nu]\lambda})v^\lambda + \Gamma^\rho_{[\nu|\lambda}\,\partial_{|\mu]}v^\lambda + \Gamma^\rho_{[\mu|\sigma}\partial_{|\nu]}v^\sigma + \Gamma^\rho_{[\mu|\sigma}\Gamma^\sigma_{|\nu]\lambda}v^\lambda\\ &= (\partial_{[\mu}\Gamma^\rho_{\nu]\lambda})v^\lambda + \Gamma^\rho_{[\nu|\lambda}\,\partial_{|\mu]}v^\lambda - \Gamma^\rho_{[\nu|\lambda}\partial_{|\mu]}v^\lambda + \Gamma^\rho_{[\mu|\sigma}\Gamma^\sigma_{|\nu]\lambda}v^\lambda\\ &= [(\partial_{[\mu}\Gamma^\rho_{\nu]\lambda}) + \Gamma^\rho_{[\mu|\sigma}\Gamma^\sigma_{|\nu]\lambda}]v^\lambda = \frac{1}{2}R^\rho{}_{\lambda\mu\nu}v^\lambda\end{aligned}$$

In the third term of the second line, the index σ is renamed to λ since it is summed over and $\mu\nu$ are exchanged by adding a minus sign.

The cancellations that have happened are similar to the ones for the electromagnetism analogy. Though, the counterpart term of A_iA_j is $\Gamma^\rho_{[\mu|\sigma}\Gamma^\sigma_{|\nu]\lambda}$ which does not cancel thanks to the presence of the matrix indices. In particular, the terms involving the derivatives of v^μ simplify.

Since the covariant derivative of a tensor is a tensor, then the double covariant derivative is also a tensor. This implies that the right hand side of the equation above is a tensor. In particular, since v^λ is vector field, the term in parentheses is a type $(1, 3)$ tensor called Riemann tensor

$$R^\rho{}_{\sigma\mu\nu} \equiv \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}$$

This is a particular combination of Christoffel symbols that is a tensor. This is analogous to how the four-vector is not gauge invariant, but the combination that gives the field strength $F_{\mu\nu}$ is.

The argument above regarding the double covariant derivative can now be understood in terms of the Riemann tensor. The infinitesimal difference between a vector transported first along direction μ then ν, and a vector transported along ν then μ is of order

$$\varepsilon\varepsilon' R^\rho{}_{\sigma\mu\nu}v^\sigma$$

The Riemann tensor quantifies how much one cannot keep a vector pointing in the same direction during parallel transport and as such it is a way to measure curvature.

Lecture 9

lun 30 ott 2023 14:30

4.3.1 Properties

These properties are algebraic and relate its components.

First. The Riemann tensor is anti-simmetric in the last two indices

$$R^\rho{}_{\sigma\mu\nu} = -R^\rho{}_{\sigma\nu\mu}$$

The last two indices describe the directions along which a vector is transported while the first two indices describes the action done on those vectors. [r]

Second. One would like to study the anti-symmetrization of the covariant derivative for any tensor. For a function one has

$$\nabla_\nu f = \partial_\nu f \,, \quad \nabla_\mu \nabla_\nu f = \nabla_\mu \partial_\nu f = \partial_\mu \partial_\nu f - \Gamma^\rho_{\mu\nu} \partial_\rho f$$

therefore

$$\nabla_{[\mu} \nabla_{\nu]} f = \partial_{[\mu} \partial_{\nu]} f - \Gamma^\rho_{[\mu\nu]} \partial_\rho f = 0$$

the first term is null thanks to Schwarz while the second is null because the Christoffel symbol is torsionless. Applying Leibinz rule one can deduce the expression for a one-form

$$\begin{aligned} 0 &= \nabla_{[\mu} \nabla_{\nu]} (v^\rho \omega_\rho) = \nabla_{[\mu} [(\nabla_{\nu]} v^\rho) \omega_\rho + v^\rho \nabla_{\nu]} \omega_\rho] \\ &= (\nabla_{[\mu} \nabla_{\nu]} v^\rho) \omega_\rho + \nabla_{[\nu} v^\rho \, \nabla_{\mu]} \omega_\rho + \nabla_{[\mu} v^\rho \, \nabla_{\nu]} \omega_\rho + v^\rho \nabla_{[\mu} \nabla_{\nu]} \omega_\rho \\ &= \frac{1}{2} R^\rho{}_{\sigma\mu\nu} v^\sigma \omega_\rho + v^\rho \nabla_{[\mu} \nabla_{\nu]} \omega_\rho \end{aligned}$$

Therefore

$$v^\rho \nabla_{[\mu} \nabla_{\nu]} \omega_\rho = -\frac{1}{2} R^\rho{}_{\sigma\mu\nu} v^\sigma \omega_\rho = -\frac{1}{2} v^\rho R^\sigma{}_{\rho\mu\nu} \omega_\sigma \,, \quad \forall v$$

since it is true for any v, then

$$\nabla_{[\mu} \nabla_{\nu]} \omega_\rho = -\frac{1}{2} R^\sigma{}_{\rho\mu\nu} \omega_\sigma$$

This analogous to the partial derivative and Lie derivative. One can generalize the formula for any tensor because one knows how to act on both contravariant and covariant indices.

If two indices are anti-symmetric, one can add a third by anti-symmetrizing all three at once

$$\nabla_{[\mu} \nabla_\nu \omega_{\rho]} = \nabla_{[[\mu} \nabla_{\nu]} \omega_{\rho]}$$

From this follows

$$-\frac{1}{2} R^\sigma{}_{[\rho\mu\nu]} \omega_\sigma = \nabla_{[\mu} \nabla_\nu \omega_{\rho]} = \nabla_{[\mu} \nabla_{[\nu} \omega_{\rho]]} = \partial_{[\mu} \nabla_{[\nu} \omega_{\rho]]} = \partial_{[\mu} \partial_{[\nu} \omega_{\rho]]} = \partial_{[\mu} \partial_\nu \omega_{\rho]} = 0$$

at the third and fourth equality one applies

$$\nabla_{[\nu} \omega_{\rho\sigma]} = \partial_{[\nu} \omega_{\rho\sigma]} \,, \quad \nabla_{[\nu} \omega_{\rho]} = \partial_{[\nu} \omega_{\rho]} \,,$$

thanks to the Christoffel symbol being torsionless. Therefore, one gets the first Bianchi identity

$$R^\sigma{}_{[\rho\mu\nu]} = 0 \implies R^\sigma{}_{\rho\mu\nu} + R^\sigma{}_{\mu\nu\rho} + R^\sigma{}_{\nu\rho\mu} = 0$$

Third. The anti-symmetrization can be done for a type $(0,2)$ tensor

$$\nabla_{[\mu} \nabla_{\nu]} T_{\rho\sigma} = -\frac{1}{2} R^\lambda{}_{\rho\mu\nu} T_{\lambda\sigma} - \frac{1}{2} R^\lambda{}_{\sigma\mu\nu} T_{\rho\lambda}$$

Applying this formula to the metric, one gets

$$0 = \nabla_{[\mu} \nabla_{\nu]} g_{\rho\sigma} = -\frac{1}{2} R^\lambda{}_{\rho\mu\nu} g_{\lambda\sigma} - \frac{1}{2} R^\lambda{}_{\sigma\mu\nu} g_{\rho\lambda} = -\frac{1}{2} R_{\sigma\rho\mu\nu} - \frac{1}{2} R_{\rho\sigma\mu\nu}$$

This means that the Riemann tensor is anti-symmetric in the first two indices

$$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$$

Fourth. From the first Bianchi identity one gets

$$R_{\mu\nu\rho\sigma} + R_{\mu\rho\sigma\nu} + R_{\mu\sigma\nu\rho} = 0 \implies R_{\mu\nu\rho\sigma} = -R_{\mu\rho\sigma\nu} + R_{\mu\sigma\rho\nu} = 2R_{\mu[\sigma\rho]\nu} = -2R_{[\sigma|\mu|\rho]\nu} = 2R_{[\sigma|\mu\nu|\rho]}$$

Raising the indices with the inverse metric, one obtains

$$R^{\mu\nu}{}_{\rho\sigma} = 2R^\mu{}_{[\sigma\rho]}{}^\nu \implies R^{[\mu\nu]}{}_{\rho\sigma} = 2R^{[\mu}{}_{[\sigma\rho]}{}^{\nu]} = 2R_{[\sigma}{}^{[\mu\nu]}{}_{\rho]}$$

[r]

$$R^{\rho\sigma}{}_{\mu\nu} = 2R^{[\rho}{}_{[\nu\mu]}{}^{\sigma]} = 2R^{[\sigma}{}_{[\mu\nu]}{}^{\rho]} \implies R_{\rho\sigma}{}^{\mu\nu} = 2R_{[\sigma}{}^{[\mu\nu]}{}_{\rho]} = \cdots$$

Therefore, the Riemann tensor is symmetric under exchange of the pairs of indices

$$R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu}$$

Independent components. One can count the number of independent components. One can count how many pairs of anti-symmetric indices are present. Two indices $\mu\nu$ give

$$\binom{d}{2} = \frac{1}{2}d(d-1)$$

choices. Also, the first Bianchi identity puts

$$d\binom{d}{3}$$

constraints. So the number of independent components is

$$\binom{d}{2} - d\binom{d}{3} = \frac{d^2}{12}(d^2-1)$$

For space-time, $d = 4$, there are 20 independent components. In two dimensions there is only one independent component R_{0101}.

Fifth. Consider the third property applied to

$$2\nabla_{[\mu}\nabla_{\nu]}\nabla_\rho\omega_\sigma = -R^\lambda{}_{\rho\mu\nu}\nabla_\lambda\omega_\sigma - R^\lambda{}_{\sigma\mu\nu}\nabla_\rho\omega_\lambda$$

[r] Therefore

$$2\nabla_{[\mu}\nabla_\nu\nabla_{\rho]}\omega_\sigma = -R^\lambda{}_{\sigma[\mu\nu}\nabla_{\rho]}\omega_\lambda$$

Also it holds

$$2\nabla_{[\mu}\nabla_\nu\nabla_{\rho]}\omega_\sigma = \nabla_{[\mu|}(-R^\lambda{}_{\sigma|\nu\rho]}\omega_\lambda) = -\nabla_{[\mu}R_{\nu\rho]}{}^\lambda{}_\sigma\omega_\lambda - R^\lambda{}_{\sigma[\nu\rho}\nabla_{\mu]}\omega_\lambda$$

These two equations mean that

$$\nabla_{[\mu}R_{\nu\rho]\lambda\sigma} = 0$$

This is the second Bianchi identity.

4.4 Geodesics

One can apply the concept of curvature to Physics. Free particles no longer move in straight lines. In space, the new definition of straight line is a line that minimizes length. In space-time a straight line maximizes proper time (as one can already see from Special Relativity). In Minkowski space, the action of a free relativistic particle is proportional to proper time.[r]

$$S = -m\int \mathrm{d}\lambda\,\sqrt{-\eta_{\mu\nu}\,\mathrm{d}_\lambda x^\mu\,\mathrm{d}_\lambda x^\nu}$$

where one chooses $\lambda = \tau$, $\mathrm{d}\tau^2 = -g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu$. In space-time the expression becomes

$$S = -m\int \mathrm{d}\lambda\,\sqrt{-g_{\mu\nu}\,\mathrm{d}_\lambda x^\mu\,\mathrm{d}_\lambda x^\nu}$$

where a particle is considered in free fall: only gravity acts. The metric represents the gravitational field. This describes how space-time tells matter how to move, but not how mass curves space-time. A path that minimizes the action is a geodesic. [r]

Equations of motion. To find the equations of motion, one needs to compute the variation of the action δS. In the action one considers the fields $x^\mu(\lambda)$ as functions of the parameter λ. So one has

$$\begin{aligned}\delta S &\propto \int \mathrm{d}\lambda\,\delta\sqrt{-g_{\mu\nu}\,\mathrm{d}_\lambda x^\mu\,\mathrm{d}_\lambda x^\nu} = \int \frac{\mathrm{d}\lambda}{2\sqrt{-g_{\mu\nu}\,\mathrm{d}_\lambda x^\mu\,\mathrm{d}_\lambda x^\nu}}\delta(-g_{\mu\nu}\,\mathrm{d}_\lambda x^\mu\,\mathrm{d}_\lambda x^\nu)\\ &= -\int \frac{\mathrm{d}\lambda}{2\sqrt{-g_{\mu\nu}\,\mathrm{d}_\lambda x^\mu\,\mathrm{d}_\lambda x^\nu}}[g_{\mu\nu}\partial_\lambda(\delta x^\mu)\,\partial_\lambda x^\nu + g_{\mu\nu}\,\partial_\lambda x^\mu\,\partial_\lambda(\delta x^\nu) + \partial_\rho g_{\mu\nu}\,\delta x^\rho\,\partial_\lambda x^\mu\,\partial_\lambda x^\nu]\end{aligned}$$

[r] The metric depends on the space-time point.

Lecture 10

gio 02 nov 2023 14:30

Introducing

$$\dot{x}^2 \equiv g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \,, \quad \dot{x}^\mu \equiv \partial_\lambda x^\mu$$

one gets

$$0 = \delta S \propto \int \mathrm{d}\lambda \, \delta\sqrt{-\dot{x}^2}$$

The first two terms of the variation are the same: it suffices to rename $\mu \leftrightarrow \nu$ and using the fact that the metric is symmetric. Therefore

$$\begin{aligned}\int \mathrm{d}\lambda \, \delta\sqrt{-\dot{x}^2} &= -\int \mathrm{d}\lambda \, \frac{1}{2\sqrt{-\dot{x}^2}} \left[2g_{\mu\nu}\,\partial_\lambda(\delta x^\mu)\partial_\lambda x^\nu + \partial_\rho g_{\mu\nu}\,\delta x^\rho\,\partial_\lambda x^\mu\,\partial_\lambda x^\nu\right] \\ &= \frac{1}{2}\int \mathrm{d}\lambda \, \delta x^\mu \left[\partial_\lambda\left(\frac{2g_{\mu\nu}\,\partial_\lambda x^\nu}{\sqrt{-\dot{x}^2}}\right) - \frac{\partial_\rho g_{\mu\nu}\,\partial_\lambda x^\mu\,\partial_\lambda x^\nu}{\sqrt{-\dot{x}^2}}\right]\end{aligned}$$

[r] second addendum does not have $\sqrt{-\dot{x}^2}$?
At the second line, one integrates by parts remembering that the variation at the extrema is null. The variation is zero for all paths δx^μ if and only if the square bracket is zero

$$\partial_\lambda\left(\frac{2g_{\mu\nu}\,\partial_\lambda x^\nu}{\sqrt{-\dot{x}^2}}\right) - \frac{\partial_\rho g_{\mu\nu}\,\partial_\lambda x^\mu\,\partial_\lambda x^\nu}{\sqrt{-\dot{x}^2}} = 0$$

To simplify the result one can see that the parameter λ can be arbitrary

$$\mathrm{d}\lambda = \partial_{\lambda'}\lambda\,\mathrm{d}\lambda' \,, \quad \partial_\lambda = \partial_\lambda\lambda'\,\partial_{\lambda'}$$

and not change the action. The action has a reparameterization invariance. This freedom can be used to simplify the equations of motion. One chooses λ so that

$$\dot{x}^2 = -1 \implies g_{\mu\nu}\,\mathrm{d}_\lambda x^\mu\,\mathrm{d}_\lambda x^\nu = -1 \implies \mathrm{d}\tau = \mathrm{d}\lambda$$

where τ is the proper time. Therefore

$$2\,\partial_\tau(g_{\nu\rho}\dot{x}^\nu) - \partial_\rho g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = 0$$

Knowing that

$$\partial_\tau g_{\nu\rho} = \partial_\mu g_{\nu\rho}\,\partial_\tau x^\mu = \partial_\mu g_{\nu\rho}\dot{x}^\mu$$

[r] means

$$\begin{aligned}0 &= 2\,\partial_\mu g_{\nu\rho}\dot{x}^\mu\dot{x}^\nu + 2g_{\nu\rho}\ddot{x}^\nu - \partial_\rho g_{\mu\nu}\,\dot{x}^\mu\dot{x}^\nu = g_{\nu\rho}\ddot{x}^\nu + \dot{x}^\mu\dot{x}^\nu(\partial_\mu g_{\nu\rho} - \frac{1}{2}\partial_\rho g_{\mu\nu}) \\ &= \ddot{x}^\sigma + g^{\rho\sigma}(\partial_{(\mu}g_{\nu)\rho} - \frac{1}{2}\partial_\rho g_{\mu\nu})\dot{x}^\mu\dot{x}^\nu\end{aligned}$$

[r] Therefore

$$\ddot{x}^\sigma + \Gamma^\sigma_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = 0$$

In general

$$\partial_\lambda(\dot{x}^2) = \partial_\lambda(g_{\mu\nu}\,\partial_\lambda x^\mu\,\partial_\lambda x^\nu) = \partial_\rho g_{\mu\nu}\,\dot{x}^\rho\dot{x}^\mu\dot{x}^\nu + g_{\mu\nu}\ddot{x}^\mu\dot{x}^\nu + g_{\mu\nu}\dot{x}^\mu\ddot{x}^\nu = 2g_{\mu\sigma}(\ddot{x}^\mu + \Gamma^\mu_{\nu\rho}\dot{x}^\nu\dot{x}^\rho)\dot{x}^\sigma$$

From this, one can rewrite the equations of motion as

$$(\dot{x}^2\delta^\mu{}_\nu - \dot{x}^\mu g_{\nu\alpha}\dot{x}^\alpha)(\ddot{x}^\nu + \Gamma^\nu_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma) = 0$$

If the matrix is not invertible, the vector is in the kernel of the matrix. [r]

To solve $S = 0$ one can consider

$$S' = \frac{1}{2}\int \mathrm{d}\lambda \left(\frac{\dot{x}^2}{e} - em^2\right)$$

where e is a new dynamical object. When $m \neq 0$, its equations of motion of e can be used to obtain $S' = S$. When $m = 0$, the action S' implies again

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\rho}\dot{x}^\nu\dot{x}^\rho = 0$$