General Relativity

Maso*

March 23, 2024

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.

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 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.

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

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 $\tau^2 = t^2 - |\mathbf{x}|^2$. 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: $P^\mu = (E, \mathbf{p})$. The invariant associated with the norm of the vector is mass $m^2 = P_\mu P^\mu$. For an object at rest, $\mathbf{p} = 0$, 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 $\mathbf{p} = \gamma m \mathbf{v}$. In the limit of low speeds, one recovers the classical relations

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

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

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

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

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

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]

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

[r] and image

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

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

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

then

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

The metric is then

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

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 $-\mathbf{v}$. 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

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

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

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

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 $v^\mu(x^\mu) = v_0^\mu$ 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

$$v^\mu(x+\varepsilon z) \sim (v^\mu - \varepsilon\Gamma^\mu_{\nu\rho} v^\rho z^\nu)(x) = v_0^\mu - \varepsilon\Gamma^\mu_{\nu\rho}(x) v^\rho z^\nu(x)$$

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

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

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.

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

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 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

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

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

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

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

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

Considering the four-velocity to be $u^\mu \equiv \dot{x}^\mu$, the covariant derivative of a vector is

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

which implies the geodesic equation

$$u^\mu\nabla_\mu u^\nu = u^\mu\partial_\mu u^\nu + u^\mu\Gamma^\nu_{\mu\rho}u^\rho = \partial_\tau u^\nu + u^\mu\Gamma^\nu_{\mu\rho}u^\rho = \ddot{x}^\nu + u^\mu\Gamma^\nu_{\mu\rho}u^\rho = \ddot{x}^\nu + \Gamma^\nu_{\mu\rho}\dot{x}^\mu\dot{x}^\rho = 0$$

In the second equality it holds

$$u^\mu\,\partial_\mu = \dot{x}^\mu\,\partial_\mu = \partial_\tau x^\mu\,\partial_\mu = \partial_\tau$$

From this, when $\lambda = \tau$, the equation of motion can be written as a covariant derivative. Remembering that a vector z^μ is parallel transported along a curve defined by the vector field v^μ if the directional derivative $v^\mu\nabla_\mu z^\nu = 0$; therefore, the equation of motion is describing the parallel transport of the four-velocity along the world line by keeping the four-velocity u^ν tangent to the trajectory defined by itself u^μ: the velocity does not change in direction nor absolute value and is constantly parallel to the direction of motion. The equation of motion is called geodesic equation.

For $\lambda \neq \tau$, the directional derivative is proportional to the four-velocity itself: the four-velocity cannot change direction, but can grow and shrink because not imposing $\dot{x}^2 = -1$ means not fixing the norm.

This computation is exactly the same when one seeks the shortest distance between two points on an arbitrary manifold where there is no time, just space. For example, this computation gives the great circle for two points on a sphere, S^2.

When the metric $g_{\mu\nu}$ is static (which is a coordinate dependent notion), that is it does not depends on time $\partial_0 g_{\mu\nu} = g_{0i} = 0$, then one can choose $\lambda = x^0$ in the action

$$S = -m \int \mathrm{d}\lambda \, \sqrt{-\dot{x}^2} = -m \int \mathrm{d}\lambda \, \sqrt{-g_{00} + g_{ij}\dot{x}^i \dot{x}^j} \sim \int \mathrm{d}\lambda \, (T - V)$$

where at the last step one assumes small velocities $\dot{x}^i$, one K is the kinetic energy while V is the potential energy

$$V = m\sqrt{-g_{00}}$$

This notion is useful to build intuition, for example to derive the equation of motion for the metric.

Rindler geodesics. As an example of geodesics, one can calculated the ones associated to the Rindler metric with $a = 1$

$$\mathrm{d}s^2 = \mathrm{e}^{2\xi}(-\mathrm{d}\eta^2 + \mathrm{d}\xi^2)$$

Instead of solving the geodesic equation

$$u^\mu \nabla_\mu u^\nu = 0$$

which in some cases is actually easier, one can vary the action

$$S = -m \int \mathrm{d}\lambda \, \sqrt{-\dot{x}^2} = -m \int \mathrm{d}\lambda \, \sqrt{\mathrm{e}^{2\xi}(-\dot{\eta}^2 + \dot{\xi}^2)} = -m \int \mathrm{d}\lambda \, \mathrm{e}^{\xi} \sqrt{-\dot{\eta}^2 + \dot{\xi}^2} \,, \quad \lambda = \tau$$

Skipping the calculations, the equations of motion are

$$\partial_\tau(\mathrm{e}^{2\xi}\dot{\eta}) = 0 \,, \quad \partial_\tau(\mathrm{e}^{2\xi}\dot{\xi}) = -1$$

The first implies a conserved quantity

$$\mathrm{e}^{2\xi}\dot{\eta} = -E$$

Later one can see that this quantity can be directly obtained from Noether's theorem. Imposing $\dot{x}^2 = -1$ one gets a linear combination of the equations of motion from which the second one can be replaced by

$$-\dot{\eta}^2 + \dot{\xi}^2 = -\mathrm{e}^{-2\xi} \implies \mathrm{d}_\eta \xi = \sqrt{1 - E^{-2}\mathrm{e}^{2\xi}} \implies \cosh(\eta - \eta_0) = E\mathrm{e}^{-\xi}$$

These trajectories in the plane (ξ, η) are straight lines only asymptotically. The potential is an exponential barrier

$$V = m\mathrm{e}^{\xi}$$

For the massless case, one has

$$\dot{x}^2 = 0 \implies \dot{\eta}^2 = \dot{\xi}^2 \implies \mathrm{d}\eta = \pm\,\mathrm{d}\xi$$

Lecture 11

lun 06 nov 2023 14:30

5 Einstein's field equations

One can study how mass creates a gravitational field by bending space-time.

5.1 Original derivation

Newtonian mechanics can be recovered from General Relativity when the metric is almost the Minkowski metric

$$g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}\,, \quad h_{\mu\nu} \ll 1$$

which corresponds to a weak gravitational field. One wants to look for equations of motion of the metric such that they are written in terms of tensors and reproduce Newtonian gravity in the weak gravitational field regime. One may notice that the metric is dimensionless, so it makes sense to state $h_{\mu\nu} \ll 1$.

To get Newtonian gravity one can calculate the Newtonian potential energy. Let m be the test particle mass. In electromagnetism, the potential ϕ obeys Poisson's law

$$\nabla^2 \phi = -\frac{\rho}{\varepsilon_0}$$

Since Newtonian gravity is very similar to electromagnetism, one expects such a law to appear. In fact, it holds

$$\nabla^2 V_{\text{grav}} = 4\pi m G_{\text{N}} \rho = m \nabla^2 \sqrt{-g_{00}}$$

where ρ is the mass density and G_{N} is the Newtonian constant of universal gravitation [r]. The sign difference is given by the fact that gravity is attractive. In the weak field regime, one can expand the square root in a power series to get

$$\sqrt{-g_{00}} \sim \sqrt{1 - h_{00}} \sim 1 - \frac{1}{2} h_{00}$$

Therefore

$$\nabla^2 h_{00} = -8\pi G_{\text{N}} \rho$$

This equation already has some geometrical flavor. [r] The metric is described in terms of the mass density. This equation is not written in tensor form: there's only one component of a type 2 tensor and it contains a partial derivative (which of a tensor is not a tensor).

The idea it is to promote these quantities so that they are parts of tensors and obtain the above expressions in the weak and static field assumptions. This guessing method (which is Einstein's) is not the only method possible. Looking at the right-hand side, the mass (more appropriately energy) density is the first component of the stress-energy tensor. This tensor has been seen in Special Relativity, but it does not follow that it is a tensor also in General Relativity: one wants to promote a lagrangian and its stress-energy tensor to tensor in General Relativity. For now, one can only assume its existence. The left-hand side is a bit more difficult because one does not know a tensor whose 00-th component is the one above. This tensor cannot be $\nabla^2 g_{\mu\nu}$ because there are only partial derivatives. One can try the covariant derivative, but it is zero. One can look at the tensors already constructed: in particular, the Riemann tensor contains second derivatives of the metric (derivative of the Christoffel symbols which contain derivatives of the metric). Though the left-hand side has only two indices, not four. One can hope to modify the Riemann tensor into the Ricci tensor

$$R_{\mu\nu} \equiv R_{\rho\mu\sigma\nu} g^{\rho\sigma} = R^{\rho}{}_{\mu\rho\nu}$$

contracting other indices gives zero thanks to the anti-symmetry of the Riemann tensor. This tensor is symmetric

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

One needs to check that its 00-th derivative in the weak and static field regime gives the expression above. The Christoffel symbols are

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

In the weak field only assumption above, the metric is

$$g_{\mu\nu} \sim \eta_{\mu\nu} + h_{\mu\nu} \quad g^{\mu\nu} \sim \eta^{\mu\nu} - h^{\mu\nu}$$

Now the indices are moved up-down with the Minkowski metric. The Christoffel symbols become

$$\Gamma^\mu_{\nu\rho} \sim \frac{1}{2}(\eta^{\mu\nu} - h^{\mu\nu})(\partial_\nu h_{\rho\sigma} + \partial_\rho h_{\nu\sigma} - \partial_\sigma h^{\nu\rho}) \sim \frac{1}{2}\eta^{\mu\nu}(\partial_\nu h_{\rho\sigma} + \partial_\rho h_{\nu\sigma} - \partial_\sigma h^{\nu\rho})$$

Also

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

The Riemann tensor is

$$R^\mu{}_{\nu\rho\sigma} = \partial_\rho \Gamma^\mu_{\sigma\nu} - \partial_\sigma \Gamma^\mu_{\rho\nu} + \Gamma^\mu_{\rho\lambda}\Gamma^\lambda_{\sigma\nu} - \Gamma^\mu_{\sigma\lambda}\Gamma^\lambda_{\rho\nu} \sim \partial_\rho \Gamma^\mu_{\sigma\nu} - \partial_\sigma \Gamma^\mu_{\rho\nu}$$

Lowering the index with Minkowski metric one gets

$$\begin{aligned} R_{\mu\nu\rho\sigma} &\sim 2\,\partial_{[\rho|}\Gamma_{\mu|\sigma]\nu} = \partial_{[\rho|}(\partial_{\sigma]} h_{\nu\mu} + \partial_\nu h_{|\sigma]\mu} - \partial_\mu h_{|\sigma]\nu}) \\ &= \partial_\nu \partial_{[\rho} h_{\sigma]\mu} - \partial_\mu \partial_{[\rho} h_{\sigma]\nu} = \partial_{[\nu|}\partial_{[\rho} h_{\sigma]|\mu]} \\ &= \frac{1}{2}(\partial_\nu\partial_\rho h_{\sigma\mu} - \partial_\nu\partial_\sigma h_{\rho\mu} - \partial_\mu\partial_\rho h_{\sigma\nu} + \partial_\mu\partial_\sigma h_{\rho\nu}) \end{aligned}$$

The Ricci tensor is then

$$\begin{aligned} R_{\nu\sigma} = g^{\mu\rho} R_{\mu\nu\rho\sigma} \sim \eta^{\mu\rho} R_{\mu\nu\rho\sigma} = \frac{1}{2}(\partial_\nu \partial^\rho h_{\sigma\rho} - \partial_\nu\partial_\sigma h^\rho{}_\rho - \partial_\rho\partial^\rho h_{\nu\sigma} + \partial^\rho\partial_\sigma h_{\nu\rho}) \\ = -\frac{1}{2}\Box h_{\nu\sigma} + \partial_{(\nu}\partial^\rho h_{\sigma)\rho} - \frac{1}{2}\partial_\nu\partial_\sigma h^\rho{}_\rho \end{aligned}$$

Adding the static assumption one gets

$$R_{00} = -\frac{1}{2}\nabla^2 h_{00}$$

[r]

Therefore, assembling the equation gives

$$R_{\mu\nu} = 4\pi G_{\mathrm{N}} T_{\mu\nu}$$

Though this relation is not totally correct. In Special Relativity one express the conservation of four-momentum as

$$\partial_\mu T^{\mu\nu} = 0$$

while in General Relativity one has

$$\nabla_\mu T^{\mu\nu} = 0$$

Substituting in the equation derived above has

$$\nabla_\mu R^{\mu\nu} = 4\pi G_{\mathrm{N}} \nabla_\mu T^{\mu\nu} = 0 \implies \nabla_\mu R^{\mu\nu} = 0$$

Though, from the second Bianchi identity

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

then contracting μ and σ (since they don't give zero) one has

$$0 = \nabla^\mu R_{\nu\rho\mu\lambda} + \nabla_\nu R_{\rho\mu\mu\lambda} + \nabla_\rho R_{\nu\lambda} = \nabla^\mu R_{\nu\rho\mu\lambda} - \nabla_\nu R_{\rho\lambda} + \nabla_\rho R_{\nu\lambda}$$

where the second addendum is

$$R_{\rho\mu\mu\lambda} = g^{\mu\sigma} R_{\rho\mu\sigma\lambda} = -R_{\rho\lambda}$$

Contracting ν and λ one gets

$$0 = -\nabla^\mu R_{\rho\mu} - \nabla^\nu R_{\rho\nu} + \nabla_\rho R_\nu{}^\nu \implies \nabla_\rho R = 2\nabla^\mu R_{\mu\rho}$$

where one has the Ricci scalar (also called scalar curvature)

$$R \equiv R_\mu{}^\mu = g^{\mu\nu} R_{\mu\nu}$$

If the equation of motion found is correct, then R would be constant. There would be a constraint with three derivatives

$$\partial_\mu R = 0$$

which does not even depend on matter. The Riemann tensor can only be contracted in one way and from the Ricci tensor one can only do one contraction to get the Ricci scalar. Multiplying it by the metric, one gets another type 2 tensor. From the covariant derivative of the Ricci tensor one has

$$\nabla_\rho R = 2\nabla^\mu R_{\mu\rho} \implies 0 = \nabla^\mu \left(R_{\mu\rho} - \frac{1}{2} g_{\mu\rho} R \right) \equiv \nabla^\mu G_{\mu\rho}$$

where $G_{\mu\rho}$ is the Einstein tensor. The equation of motion should then be corrected by

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \alpha T_{\mu\nu}$$

where it holds

$$\nabla^\mu T_{\mu\nu} = 0 \implies \nabla^\mu \left(R_{\mu\rho} - \frac{1}{2} g_{\mu\rho} R \right) = 0$$

To determine the constant one can take the trace (that is multiply by $g^{\mu\nu}$)

$$R - \frac{1}{2} g^{\mu\nu} g_{\mu\nu} R = \alpha g^{\mu\nu} T_{\mu\nu} \implies R - \frac{1}{2} 4R = \alpha T \implies R = -\alpha T$$

[r] The trace-reversed form is

$$R_{\mu\nu} = \alpha \left[T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right]$$

In the weak and static field limit one has

$$R_{00} = \alpha \left[T_{00} - \frac{1}{2} g_{00} T \right] \implies -\frac{1}{2} \nabla^2 h_{00} = \alpha \left(T_{00} - \frac{1}{2} T_{00} \right) = \frac{1}{2} \alpha T_{00}$$

where

$$T = T^0{}_0 + T^i{}_i = T^0{}_0$$

Remembering that the derived equation was

$$\nabla^2 h_{00} = 8\pi G_\mathrm{N} \rho$$

then one has

$$\alpha = 8\pi G_\mathrm{N}$$

This leads to Einstein's field equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G_\mathrm{N} T_{\mu\nu}$$

The new method (Hilbert's) guesses the action, not the equations of motion: the action is a scalar and there are not that many options.

Lecture 12

gio 09 nov 2023 14:30

5.2 Action

For any dynamics one integrates the lagrangian over time. The action is expressed as an integral of a lagrangian density

$$S = \int \mathrm{d}^4 x \, \mathcal{L}$$

The lagrangian density has a geometrical interpretation in Special Relativity, but in General Relativity one needs to modify the expression above because the measure is not invariant

$$\mathrm{d}^4 x' = (\det J) \, \mathrm{d}^4 x \,, \quad J^\mu{}_\nu = \partial_\nu x'^\mu$$

[r] absolute value on determinant? The solution is to look at the metric. The measure is related to the infinitesimal volume but it does not involve the metric. Recalling that the metric transforms as

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

One notices that

$$\det g' = \det\bigl(J^{-1} g J^{-1}\bigr) = (\det J^{-1})^2 \det g = \frac{\det g}{(\det J)^2}$$

Denoting $g \equiv \det g < 0$, one can construct an invariant measure

$$\sqrt{-g'} \, \mathrm{d}^4 x' = \sqrt{-g} \, \mathrm{d}^4 x$$

An action in a theory that is well-defined under general coordinate changes is of the form

$$S = \int \mathrm{d}^4 x \sqrt{-g} \mathcal{L}$$

The lagrangian density is invariant and thus is a function. One may look for a lagrangian density for the metric. For the metric to have a dynamic, the lagrangian has to contain the metric and its derivatives. One such scalar is the Ricci scalar. [r] One obtains the Einstein–Hilbert action

$$S \propto \int \mathrm{d}^4 x \sqrt{-g} R$$

Equations of motion. Varying the action, one gets

$$\delta S \propto \int \mathrm{d}^4 x \, (\delta \sqrt{-g} R + \sqrt{-g} \, \delta R)$$

where

$$R = g^{\mu\nu} R_{\mu\nu} \,, \quad \delta R = \delta g^{\mu\nu} \, R_{\mu\nu} + g^{\mu\nu} \, \delta R_{\mu\nu}$$

The variation of the square root is

$$\delta \sqrt{-g} = -\frac{\delta g}{2\sqrt{-g}} \implies \frac{\delta\sqrt{-g}}{\sqrt{-g}} = \frac{1}{2} \frac{\delta g}{g}$$

The variation of the determinant is lengthier. Consider a diagonal matrix D, its determinant is the product of the non-zero entries

$$\det D = d_1 \cdots d_n \,, \quad \log \det D = \sum_i \log d_i$$

The variation is then

$$\frac{\delta \det D}{\det D} = \delta(\log \det D) = \sum_i \delta(\log d_i) = \sum_i \frac{\delta d_i}{d_i} = \sum_i d_i^{-1} \, \delta d_i = \operatorname{Tr} D^{-1} \, \delta D$$

This result is true for any diagonalizable matrix $M = C^{-1} D C$:

$$\delta(\log \det M) = \operatorname{Tr} M^{-1} \, \delta M$$

and also for non diagonalizable matrices. Applying this result to the case above, one gets

$$\frac{\delta g}{g} = \frac{\delta \det g}{\det g} = \operatorname{Tr}\bigl(g^{-1} \, \delta g\bigr) = \operatorname{Tr}(g^{\mu\rho} \, \delta g_{\rho\nu}) = g^{\mu\rho} \, \delta g_{\rho\mu} = g^{\mu\rho} \, \delta g_{\mu\rho}$$

[r] One would like to write the variation for the inverse metric

$$g^{\mu\rho} g_{\rho\nu} = \delta^\mu{}_\rho \implies 0 = \delta(g^{\mu\rho} g_{\rho\nu}) = \delta g^{\mu\rho} \, g_{\rho\nu} + g^{\mu\rho} \, \delta g_{\rho\nu} \implies \delta g^{\mu\nu} = -g^{\mu\rho} \, (\delta g_{\rho\sigma}) g^{\sigma\nu}$$

The middle expression also implies

$$-\delta g^{\mu\nu} \, g_{\mu\nu} = g^{\mu\nu} \, \delta g_{\mu\nu} = \frac{\delta g}{g} = 2\frac{\delta\sqrt{-g}}{\sqrt{-g}}$$

The variation of the Ricci scalar is

$$\delta R = \delta g^{\mu\nu}\, R_{\mu\nu} + g^{\mu\nu}\, \delta R_{\mu\nu}$$

The only term to calculate is the last variation. One needs to evaluate the variation of the Riemann tensor

$$\frac{1}{2} R^{\mu}{}_{\nu\rho\sigma} = \partial_{[\rho} \Gamma^{\mu}_{\sigma]\nu} + \Gamma^{\mu}_{[\rho|\lambda} \Gamma^{\lambda}_{\sigma]\nu}$$

getting

$$\frac{1}{2} \delta R^{\mu}{}_{\nu\rho\sigma} = \partial_{[\rho} \delta\Gamma^{\mu}_{\sigma]\nu} + \delta\Gamma^{\mu}_{[\rho|\lambda} \Gamma^{\lambda}_{\sigma]\nu} + \Gamma^{\mu}_{[\rho|\lambda} \delta\Gamma^{\lambda}_{\sigma]\nu}$$

In this expression it is not clear that the right-hand side is a tensor. Since the finite difference between two Christoffel symbols is a tensor, this also happens for the variation. One would like to manifest the tensor nature. One substitutes the partial derivative for the covariant derivative

$$\nabla_{[\rho} \delta\Gamma^{\mu}_{\sigma]\nu} = \partial_{[\rho} \delta\Gamma^{\mu}_{\sigma]\nu} + \Gamma^{\mu}_{[\rho|\lambda}\, \delta\Gamma^{\lambda}_{\sigma]\nu} - \Gamma^{\lambda}_{[\rho\sigma]}\, \delta\Gamma^{\mu}_{\lambda\nu} - \Gamma^{\lambda}_{[\rho|\nu}\, \delta\Gamma^{\mu}_{\sigma]\lambda}$$

The third addendum is zero. The second term is the last of the variation of the Riemann tensor, while the last term is the same as Riemann's second

$$\Gamma^{\lambda}_{[\rho|\nu}\, \delta\gamma^{\mu}_{\sigma]\lambda} = \delta\Gamma^{\mu}_{[\sigma|\lambda}\, \Gamma^{\lambda}_{\rho]\nu} = -\delta\Gamma^{\mu}_{[\rho|\lambda}\, \Gamma^{\lambda}_{\sigma]\nu}$$

Therefore

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

[r] The variation of the Ricci tensor is then

$$\frac{1}{2} R_{\nu\sigma} = \frac{1}{2} R^{\mu}{}_{\nu\rho\sigma} = \frac{1}{2} (\nabla_{\mu} \delta\Gamma^{\mu}_{\sigma\nu} - \nabla_{\sigma} \delta\Gamma^{\mu}_{\mu\nu})$$

The contraction with an inverse metric is

$$g^{\nu\sigma} \delta R_{\nu\sigma} = \frac{1}{2} (\nabla_{\mu} \delta\Gamma^{\mu}{}_{\sigma}{}^{\sigma} - \nabla_{\sigma} \delta\Gamma^{\mu}{}_{\mu}{}^{\sigma}) = \nabla_{\mu} (\delta\Gamma^{\mu}{}_{\sigma}{}^{\sigma} - \delta\Gamma^{\sigma}{}_{\sigma}{}^{\mu}) \equiv \nabla_{\mu} V^{\mu}$$

This term looks like a total derivative, in fact

$$\nabla_{\mu} V^{\nu} = \partial_{\mu} V^{\nu} + \Gamma^{\nu}_{\mu\rho} V^{\rho} \implies \nabla_{\mu} V^{\mu} = \partial_{\mu} V^{\mu} + \Gamma^{\mu}_{\mu\rho} V^{\rho}$$

The second addendum is given by contracting

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

to get

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

Since the first addendum is symmetric and the metric before is symmetric, then

$$g^{\mu\sigma}\, \partial_{\mu} g_{\rho\sigma} = g^{\sigma\mu}\, \partial_{\sigma} g_{\rho\mu} = g^{\mu\sigma}\, \partial_{\sigma} g_{\rho\mu}$$

in the first equality one renames the indices, in the second one uses the symmetry of the metric. Therefore

$$\Gamma^{\mu}_{\mu\rho} = \frac{1}{2} g^{\mu\rho}\, \partial_{\rho} g_{\mu\sigma}$$

which is similar to

$$\frac{1}{2} g^{\mu\rho}\, \delta g_{\mu\rho} = \frac{\delta\sqrt{-g}}{\sqrt{-g}}$$

[r] So one has

$$\Gamma^{\mu}_{\mu\rho} = \frac{1}{\sqrt{-g}}\, \partial_{\rho} \sqrt{-g}$$

Going back to the beginning of the discussion, one has

$$\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu_{\mu\rho} V^\rho = \partial_\mu V^\mu + (\frac{1}{\sqrt{-g}}\,\partial_\rho \sqrt{-g}) V^\rho = \frac{1}{\sqrt{-g}} \partial_\rho (\sqrt{-g} V^\rho)$$

which is explicitly a total derivative:

$$\int \mathrm{d}^4 x\, \sqrt{-g} g^{\mu\nu}\, \delta R_{\mu\nu} = \int \mathrm{d}^4 x\, \sqrt{-g}\, \nabla_\mu V^\mu = \int \mathrm{d}^4 x\, \partial_\mu (\sqrt{-g} V^\mu)$$

Finally, putting everything together, one has

$$\delta S \propto \delta \int \mathrm{d}^4 x\, \sqrt{-g} R = \int \mathrm{d}^4 x\, (\delta \sqrt{-g}\, R + \sqrt{-g}\, \delta g^{\mu\nu}\, R_{\mu\nu}) = \int \mathrm{d}^4 x\, \sqrt{-g}\, \delta g^{\mu\nu} \left(-\frac{1}{2} g_{\mu\nu} R + R_{\mu\nu} \right)$$

The term inside the parentheses is the Einstein tensor which has to be guesses when one uses the original derivation. Therefore, the equations of motion are

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 0$$

This action describes gravity without any matter: this is the reason why there is a zero. A consequence of this equation is that, taking the trace, one gets that $R = 0$.

Adding matter fields. In general, the matter action is

$$S_\mathrm{m} = \int \mathrm{d}^4 x\, \sqrt{-g} \mathcal{L}_\mathrm{m}$$

The total action is then

$$S = S_\mathrm{EH} + S_\mathrm{m} = \int \mathrm{d}^4 x\, \sqrt{-g} (\alpha R + \mathcal{L}_\mathrm{m})$$

whose variation is

$$\delta S = \int \mathrm{d}^4 x\, \sqrt{-g}\, \delta g^{\mu\nu} \left[\alpha \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \right] + \frac{\delta(\sqrt{-g} \mathcal{L}_\mathrm{m})}{\sqrt{-g}\, \delta g^{\mu\nu}}$$

where one defines the stress-energy tensor like

$$T_{\mu\nu} \equiv -2 \frac{\delta(\sqrt{-g} \mathcal{L}_\mathrm{m})}{\sqrt{-g}\, \delta g^{\mu\nu}}$$

from which Einstein's field equations are

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} = \frac{1}{2\alpha} T_{\mu\nu}\,, \quad \alpha = \frac{1}{16\pi G_\mathrm{N}}$$

Notice that there are multiple definitions of the stress-energy tensor. The definition provided is naturally symmetric. In Theoretical Physics I, one has seen the Belinfante tensor which is symmetric and equivalent to the above.

Example of stress-energy tensor. For electromagnetism, the free field action in Special Relativity is

$$S = -\frac{1}{4} \int \mathrm{d}^4 x\, F_{\mu\nu} F^{\mu\nu}\,, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

while in General Relativity it is

$$S = -\frac{1}{4} \int \mathrm{d}^4 x\, \sqrt{-g} F_{\mu\nu} F^{\mu\nu}\,, \quad F^{\mu\nu} = g^{\mu\rho} g^{\nu\sigma} F_{\rho\sigma}$$

where one does not have to generalize the field tensor because the anti-symmetrization keeps the tensorial structure. [r] Varying the action with respect to A^μ one gets Maxwell's equations. The homogeneous ones do not need the covariant derivative and follow directly from the definition of the field strength tensor, while the inhomogeneous ones have

$$\nabla_\mu F^{\mu\nu} = 0$$

and are found by varying the action.

For the stress-energy tensor, one varies with respect to the metric:

$$\delta S = -\frac{1}{4}\int \mathrm{d}^4x\,(\delta\sqrt{-g}\,F_{\mu\nu}F^{\mu\nu} + 2F_{\mu\nu}\,\delta g^{\mu\rho}\,g^{\nu\sigma}F_{\rho\sigma})$$

Also it holds

$$\delta\sqrt{-g} = \frac{1}{2}\sqrt{-g}(-\delta g^{\rho\sigma}\,g_{\rho\sigma})$$

giving

$$\delta S = -\frac{1}{2}\int \mathrm{d}^4x\,\sqrt{-g}\,\delta g^{\mu\nu}\left(F_\mu{}^\rho F_{\rho\nu} - \frac{1}{4}g_{\mu\nu}F^{\rho\sigma}F_{\rho\sigma}\right)$$

Lecture 13

lun 13 nov 2023 14:30

Remark. The quantum properties of the action are not desirable

$$S = \frac{1}{16\pi G_\mathrm{N}}\int \mathrm{d}^4x\,\sqrt{-g}R$$

because the action has mass dimension 2 (instead of [r]). The quantum analysis shows that the action leads to non-renormalizable theories. One must add new terms like any power of R or any scalar. In the context of the renormalization group one must focus on the mass dimension of the lagrangian: mass dimensions up and including 4 are important at low energies. From General Relativity, a key point is the importance of the equivalence principle. From a quantum point of view, it becomes necessary to include more general functions of the curvature.

Normal. One would like to see manifest the equivalence principle from the theory developed. One expects that at every point of space-time, there is a change of coordinates such that the metric is approximately the flat metric remembering that the metric is identified with the gravitational field.

One would like to find such coordinates system. At a single point one can apply Linear Algebra to know that the metric is diagonalizable (because it is symmetric) and then one can normalize the basis vectors to obtain the flat metric. One is interested in a set of points. Consider a point p and a geodesic $x^\mu(\lambda)$ to another point. The velocity vector at p is

$$a^\mu = \partial_\lambda x(\lambda = 0) = \dot{x}^\mu(0)$$

One chooses the coordinates such that the second point is at λa^μ. A point is labeled by how long in terms of λ it takes to go along a geodesic with velocity a^μ. A point at a parameter λ along a geodesic with $\dot{x}^\mu(0) = a^\mu$ has coordinates $x^\mu = \lambda a^\mu$. From the geodesic equation

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

The velocity a^μ is constant so the acceleration is zero [r] while the Christoffel symbols depend on the point. At $\lambda = 0$ one has

$$\Gamma^\mu_{\nu\rho}(0)a^\nu a^\rho = 0\,,\quad \forall a \implies \Gamma^\mu_{\nu\rho}(p) = 0$$

One condition for [r] is

$$\nabla_\mu g_{\nu\rho} = 0 \implies (\partial_\mu g_{\nu\rho})(p) = 0$$

[r] At first order in displacement, the metric does not change. In fact, the power series of the metric is

$$g_{\mu\nu} = g_{\mu\nu}(p) + (\partial_\rho g_{\mu\nu})(p)x^\rho + \frac{1}{2}(\partial_\rho\partial_\sigma g_{\mu\nu})(p)x^\rho x^\sigma + o(x^2)$$

From Linear Algebra one sets the first addendum to $\eta_{\mu\nu}$, while the procedure above gives zero first derivative. At second order one can differentiate the geodesic equation

$$0 = \partial_\lambda(\Gamma^\mu_{\nu\rho}(\lambda a) a^\nu a^\rho)$$

and then set $\lambda = 0$. This gives

$$\partial_\sigma \Gamma^\mu_{\nu\rho} a^\sigma a^\nu a^\rho = 0\,, \quad \forall a \implies \partial_{(\sigma}\Gamma^\mu_{\nu\rho)}(p) = 0$$

This implies (see notes) that

$$\partial_\sigma\partial_\mu g_{\nu\rho}(p) = \frac{2}{3} R_{\mu(\nu\rho)\sigma}(p)$$

Therefore, the power series becomes

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{3} R_{\mu\rho\sigma\nu}(p) x^\rho x^\sigma + o(x^2)$$

These are called normale coordinates. The deviation from the flat metric is given by curvature, the tidal forces.

6 Solutions

Given Einstein's fields equations one can look for solutions. One first searches simple solutions, so ones with lots of symmetries.

6.1 Isometries

A symmetry is a coordinate transformation under which the metric does not change

$$g'_{\mu\nu} = \partial'_\mu x^\rho \, \partial'_\nu x^\sigma g_{\rho\sigma} \equiv g_{\mu\nu}$$

When a coordinate change has this property, it is called isometry. In euclidean space, the isometries of $\mathrm{d}x^i \, \mathrm{d}x^i$ are three translations, three rotations and reflections. In Minkowski space, the isometries of $\eta_{\mu\nu} \, \mathrm{d}x^\mu \, \mathrm{d}x^\nu$ are four translations, three spatial rotations, three boosts.

In general, isometries form a group. Often the group is a Lie group. The isometries of the euclidean space constitute the euclidean group. Working with a Lie group, one considers maps near the identity with the notion of Lie algebra. Any transformation near the identity is associated with a vector field

$$x'^\mu \sim x^\mu + \varepsilon v^\mu(x)$$

This implies an infinite dimensional Lie algebra. [r] One may look for a transformation near the identity that is also an isometry of the metric

$$g_{\mu\nu} = g'_{\mu\nu} \sim g_{\mu\nu} + \varepsilon(L_v g)_{\mu\nu} \implies (L_v g)_{\mu\nu} = 0$$

A vector v with this property is said to be a Killing vector. The Lie derivative is

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

Consider

$$\begin{aligned}
\nabla_\mu v_\nu + \nabla_\nu v_\mu = 2\nabla_{(\mu} v_{\nu)} &= 2\,\partial_{(\mu} v_{\nu)} - 2\Gamma^\rho_{\mu\nu} v_\rho = 2\,\partial_{(\mu}(g_{\nu)\rho} v^\rho) - 2\Gamma^\rho_{\mu\nu} g_{\rho\sigma} v^\sigma \\
&= 2(\partial_{(\mu} g_{\nu)\rho}) v^\rho + 2 g_{\rho(\nu}\,\partial_{\mu)} v^\rho - (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) v^\rho \\
&= v^\rho \, \partial_\rho g_{\mu\nu} + \partial_{\rho\nu} \, \partial_\mu v^\rho + g_{\rho\mu} \, \partial_\nu v^\rho = (L_v g)_{\mu\nu}
\end{aligned}$$

So v is a Killing vector if and only if

$$\nabla_{(\mu} v_{\nu)} = 0$$

More generally, the expression of the Lie derivative does not change when replacing partial derivatives with the covariant derivatives. Also it holds

$$[L_v, L_w] = L_{[v,w]}$$

From this it follows that the Lie bracket of two Killing vectors is also a Killing vector. Killing vectors constituted a finite dimensionale Lie algebra.

In the euclidean space, the generators of translations and rotations are Killing vectors

$$\partial_i \,, \quad -l^3 = x^1\, \partial_2 - x^2 \partial_1 \,, \quad \text{etc.}$$

Exercise: check that these are Killing vectors and that the Lie bracket is also a Killing vector.

6.2 Spherical symmetry

One can study the symmetries for a single point particle. One expects the same symmetries of euclidean space: rotations but not translations. For the spherical symmetry one would like to find a group to be isomorphic to the isometry group as the rotations in euclidean space SO(3). In spherical coordinates, the euclidean line element s

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

The rotations leave the radius invariant and act only on the angles. A metric of the form

$$g_{tt}\,\mathrm{d}t^2 + 2g_{tr}\,\mathrm{d}t\,\mathrm{d}r + g_{rr}\,\mathrm{d}r\,\mathrm{d}r + f(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)$$

has an isometry group isomorphic to SO(3) if the coefficients g and f do not depend on the angles. More concretely, the metric above has three Killing vector fields. In the spherical coordinates, the rotations generators are

$$l_3 = -\partial_\varphi \,, \quad l_1 = \sin\varphi\,\partial_\theta + \cos\varphi\cot\theta\,\partial_\varphi \,, \quad l_2 = -\cos\varphi\,\partial_\theta + \sin\varphi\cot\theta\,\partial_\varphi$$

These can be obtained from the cartesian coordinates using a change of basis and the associated Jacobian. [r]

It is always possible to change coordinates to set g_{tr} to zero.

Lecture 14

gio 16 nov 2023 14:30

6.3 Schwarzschild solution

One can make a few simplifying assumptions. There is still a freedom in the parametrization thanks to the value of f. Without loss of generality, one sets $f = r^2$. One can also assume that g_{tt} and g_{rr} do not depend on time t. This is equivalent to the assumption that there is a time translation symmetry. The corresponding Killing vector is $K^\mu\,\partial_\mu = K = \partial_t$. One then focuses on static solutions

$$(L_K g)_{\mu\nu} = 0 \iff \partial_t g_{\mu\nu} = 0$$

Since one also has $g_{ti} = 0$, the solution is static. Therefore

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

and the metric is diagonal

$$g_{\mu\nu} = \mathrm{diag}(g_{tt}, g_{rr}, r^2, r^2\sin^2\theta)$$

One can see that the first eigenvalue must be negative, while the second one must be positive, since the other two are already positive. It is convenient to define

$$U(r)\,, V(r) \implies g_{tt} = -\mathrm{e}^{2U} \,, \quad g_{rr} = \mathrm{e}^{2V}$$

Starting with this ansatz, one computes the equations of motion. The Christoffel symbols are

$$\Gamma^\mu_{\nu\rho} = \frac{1}{2}g^{\mu\sigma}(\partial_\nu g_{\rho\sigma} + \partial_\rho g_{\nu\sigma} - \partial_\sigma g_{\mu\nu}) \,, \quad x^0 = t\,, \quad x^1 = r\,, \quad x^2 = \theta\,, \quad x^3 = \varphi$$

First one has

$$\Gamma^0_{00} = \frac{1}{2}g^{00}\,\partial_0 g_{00} = 0$$

Without the above assumptions this would not have been zero. Second

$$\Gamma^0_{01} = \frac{1}{2} g^{00}(\partial_0 g_{10} - \partial_1 g_{00} - \partial_0 g_{10}) = \frac{1}{2}(-\mathrm{e}^{2U})\,\partial_1(-\mathrm{e}^{2U}) = U'\,, \quad \partial_1 U \equiv U'$$

[r] The Christoffel symbols are 40, because the lower two indices are symmetric. Alternatively one can use the action of a massive particle

$$S \propto \int \mathrm{d}\lambda \sqrt{-\dot{x}^2} = \int \mathrm{d}\lambda \sqrt{-g\mathrm{e}^{2U}\dot{t}^2 + \mathrm{e}^{2V}\dot{r}^2 + r^2(\dot{\theta} + \sin^2\theta\dot{\varphi}^2)}$$

The variation is

$$\begin{aligned}
\delta \int \mathrm{d}\lambda \sqrt{-\dot{x}^2} &= \int \frac{\mathrm{d}\lambda}{2\sqrt{\cdots}} 2[-\mathrm{e}^{2U} 2\dot{t}\,\delta\dot{t} + \mathrm{e}^{2V}\dot{r}\,\delta\dot{r} + r^2(\dot{\theta}\,\delta\dot{\theta} + \sin^2\theta\,\dot{\varphi}\,\delta\dot{\varphi}) \\
&\quad + \delta r(-\mathrm{e}^{2U}U'\dot{t}^2 + \mathrm{e}^{2V}V'\dot{r}^2 + r(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2)) + r^2\sin\theta\cos\theta\,\delta\theta\dot{\varphi}^2] \\
&= \int \frac{\mathrm{d}\lambda}{\sqrt{\cdots}} [\delta t\,\partial_\tau(\mathrm{e}^{2U}\dot{t}) - \delta r\,\partial_\tau(\mathrm{e}^{2V}\dot{r}) - \delta\theta\,\partial_\tau(r^2\dot{\theta} - \delta\varphi\,\partial_\tau(r^2\sin^2\theta\,\dot{\varphi}) \\
&\quad + \delta r(-\mathrm{e}^{2U}U'\dot{t}^2 + \mathrm{e}^{2V}V'\dot{r}^2 + r(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2)) + r^2\sin\theta\cos\theta\,\delta\theta\dot{\varphi}^2]
\end{aligned}$$

at the third line one sets $\lambda = \tau$ and integrates by parts the first line. The equations of motion by collecting the terms multiplied by δt and δr

$$\partial_\tau(\mathrm{e}^{2U}\dot{t}) = 0\,, \quad \partial_\tau(\mathrm{e}^{2V}\dot{r}) + (-\mathrm{e}^{2U}U'\dot{t}^2 + \mathrm{e}^{2V}V'\dot{r}^2 + r(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2)) = 0$$

et cetera. Then one can compare this results to what one already knows the expression already is

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

In the first one, one gets

$$\partial_\tau(\mathrm{e}^{2U}\dot{t}) = \mathrm{e}^{2U}(\ddot{t} + 2U'\dot{r}\dot{t}) = 0$$

One sees that the only non zero symbol with upper 0 is Γ^0_{01}. One then has

$$U' = \Gamma^0_{01} = \Gamma^0_{10}$$

The second equation above, one gets which Γ^1 are non zero.

Both methods give

$$\Gamma^0_{01} = U'\,, \quad \Gamma^1_{00} = \mathrm{e}^{2(U-V)}V'\,, \quad \Gamma^1_{11} = V'\,, \quad \Gamma^1_{33} = \sin^2\theta\,\Gamma^1_{22} = \sin^2\theta r\sin^2\theta\mathrm{e}^{-2V}$$

and also

$$\Gamma^2_{12} = \frac{1}{r}\,, \quad \Gamma^2_{33} = -\sin\theta\cos\theta\,, \quad \Gamma^3_{13} = \frac{1}{r}\,, \quad \Gamma^3_{23} = \cot\theta$$

The Ricci tensor is

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

One notices that

$$\Gamma^\rho_{\rho\mu} = \frac{1}{2}\,\partial_\mu\log|g|\,, \quad |g| = \mathrm{e}^{2(U+V)}r^4\sin^2\theta\,, \quad \frac{1}{2}\log|g| = U + V + 2\log r + \log\sin\theta$$

Also one can easily organize terms by computing

$$\begin{aligned}
R_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \propto -\partial_\nu\Gamma^\rho_{\rho\mu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu &= \frac{1}{2}\mathrm{d}x^\nu\,\partial_\nu(\delta x^\mu\,\partial_\mu\log|g|) \\
&= \frac{1}{2}\mathrm{d}x^\nu\,\partial_\nu\left[\mathrm{d}r\left(U' + V' + \frac{2}{r}\right) + \mathrm{d}\theta\,\cot\theta\right] \\
&= \frac{1}{2}\,\mathrm{d}r^2\left(U'' + V'' - \frac{2}{r^2}\right) + \mathrm{d}\theta^2\frac{1}{\sin^2\theta}
\end{aligned}$$

[r] Similarly one can do for the other addenda. Defining

$$\Gamma^{\rho}{}_{\lambda} \equiv \Gamma^{\rho}_{\mu\lambda}\,\mathrm{d}x^{\mu}$$

for the last term, one has

$$\Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\rho}\,\mathrm{d}x^{\mu}\,\mathrm{d}\nu = \mathrm{Tr}\big(\Gamma^2\big)$$

The final result is

$$\begin{aligned}R_{\mu\nu}\,\mathrm{d}x^{\mu}\,\mathrm{d}x^{\nu} &= \mathrm{e}^{2(U-V)}\left[U'' + U'\left(\frac{2}{r} + U' - V'\right)\right]\mathrm{d}t^2 + \left[-U'' + U'(V'-U') + \frac{2}{r}V'\right]\mathrm{d}r^2 \\ &\quad + [r\mathrm{e}^{-2V}(V'-U') + 1 - \mathrm{e}^{-2V})(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2) \\ &= R_{tt}\,\mathrm{d}t^2 + R_{rr}\,\mathrm{d}r^2 + R_{\theta\theta}\,\mathrm{d}\theta^2 + R_{\theta\theta}\sin^2\theta\,\mathrm{d}\varphi^2\end{aligned}$$

The Ricci tensor also has spherical symmetry.

For a single particle, the stress-energy tensor $T_{\mu\nu}$ is null outside the origin. One cannot use the Dirac delta function because the theory is not linear. One solves the equations of motions outside the origin

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0 \implies R_{\mu\nu} = 0$$

which can be seen by taking the trace or using the trace-reversed form of the equation. From this, one has to set each term of the Ricci to zero. [r] There are two parameters but three equations and also a first order equation. This is because there is a gauge redundancy. In fact, gravity is a gauge theory and the gauge transformation is the coordinates transformation. There is a general theory that relates gauge invariance to symmetry and constraints [r]. In hamitlonian formalism, the second order equations of lagrangian mechanics become first order, and first order become constraints. In this sense gravity has constraints.

There is another first order equation given by combining two others

$$R_{tt} = R_{rr} = 0 \implies 0 = \frac{1}{2}(R_{tt} + R_{rr}) = U'\frac{1}{r} + \frac{1}{r}V' \implies (U+V)' = 0$$

so the sum of the two is constant which one sets to zero (the constant can be absorbed by $\mathrm{d}t$) without loss of generality. Similarly

$$0 = \frac{1}{2}(R_{tt} - R_{rr}) = U'' + U'\left(\frac{1}{r} + U' - V'\right) - \frac{1}{r}U'$$

finally

$$R_{\theta\theta} = 0 \implies r\mathrm{e}^{-2V}(V'-U') + 1 - \mathrm{e}^{-2V} = 0$$

From the first one, one gets

$$U'' + 2U'\left(U' + \frac{1}{r}\right) = 0\,, \quad 2rU' = \mathrm{e}^{-2U} - 1$$

The first one of these is given by the second by differentiation [r], so it is redundant. One can solve the second by integration or by observing that

$$2rU' = \mathrm{e}^{-2U} - 1 \iff \partial_r(r\mathrm{e}^{2U}) = 1 \implies r\mathrm{e}^{2U} = r - r_{\mathrm{s}} \implies \mathrm{e}^{2U} = 1 - \frac{r_{\mathrm{s}}}{r}$$

where r_{s} is the integration constant. Going back to the metric, and remembering $g_{tt} = -\mathrm{e}^{2U}$, one has

$$\mathrm{d}s^2 = -\left(1 - \frac{r_{\mathrm{s}}}{r}\right)\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{1 - \frac{r_{\mathrm{r}}}{r}} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)$$

This is the Schwarzschild metric and the integration constant is the Schwarzschild radius. One needs also to understand the physics of the solution.

Interpretation. This solution was found for point-like particles, but it can be applied to spherically symmetric bodies. The solution holds only outside the body because the solution has been calculated for the stress-energy tensor equal to zero. For an object close to be point-like, one finds some problems. For large radii, R_{tt} is almost one, but for radii close to the Schwarzschild radius, then time is null. A similar problem holds for R_{rr}. One needs to study if the theory stop working at these points or if the metric in a certain coordinate system is pathological, but not in other (like for the sphere).

Lecture 15

lun 20 nov 2023 14:30

One compares the solution with Newtonian gravity. In the limit $r \to \infty$, the metric $\mathrm{d}s^2$ tends to the Minkowski metric (in radial coordinates). When the radius is much greater than the Schwarzschild radius then the metric is almost Minkowski $g_{\mu\nu} \sim \eta_{\mu\nu}$ [r]. The gravitational energy used to motivate the theory is

$$V_{\mathrm{grav}} = m\sqrt{-g_{tt}} = m\sqrt{1 - \frac{r_{\mathrm{s}}}{r}} \sim m\left[1 - \frac{1}{2}\frac{r_{\mathrm{s}}}{r} + o({}^{r_{\mathrm{s}}}\!/_{r})\right] \sim m - \frac{r_{\mathrm{s}} m}{2r}$$

which agrees with the potential energy in Newtonian gravity

$$V_{\mathrm{grav}} = -G_{\mathrm{N}}\frac{mM}{r}$$

Comparing the two, one gets

$$r_{\mathrm{s}} = 2G_{\mathrm{N}}M$$

The potential is not defined for radii smaller than Schwarzschild's. In fact, one has already noticed that the metric for $r \leq r_{\mathrm{s}}$ has no immediate physical meaning, but needs clarification. For most situations, the Schwarzschild radius is not relevant. Often times, the radius of a celestial body is much greater than Schwarzschild's, so if one wants to study gravity below the radius of the body, one has to use another metric, since the one studied assumes that the stress-energy tensor (which describes matter) is null.

Though, there are objects for with the metric applies. The theoretical understanding of the solution came well after the solution, and the empirical evidence came even later.

Horizon. [r] One would like to understand if locus of $r = r_{\mathrm{s}}$ is an artefact of the coordinates system or not. One can study the radial light-like geodesics (which are simpler): so from outside the radius, one shines light directly towards the center. The geodesics equations can be avoided, like it was done for the Rindler metric noting that $\dot{x}^2 = 0$. [r] Remembering that $g_{tt} = -\mathrm{e}^{2U}$ one has

$$0 = \dot{x}^2 = -\mathrm{e}^{2U}\dot{t}^2 + \mathrm{e}^{-2U}\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2) = -\mathrm{e}^{2U}\dot{t}^2 + \mathrm{e}^{-2U}\dot{r}^2$$

The angles are constant since the geodesics is radial. One has

$$\pm\dot{t} = \mathrm{e}^{-2U}\dot{r} \implies \pm\mathrm{d}t = \mathrm{e}^{-2U}\,\mathrm{d}r = \frac{\mathrm{d}r}{1 - \frac{r_{\mathrm{s}}}{r}}$$

from which

$$\begin{aligned}\pm(t - t_0) &= \int \frac{\mathrm{d}r}{1 - \frac{r_{\mathrm{s}}}{r}} = \int \frac{r\,\mathrm{d}r}{r - r_{\mathrm{s}}} = \int \frac{r - r_{\mathrm{s}} + r}{r - r_{\mathrm{s}}}\,\mathrm{d}r \\ &= \int \left(1 + \frac{r_{\mathrm{s}}}{r - r_{\mathrm{s}}}\right)\mathrm{d}r = r + r_{\mathrm{s}}\ln\frac{r - r_{\mathrm{s}}}{r_{\mathrm{s}}} \equiv r_\star\end{aligned}$$

This is called tortoise coordinate: it is as if the light ray slows down at the Schwarzschild radius. The plus is for ingoing light rays, while the minus is for outgoing light rays. There is a divergence at $r = r_{\mathrm{s}}$. From this one can suspect that the radius is an artefact since ray never reaches such point.

One can send a massive object and consider the proper time of such observer. For an observer at infinity, the above time is also the proper time. Consider an object at fixed $r_1 > r_{\mathrm{s}}$ and consider an observer far away $r_0 \gg r_\mathrm{s}$. The object must accelerate away from the origin to counteract the gravitational potential. The light rays from the object are not straight lines and if they are spaced by Δt when they are being sent, they are also Δt when they are being received by the observer (as measured in the reference of the observer? [r]): there is no explicit dependance on time [r]. However the proper time of the object is

$$\mathrm{d}\tau = \sqrt{-\mathrm{d}s^2} = \sqrt{1 - \frac{r_\mathrm{s}}{r}}\,\mathrm{d}t$$

[r] For an object at $r = r_1$ one has

$$\Delta\tau = \sqrt{1 - \frac{r_\mathrm{s}}{r_1}}\,\Delta t$$

while for an observer at $r = r_0$ one has

$$\Delta\tau_0 = \sqrt{1 - \frac{r_\mathrm{s}}{r_0}}\Delta t = \frac{\sqrt{1 - \frac{r_\mathrm{s}}{r_0}}}{\sqrt{1 - \frac{r_\mathrm{s}}{r_1}}}\Delta\tau_1$$

The observer sees a greater interval between the pulses. More generally one has

$$\frac{\Delta\tau_0}{\Delta\tau_1} = \frac{\sqrt{-g_{tt}(r_0)}}{\sqrt{-g_{tt}(r_1)}}$$

This is the gravitational redshift.

One can calculate the proper time of a massive object moving towards the origin: an object on a massive geodesic. Therefore

$$-1 = \dot{x}^2 = \mathrm{e}^{2U}\dot{t}^2 = \mathrm{e}^{-2U}\dot{r}^2$$

One of the geodesic equations is

$$\partial_\tau(\mathrm{e}^{2U}\dot{t}) = 0 \implies \mathrm{e}^{2U}\dot{t} \equiv -E = \text{const.}$$

So one has

$$-1 = \mathrm{e}^{-2U}(E^2 - \dot{r}^2) \implies E^2 - \dot{r}^2 = -\mathrm{e}^{2U} \implies \dot{r} = \sqrt{E^2 - \mathrm{e}^{-2U}} = \sqrt{E^2 - 1 + \frac{r_\mathrm{s}}{r}}$$

The proper time along such a geodesic is

$$\tau = \int \mathrm{d}\tau = \int \frac{\mathrm{d}r}{\dot{r}} = \int \frac{\mathrm{d}r}{\sqrt{E^2 - \mathrm{e}^{2U}}}$$

For a particle at rest $\dot{r} = 0$ at $r = r_0$ then

$$E^2 = 1 - \frac{r_\mathrm{s}}{r_0} \implies \tau = \int_{r_0}^{r_\mathrm{s}} \frac{\mathrm{d}r}{\sqrt{\frac{r_\mathrm{s}}{r} - \frac{r_\mathrm{s}}{r_0}}} < \infty$$

A massive objects measures a finite time: its time is slower than an observer at infinity for which a light ray never arrives at the Schwarzschild radius.

The inner region is physical for the massive observer, but for an observer at infinity is not really physical. This analysis is inconclusive and requires additional mathematical tools to give a proper answer. There exists another coordinate system in which the Schwarzschild solution is not singular at the Schwarzschild radius. This relates to the possibility to connect every metric to the flat metric which depends on the curvature. [r] Changing the coordinates, one can have the Jacobian to go to zero faster than the curvature diverges. One can use scalars because they do not have Jacobians.[r] In this case, outside the origin, the scalar curvature is zero. One can build another scalar

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} = 12\frac{r_\mathrm{s}^2}{r^6}$$

This scalar is non zero, because if it were, one would have flat space-time. This scalar is finite at $r = r_\mathrm{s}$. Though, the origin diverges which agrees with the fact that it is also the location where one put the particle creating the gravitational field.

The correct coordinate change took several decades to find. One begins from the Schwarzschild metric and the tortoise coordinate to have

$$\mathrm{d}s^2 = \left(1 - \frac{r_\mathrm{s}}{r}\right)(-\mathrm{d}t^2 + \mathrm{d}r_\star^2) + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)$$

In fact, the tortoise coordinate, the light like geodesics are $\frac{\pi}{4}$ lines. One can think of r as a function of $r_\star$. In this coordinate system, the Schwarzschild radius has been pushed to infinity. [r] One would like to change coordinates in the first term so that the Jacobian cancels the first parenthesis. One may not use a one-coordinate change. [r] Defining the light-like coordinates

$$u_\pm = t \pm r_\star$$

one gets

$$-\mathrm{d}t^2 + \mathrm{d}r_\star^2 = -\mathrm{d}u_+\mathrm{d}u_-$$

Taking

$$u_\pm = u_\pm(v_\pm)$$

one then has

$$-\mathrm{d}u_+\mathrm{d}u_- = -\partial_{v_+}u_+\,\partial_{v_-}u_-\,\mathrm{d}v_+\,\mathrm{d}v_-$$

To cancel the parenthesis, one may remember that

$$r_\star = r + r_\mathrm{s}\ln\frac{r - r_\mathrm{s}}{r_\mathrm{s}} \implies \mathrm{e}^{\frac{r_\star}{r_\mathrm{s}}} = \mathrm{e}^{\frac{r}{r_\mathrm{s}}}\frac{r - r_\mathrm{s}}{r_\mathrm{s}}$$

Noting

$$r_\star = \frac{1}{2}(u_+ - u_-) \implies \mathrm{e}^{-r_\star} = \mathrm{e}^{-\frac{1}{2}u_+}\mathrm{e}^{\frac{1}{2}u_-}$$

and

$$v_\pm = \pm 2r_\mathrm{s}\mathrm{e}^{\pm\frac{u_\pm}{2r_\mathrm{s}}}, \quad \mathrm{d}v_\pm = \mathrm{e}^{\pm\frac{u_\pm}{2r_\mathrm{s}}}\,\mathrm{d}u_\pm$$

one gets

$$\mathrm{d}s^2 = \frac{r_\mathrm{s}}{r}\mathrm{e}^{-\frac{r}{r_\mathrm{s}}}(-\mathrm{d}v_+\,\mathrm{d}v_-) + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)$$

One can also define

$$v_\pm = T \pm R$$

so that

$$\mathrm{d}s^2 = \frac{r_\mathrm{s}}{r}\mathrm{e}^{-\frac{r}{r_\mathrm{s}}}(-\mathrm{d}T^2 + \mathrm{d}R^2) + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)$$

At the Schwarzschild radius, the first term is neither zero nor it diverges. The metric is non-singular at $r = r_\mathrm{s}$ (and closer to the origin), but the origin is still singular. So the Schwarzschild radius is a physical locus, but now one wonders what happens beyond it, for $r < r_\mathrm{s}$. This new coordinate system, called Krustal–Szekeres coordinates, is a valuable tool to analyse the metric.

Returning back to the radial light-like geodesic, one can analytically continue the solution and fix the tortoise coordinate. Inside the Schwarzschild radius, the term g_{tt} becomes positive, so time is a spatial coordinate, while the time coordinate is r. Inside, a light cone is flipped on its side and the light ray comes from an infinite t at $r < r_\mathrm{s}$ and goes to a finite t for $r = 0$. The Schwarzschild metric becomes difficult to work with when analysing the inside region.

Lecture 16

gio 23 nov 2023 14:30

In the above coordinates, inside the horizon time flows to lower radii. One can study whether the analytic continuation makes sense. To this end, one can use the above coordinates and plot the trajectory for a light ray. One would like to find the locus of the horizon

$$\mathrm{e}^{\frac{r_\star}{r_\mathrm{s}}} = -\frac{v_+v_-}{4r_\mathrm{s}^2}$$

The horizon $r = r_\mathrm{s}$ is given by the bisectors of the quadrants of the plane (R, T). The upper and lower regions have $r < r_\mathrm{s}$ while the left and right regions have $r > r_\mathrm{s}$. The locus $r = 0$ corresponds to an upper and a lower hyperbolae.[r] All the loci where r is constant correspond to hyperbolae which are in between the ones for $r = 0$. [r] The analytic continuation gives two copies of the inside and outside [r] image. The extra copies are valid when talking about the manifold, but physically are not accessible (the left and lower regions). [r]

One can study the regions. The purely radial light rays are radial both in RT and rt. The rays obey

$$\dot{x}^2 = 0 \implies \dot{T} = \pm\dot{R}$$

So again 45° lines. Rays that are not purely radial imply rays between the two diagonal lines. The possible future of a point in the right region (the outside) never involves the copy regions. The inbound light path crosses the horizon r_s at a right angle and one can notice that in this coordinate system there are no asymptotes. [r]

One can study the possible future of a point inside the horizon. The point can never exit the horizon, it is always inside. In fact, in the rt coordinates the only future is going to $r = 0$. For this reason r_s is called Schwarzschild horizon [r]. In the limiting case of being on the horizon, a light ray either goes in or stays put on the horizon. In the rt coordinates, the ray is vertical. Because light being unable to escape from the inside, the solution is called black hole. Once one gets to $r = 0$ there is no more space-time, curvature is infinite: one has a singularity.

When considering quantum mechanics, the other terms beside R become relevant for extreme curvature. So the singularity could be an artefact of using a linear action.

For a point in the below region, one can reach the outside on the right. The below region is called white hole.

In a diagram RT one must remember that R depends on T [r] so one can go infinitely far away from the black hole with a non-light speed.

From the left region, one can meet another observer only inside the black hole.

[r] For $R \to \infty$ one has $r \to \infty$ so one returns to a flat space-time. The same must happen for $R \to -\infty$. There are two copies of Minkowski space that are joined by the black hole: this is a wormhole (like worms in Newton's apple). The length of the wormhole depends on T and it becomes longer at a superluminal speed. There are also traversable wormholes.

Even though there are strong evidences of the existence of black holes, the same cannot be said about white holes and wormholes.

This solution is unrealistic because it states that black holes are eternal: they have always existed and will always exist, but this is not the case. If one also includes the existence of the star that generated the black hole, then the white holes and wormholes disappear.

Remark. The locus $r = 0$ is no longer a place in space, but a place in time. It is in the future of someone inside the black hole. Time flows in the radial direction.

6.4 Penrose diagrams

One can introduce a set of coordinates that is useful in general. The coordinates have a compact range and the future is still contained within 45° lines.

Minkowski space. Consider the flat space-time metric in radial coordinataes

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

The range is infinite because t can cover all of the real line, while r only the positive part. Though it is true that the future is contained within 45° lines. One would like to manipulate the first two coordinates to obtain a compact range. Like for the Krustal–Szekeres coordinates, one defines

$$u_\pm = t \pm r \implies -\mathrm{d}t^2 + \mathrm{d}r^2 = -\mathrm{d}u_+\mathrm{d}u_-$$

[r] One can look for a coordinate change that makes the range of us finite. One can use the tangent

$$u_\pm = \tan p_\pm\,, \quad p_\pm \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

In fact, the arctangent of the real line is a finite interval. [r] image. The region corresponding to $r > 0$ is $p_+ \geq p_-$. The range of these coordinates is compact. The metric becomes

$$-\mathrm{d}t^2 + \mathrm{d}r^2 = -\mathrm{d}u_+ \mathrm{d}u_- = \frac{-\mathrm{d}p_+\,\mathrm{d}p_-}{\cos^2 p_+ \cos^2 p_-}$$

One can also write r explicitly. One can define p^0 and p^1 such that

$$p_\pm = p^0 \pm p^1$$

from which

$$-\mathrm{d}t^2 + \mathrm{d}r^2 = \frac{-(\mathrm{d}p^0)^2 + (\mathrm{d}p^1)^2}{\cos^2 p_+ \cos^2 p_-}$$

When one studies radial rays, one has 45° lines because of

$$0 = \dot{x}^2 = \frac{1}{\dots}[-(\dot{p}^0)^2 + (\dot{p}^1)^2] \implies \mathrm{d}p^1 = \pm \mathrm{d}p^0 \implies p^1 = \pm(p^0 - \text{const.})$$

The lines of constant r are lines starting from the lower point, bending to the right and ending in the upper pint. The limit $r \to \infty$ is made of the two limiting straight lines. The future of a point is contained within two 45° lines and for a massive object the trajectories converge to the upper point. These coordinates can be represented in a Penrose diagram in which one ignores the radial coordinates.

Schwarzschild solution. Going back to the Schwarzschild solutions, one can make the change

$$v_\pm = 2r_\mathrm{s} \tan p_\pm$$

The new coordinates form a finite square, but one has to impose that the space-time accessible regions lie within the upper and lower hyperbolae of $r = 0$ in RT. The hyperbolae become two straight horizontal segments in the Penrose diagram and the horizon lines $r = r_\mathrm{s}$ are diagonal lines inside the cut diagram.

A more realistic black hole. Again, one can consider a non-eternal black hole. When a star lives time is similar to Minkowski, after it dies it changes space-time. Consider a star at constant $r = 0$ with its surface at r. When the star collapses it produces a horizon, so a diagonal line with a singularity above. The lower part is like Minkowski while the upper part is Schwarzschild. There are solutions that sketch this behaviour, like a shell that collapses. In this diagram, there is no white hole: the solution sees Minkowski space time and wormholes do not appear because one hits the surface of the star before being able to pass to the other copy of flat space-time.

The typical time from going from the horizon to the singularity can be estimated by using a light ray in the flat metric. Though, black holes can get very large, larger than any single body.

Lecture 17

lun 27 nov 2023 14:30

6.5 Schwarzschild geodesics

For the Schwarzschild metric

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

one assumes that

$$\dot{x}^2 = \begin{cases} -1\,, & m > 0 \\ 0\,, & m = 0 \end{cases}$$

to speed up calculations. One uses

$$\dot{x}^2 \equiv -\varepsilon = -\mathrm{e}^{2U}\dot{t}^2 + \mathrm{e}^{-2U}\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2)$$

[r] One can utilize the symmetries of the problem to simplify the calculation. In General Relativity a symmetry is an isometry which is a Killing vector at the infinitesimal level: ∂_t, l_i with $l_3 = -\partial_\varphi$. One expects some conserved quantities from Noether's theorem. This can be proven for a general geodesic motion. In general, given a Killing vector K^μ, the following quantity $K_\mu \dot{x}^\mu = g_{\mu\nu} K^\mu \dot{x}^\nu$ is conserved a long geodesics. In fact

$$\begin{aligned}\partial_\tau(K_\mu \dot{x}^\mu) &= \partial_\nu K_\mu \, \dot{x}^\nu \dot{x}^\mu + K_\rho \ddot{x}^\rho = (\partial_\nu K_\mu - K_\rho \Gamma^\rho_{\mu\nu}) \dot{x}^\mu \dot{x}^\nu = \nabla_\nu K_\mu \, \dot{x}^\mu \dot{x}^\nu \\ &= \nabla_{(\nu} K_{\mu)} \dot{x}^\nu \dot{x}^\mu = 0\end{aligned}$$

The Killing vector depends on the position which itself depends on τ, the proper time, parametrization of the geodesic. At the second equality one employs the geodesic equation

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

The last equation is null because the vector K is a Killing vector. There is another constant of motion, a generalization of a Killing vector called Killing tensor. In Schwarzschild it does not give anything new. For later it is important.

Letting

$$K = K^\mu \, \partial_\mu = \partial_t \,, \quad K^\mu = (1, 0, 0, 0)^\top$$

one has

$$K_\mu = g_{\mu\nu} K^\nu = (-\mathrm{e}^{2U}, 0, 0, 0)^\top$$

remembering that

$$g_{\mu\nu} = \mathrm{diag}(-\mathrm{e}^{2U}, \mathrm{e}^{-2U}, r^2, r^2 \sin^2\theta)_{\mu\nu}$$

The conserved quantity associated is

$$K_\mu \dot{x}^\mu = -\mathrm{e}^{2U} \dot{t} = E$$

which is energy. Also, from the geodesic equations one has

$$\partial_t(\mathrm{e}^{2U} \dot{t}) = 0$$

There are other conserved quantities. From

$$l_3 = -\partial_\varphi \,, \quad l_3^\mu = (0, 0, 0, 1)^\top \,, \quad (l_3)_\mu = (0, 0, 0, r^2 \sin^2\theta)^\top$$

one gets the z component of the angular momentum

$$(l_3)_\mu \dot{x}^\mu = -r^2 \sin^2\theta \, \dot{\varphi} = L_3$$

Doing the same for l_1 and l_2 one gets

$$L_1 = r^2(\sin\varphi \, \dot{\theta} + \cos\varphi \cot^2\theta \, \dot{\varphi}) \,, \quad L_2 = r^2(-\cos\varphi \, \dot{\theta} + \sin\theta \cot^2\theta \, \dot{\varphi})$$

Using these quantities, one has

$$\varepsilon = \mathrm{e}^{2U} \dot{t}^2 - \mathrm{e}^{-2U} \dot{r}^2 - r^2(\dot{\theta}^2 + \sin^2\theta \, \dot{\varphi}^2) = \mathrm{e}^{-2U} E^2 - \mathrm{e}^{-2U} \dot{r}^2 - r^2(\dot{\theta}^2 + \sin^2\theta \, \dot{\varphi}^2)$$

One can consider $\theta = \frac{\pi}{2}$ or compute

$$L^2 = L_i L_i = r^4(\dot{\theta}^2 + \sin^2\theta \, \dot{\varphi}^2)$$

One then has

$$\varepsilon = \mathrm{e}^{-2U} E^2 - \mathrm{e}^{-2U} \dot{r}^2 - \frac{L^2}{r^2}$$

In this way the problem has been reduced to a one-dimensional motion

$$E^2 = \varepsilon \mathrm{e}^{2U} + \dot{r}^2 + \frac{L^2}{r^2} \mathrm{e}^{2U} = \dot{r}^2 + V(r) \,, \quad V(r) = \left[\varepsilon + \frac{L^2}{r^2}\right] \left[1 - \frac{2GM}{r}\right]$$

where V is like an effective potential. Studying the problem in classical physics, one gets a similar expression, but with a different effective potential.

[r] One needs to solve

$$\dot{r}^2 = E - V(r) \implies \int \mathrm{d}t = \int \frac{\mathrm{d}r}{\sqrt{E^2 - V(r)}}$$

This is an elliptic integral. One can compute the precession of the perihelion. Letting $r = x^{-1}$, one has

$$\int \mathrm{d}t = \int \frac{\mathrm{d}x}{x^2\sqrt{P_3(x)}}$$

where P_4 is a cubic polynomial. To study the precession one can calculate how much more angle $\Delta\varphi$ one has when going from the aphelion to the perihelion [r]. Considering

$$(\mathrm{d}_\varphi r)^2 = \frac{\dot{r}^2}{\dot{\varphi}^2} = \dot{r}^2 \left(\frac{L}{r}\right)^{-2} = \frac{r^4}{L^2}[E^2 - V(r)]$$

from which

$$\Delta\varphi = \int \mathrm{d}\varphi = \int \frac{L}{r^2} \frac{\mathrm{d}r}{\sqrt{E^2 - V}} = L \int \frac{\mathrm{d}x}{\sqrt{P_3(x)}}$$

[r] One of the solutions of the polynomial

$$P_3(x) \propto (x - x_1)(x_2 - x)\left(1 - \frac{x}{x_3}\right)\left(1 + \frac{x_1 + x_2}{x_3}\right)^{-1}$$

is large with respect to the other two $x_3\ r_\mathrm{s}^{-1} \gg \delta = \frac{1}{2}(x_1 + x_2)$. In Newtonian mechanics, the polynomial is of second order. Using the approximation, one can ignore the last two parentheses. One then has

$$\begin{aligned}
\Delta\varphi &\sim (1 + r_\mathrm{s}\delta) \int_{r_\mathrm{min}^{-1}}^{r_\mathrm{max}^{-1}} \frac{\mathrm{d}x}{\sqrt{(x - x_1)(x_2 - x)}} \left(1 + \frac{x}{2x^3}\right) \\
&= (1 + r_\mathrm{s}\delta) \int \frac{\mathrm{d}x}{\sqrt{x_0^2 - \widetilde{x}^2}} \left[1 + \frac{r_\mathrm{s}}{2}(\widetilde{x} + \delta)\right] \\
&\sim \left[1 + \frac{3}{2} r_\mathrm{s}\delta\right]
\end{aligned}$$

at the second equality one can translate $x = \widetilde{x} + \delta$.

6.6 Charged black holes

To describe charged black holes one considers General Relativity and electromagnetism. The action is given by

$$S = S_\mathrm{EH} + S_\mathrm{Maxwell}\,, \quad S_\mathrm{Maxwell} = -\frac{1}{4e^2} \int \mathrm{d}^4x\, \sqrt{-g} F_{\mu\nu} F_{\rho\sigma} g^{\mu\rho} g^{\nu\sigma}$$

Where the interaction between the electromagnetic field and a charged particle is not included. One looks for the metric generated by a particle with mass M and electric charge e. Generating the electromagnetic field adds a term to the stress-energy tensor which influences the metric. The stress-energy tensor is

$$T_{\mu\nu} = -\frac{2\,\delta(\sqrt{-g}\mathcal{L})}{\sqrt{-g}\,\delta g^{\mu\nu}}$$

Therefore, the variation of the electromagnetic action is

$$\begin{aligned}
\delta S_\mathrm{Maxwell} &\propto \int \mathrm{d}^4x \left[-\frac{1}{2}\,\delta g^{\mu\nu}\, g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} + 2\sqrt{-g} F_{\mu\nu} F_{\rho\sigma}\, \delta g^{\mu\rho}\, g^{\nu\sigma}\right] \\
&= \int \mathrm{d}^4x\, \sqrt{-g}\, \delta g^{\mu\nu} \left[-\frac{1}{2} g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} + 2F_{\mu\rho} F_{\nu\rho} g^{\rho\sigma}\right] \\
&= \int \mathrm{d}^4x\, \sqrt{-g}\, \delta g^{\mu\nu} \left[-\frac{1}{2} g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} + 2F_{\mu\rho} F_\nu{}^\rho\right]
\end{aligned}$$

remembering that

$$\frac{\delta\sqrt{-g}}{\sqrt{-g}} = -\frac{1}{2}\,\delta g^{\mu\nu}\,g_{\mu\nu}$$

Therefore

$$T_{\mu\nu} = \frac{1}{e^2}\left[F_{\mu\rho}F_\nu{}^\rho - \frac{1}{4}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\right]$$

The trace-reversed of the equations of motion is

$$R_{\mu\nu} = 8\pi G\left[T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right]$$

in which

$$\begin{aligned} T = g^{\mu\nu}T_{\mu\nu} = T_{mu}{}^\mu &= \frac{1}{e^2}\left[g^{\mu\nu}F_{\mu\rho}F_\nu{}^\rho - \frac{1}{4}g^{\mu\nu}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\right] \\ &= \frac{1}{e^2}\left[F_{\mu\rho}F^{\mu\rho} - F_{\rho\sigma}F^{\rho\sigma}\right] = 0 \end{aligned}$$

Optional note: The stress-energy tensor is traceless and this relates to scale invariance. The coefficient of the action has no scale of length: electromagnetism is scale invariant (only classically). In flat space, the infinitesimal rescaling is

$$x'^\mu = Ax^\mu \implies \delta x^\mu = \varepsilon x^\mu\,, \quad A \approx 1+\varepsilon$$

Given a vector field v^μ, the current associated is $J^\mu = v^\nu T_\nu{}^\mu$. For $v^\mu = v^\mu$ one has

$$J^\mu = T^{\mu\nu}x_\nu \implies \partial_\mu J^\mu = \partial_\mu(T^{\mu\nu}x_\nu) = T^{\mu\nu}\eta_{\mu\nu} = T$$

where one remembers $\partial_\mu T^{\mu\nu} = 0$. End optional note.

The metric that realizes spherical symmetry is

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

[r] while keeping time translation as an isometry. This symmetry must be imposed on Maxwell's equations by using

$$(L_{l_i}F)_{\mu\nu} = 0$$

like one did for the metric

$$(L_{l_i}g)_{\mu\nu} = 0$$

This implies that the electromagnetic field is only radial

$$F_{tr} = f_e(r) \neq 0\,, \quad F_{t\theta} = F_{t\varphi} = 0$$

The last components have to be zero because they constitute a vector field on the sphere, but there is not field invariant under all three rotations. [r] The components

$$F_{r\theta} = F_{r\varphi} = 0$$

are zero by the previous argument. Finally

$$F_{\theta\varphi} \sim \sin\theta\, f_m(r)$$

which is related to the volume element in spherical coordinates. In terms of the magnetic field, [r] one sets $F_{\theta\varphi}$ to zero because one does not want magnetic charges, monopoles.

Lecture 18

lun 04 dic 2023 14:30

The action has to be varied with respect to both the vector potential and the metric. In Special Relativity, Maxwell's equations are

$$\partial_{[\mu}F_{\nu\rho]} = 0\,, \quad \partial_\mu F^{\mu\nu} = J^\nu = 0$$

The first equation is the same using the covariant derivative because one has chosen zero torsion. For the second one would like to have a covariant derivative. Though, varying the action with respect to the vector potential, one gets

$$\partial_\mu(\sqrt{-g}F^{\mu\nu}) = 0$$

which seems to be different from the equation above with a covariant derivative. [r] One has

$$0 = \nabla_\mu F^{\mu\nu} = \partial_\mu F^{\mu\nu}\Gamma^\mu_{\mu\sigma}F^{\sigma\nu} + \Gamma^\nu_{\mu\sigma}F^{\mu\sigma}\,, \quad \Gamma^\mu_{\mu\sigma} = \partial_\sigma \log\sqrt{-g} = \frac{1}{\sqrt{-g}}\,\partial_\sigma\sqrt{-g}$$

From which one has

$$0 = \partial_\mu F^{\mu\nu} + \frac{1}{\sqrt{-g}}\,\partial_\mu\sqrt{-g}\,F^{\mu\nu} + 0 = \frac{1}{\sqrt{-g}}\,\partial_\mu(\sqrt{-g}F^{\mu\nu})$$

This second version is more practical and one has

$$\sqrt{-g} = \mathrm{e}^{U+V}r^2\sin\theta$$

Using $\nu = 0$, one has

$$0 = \partial_1(\mathrm{e}^{U+V}r^2\sin\theta F^{10}) = \partial_1(\mathrm{e}^{U+V}r^2F^{10})\sin\theta \implies \partial_1(\mathrm{e}^{U+V}r^2F^{10}) = 0$$

For $\nu = 1, 2, 3$ one has $0 = 0$. The solution to Maxwell equations is

$$F^{10} = \frac{q}{r^2}\mathrm{e}^{-U-V}\,, \quad F_{10} = g_{11}F^{10}g_{00} = -\frac{q}{r^2}\mathrm{e}^{U+V}$$

One can compute Einstein's equations starting from the stress-energy tensor

$$F_{\mu\rho}F_\nu{}^\rho\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = \mathrm{d}t^2(F_{01}F_{01}g^{11}) + \mathrm{d}r^2(F_{10}F_{10}g^{00}) = \frac{q^2}{r^4}(\mathrm{e}^{2U}\,\mathrm{d}t^2 - \mathrm{e}^{2V}\,\mathrm{d}r^2)$$

also

$$\frac{1}{2}F_{\mu\nu}F^{\mu\nu} = F_{01}F_{01}g^{00}g^{11} = -\frac{q^2}{r^4}$$

Therefore

$$\begin{aligned} T_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu &= \left[F_{\mu\rho}F_\nu{}^\rho - \frac{1}{4}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\right]\mathrm{d}x^\mu\,\mathrm{d}x^\nu \\ &= \frac{q^2}{2r^4}[\mathrm{e}^{2U}\,\mathrm{d}t^2 - \mathrm{e}^{2V}\,\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)] \end{aligned}$$

[r] From which

$$T = g^{\mu\nu}T_{\mu\nu} = 0$$

The Ricci tensor is already known

$$R_{\mu\nu} = 8\pi G T_{\mu\nu}$$

[r] The first and second components of the metric are

$$tt + rr : \quad \frac{2}{r}(U' + V') = 0 \implies V = -U$$

Similarly

$$tt - rr : \quad U'' + 2U'\left(U' + \frac{1}{r}\right) = 8\pi\frac{Gq^2}{r^4}\mathrm{e}^{2(U-V)}$$

The colatitude component is

$$\theta\theta : \quad -\mathrm{e}^{2U}(2rU' + 1) + 1 = \frac{4\pi G}{e^2} F_{00}^2 r^2 = \frac{4\pi G q^2}{e^2 r^2} \implies \mathrm{e}^{2U} = 1 - \frac{2GM}{r} + G\frac{q^2}{r^2}$$

[r] e^2? This is the Reissner–Nordström solution

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

The potential energy is

$$V_{\mathrm{grav}} = m\sqrt{-g_{00}} = m\mathrm{e}^U \sim m\left(1 - \frac{GM}{r}\right)$$

where m is a test mass, not the mass M that created the metric. The potential comes down from infinity to zero, then nothing, then slowly rises from zero to a constant value [r] diagr. The radii at which the potential is null are

$$r = r_\pm = GM \pm \sqrt{(GM)^2 - Gq^2}$$

There are two notable spheres. In the middle region, r is the time coordinate, like for Schwarzschild. However, the inner region is similar to the outer-most region. This graph is valid only for the sub-extremal case

$$q^2 < GM^2$$

Otherwise for the extremal case $q^2 = GM^2$, there is only one point where the potential is null. For the super-extremal case $q^2 > GM^2$ there are no null points: there is no horizon.

In this last case, the origin is still a singularity but it is naked, not surrounded by an horizon. This answer is unsatisfactory because at the origin General Relativity may fail. One does not know what boundary conditions to impose at the origin. If someone considers this solution not to appear in nature, one is looking for the cosmic censorship hypothesis: a naked singularity can never be obtained by time evolution from physical initial conditions.

The extremal solution is important more as a theoretical device in supersymmetry.

One may then consider only the sub-extremal solution and study its geometry. [r] see notes penrose. In this case, near the singularity, t is the time coordinate, not r, so things have to be different from Schwarzschild's Penrose diagram. Inside the inner horizon, one again has the problem of a naked singularity because the singularity is a place in space, not time, it is not necessarily in the future. The strong cosmic censorship hypothesis states that timelike singularity also cannot form. Like one should not take the Schwarzschild solution to be real, as it describes an eternal black hole, one can argue that inside the inner horizon, the solution is not stable and it reorganizes itself.

One may want to continue study this solution anyway. One can analytically continue the solution to obtain an upside-down copy stitched above the current Penrose diagram. The part above is another universe.

6.7 Rotating black holes

For a simple theory like the previous there is not much one can do. One can drop the spherical symmetry and consider a rotation parameter. This holds for only one particle. Black hole solutions are very constrained.

One can assume the existence of two Killing vectors ∂_t and ∂_φ. There is a non-trivial dependence on r and θ. This is an optional discussion. The solution was found imposing a property that had no compelling reason to be true, but was found in other cases: one looks for spaces that are algebraically special. From the Riemann tensor, one can extract the Ricci tensor. This is something related to group theory: one can study the irreducible representation of the Riemann tensor space? [r]. One gets Weyl tensors which have curious properties. In particular, it has four indices and the same symmetries of the Riemann tensor and can be considered as a map from the space of anti-symmetrical two-indexed tensors to itself? [r]. Studying its eigenvalues one sees only four non trivial values. Sometimes they coincide and according to which do one obtains Petrov classification. At least two coincident eigenvalues implies the algebraically special property. Imposing this property on top of the expected symmetries means more constraints from which the solution arises.

Lecture 19

lun 11 dic 2023 14:30

There are many coordinate systems for the rotating solution. The one used in the course is the elliptic coordinates (for Schwarzschild the spherical coordinates have been used). At infinity one obtains the flat metric in elliptic coordinates. The transformation is given by

$$x = \sqrt{r^2 + a^2} \sin\theta \cos\phi \,, \quad y = \sqrt{r^2 + a^2} \sin\theta \sin\phi \,, \quad z = r\cos\phi$$

One can see that

$$\frac{x^2 + y^2}{r^2 + a^2} + \frac{z^2}{r^2} = 1$$

If $a = 0$ one gets back to spherical coordinates. For $a \neq 0$ one has ellipsoids for r constant. At $r = 0$ (and $\theta = \frac{\pi}{2}$) one has a disk of radius a. The metric is given by

$$\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \frac{\rho^2}{r^2 + a^2} \mathrm{d}r^2 + \rho^2 \, \mathrm{d}\theta^2 + (r^2 + a^2) \sin^2\theta \, \mathrm{d}\phi^2 \,, \quad \rho^2 = r^2 + a^2 \cos^2\theta$$

In this coordinate system, the metric for the rotating black hole is simpler and it is called Kerr solution. The original solution was found without charged. The solution is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \frac{2mr}{\rho^2} (\mathrm{d}t - a\sin^2\theta \, \mathrm{d}\phi)^2 + (r^2 + a^2) \sin^2\theta \, \mathrm{d}\phi^2 + \rho^2 \left(\frac{\mathrm{d}r^2}{\Delta} + \mathrm{d}\theta^2 \right)$$

where

$$m = GM \,, \quad \Delta = r^2 - 2mr + a^2$$

The third addendum is the same as the ϕ component in flat space, also $\rho^2 \, \mathrm{d}\theta^2$ appears in both. Also, for very large radii, one recovers flat space. The time component is Schwarzschild's for $a = 0$:

$$\mathrm{d}s^2 = \left(-1 + \frac{2m}{r} \right) \mathrm{d}t^2 + \frac{\mathrm{d}r^2}{1 - \frac{2m}{r}} + r^2 (\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2)$$

Ergosphere. The time component is

$$g_{00} = -1 + \frac{2mr}{\rho^2}$$

One can study the sign. Starting from outside, one has

$$g_{00} < 0 \implies \rho^2 - 2mr > 0 \implies r > m + \sqrt{m^2 - a^2 \cos^2\theta} \equiv r_0(\theta)$$

At $r_0(\theta)$ the component changes sign so t ceases to be the time coordinate, while r becomes such.

The radial component is

$$g_{11} = \frac{\rho^2}{\Delta} > 0 \implies \Delta > 0 \implies r > r_+ = m + \sqrt{m^2 - a^2}$$

This locus does not depend on θ. It coincides with $r_0(\theta)$ when $\theta = 0, \pi$. Outside the black hole, $r > r_0$ one has

$$g_{00} < 0 \,, \quad g_{11} > 0$$

In the region $r_+ < r < r_0$, the ergosphere, one has

$$g_{00} > 0 \,, \quad g_{11} > 0$$

It seems like there is no time coordinate. However, the metric is not diagonal, so one has to look at the eigenvalues. The radial component is an eigenvalue, but the time component is not. In fact

$$g_{\mu\nu} = \begin{bmatrix} g_{00} & 0 & 0 & g_{03} \\ 0 & g_{11} & 0 & 0 \\ 0 & 0 & g_{22} & 0 \\ g_{30} & 0 & 0 & g_{33} \end{bmatrix}$$

There is an inner block for which one can see that g_{11} and g_{22} are eigenvalues, but the outer block needs diagonalization. However, one can compute the determinant to know if there is one negative eigenvalue. Since the metric is symmetric, the coefficient is half the one it appears in $\mathrm{d}s^2$ above:

$$\begin{aligned}\det\begin{bmatrix} g_{00} & g_{03} \\ g_{30} & g_{33}\end{bmatrix} &= \det\begin{bmatrix} -1+\frac{2mr}{\rho^2} & -\frac{2mr}{\rho^2}a\sin^2\theta \\ -\frac{2mr}{\rho^2}a\sin^2\theta & \frac{2mr}{\rho^2}a^2\sin^4\theta + (r^2+a^2)\sin^2\theta\end{bmatrix} \\ &= \left(-1+\frac{2mr}{\rho}^2\right)(r^2+a^2)\sin^2\theta - \frac{2mr}{\rho^2}a^2\sin^4\theta \\ &= -(r^2+a^2)\sin^2\theta + \frac{2mr}{\rho^2}(r^2+a^2)\sin^2\theta - a^2\sin^4\theta \\ &= -(r^2+a^2)\sin^2\theta + \frac{2mr}{\rho^2}\sin^2\theta(r^2+a^2\cos^2\theta) \\ &= -(r^2+a^2)\sin^2\theta + \frac{2mr}{\rho^2}\rho^2\sin^2\theta \\ &= -\Delta\sin^2\theta < 0\end{aligned}$$

One of the eigenvalues is the time direction: time is a mix of the coordinates t and ϕ. Since r is not a time coordinate $g_{11} > 0$, then one is not obligated to go towards the center: r_0 is not an horizon. Since $g_{00} > 0$ then one cannot be at rest because the four-velocity would have positive norm (remember that the signature is spacelike). Since the time coordinate is a mix, one can rotate around the black hole. The light cone is directed along ϕ.

Standing in the ergosphere, one can escape it by throwing something. [r] It would possibile to extract energy from the black hole.

Inside the ergosphere. There is a second solution r_- to $\Delta = 0$. For $r_- < r < r_+$ the radial component is $g_{11} < 0$ so r is the time component. The light cones point in the radial direction.

Inside this solution $r < r_-$ then $g_{11} > 0$. The singularity is at the boundary of the $r = 0$ disk. In this region one can interact with the singularity, so the doubt of the solution being stable arises again.

Angular momentum. The angular momentum cannot be found by taking an integral of the stress-energy tensor. [r] There is no simple notion of gravitational notion and no simple integral over a single region to get [r]. In the Schwarzschild solution there is no point where is the mass because the singularity is in the future. The mass can be justified by taking integrals at infinity. Even if there is a general theory, the gist of it is similar to what one has seen for the Schwarzschild solution by using the gravitational potential. The discussion can be done for angular momentum also. The angular momentum is non-zero. As such the solution is called rotating. The light rays rotate too. Looking at the horizon r_+, there are light rays that stay on the horizon and rotate.

Consequences. If the solution is all physical, then going inside the singularity's ring, one does not stay within a single solution, similar to branch cuts of complex multivalued functions. Crossing the disk one finds a copy of the solution with $r < 0$ with closed timelike curves: time repeats itself. One must remember though that this region, $r < r_-$, may be unstable and unphysical.

Remarks. This solution describes black holes in nature since every astrophysical object rotates. This solution also presents an extremal case of maximal rotating black hole, but the one observed are all in the sub-extremal case.

One can also include charge, acceleration and other exotic things. When looking for a solution of gravity alone, this is the most general static solution. Black holes are stable outside: the perturbations vanish exponentially in time.

Coupling General Relativity with some forms of matter, then the solution depends only on a few parameters like m, q, a. Black holes are very different from stars. It is often said that black holes have no hair. Once there is an horizon, the solution is determined.

This exponential decay can be seen in gravitational waves: a few moment after the merging, the result black hole is a different shape but quickly decays to the Kerr solution and this is reflected in the shape of the gravitational waves.

The process of extracting energy from the black holes prompted thinking about the thermodynamics of a black hole, one finds the first law and the second along with entropy. One would think that entropy measures the many microstates that form a macrostate. However in General Relativity there are no multiple microstates, there is no statistical interpretation of a black hole entropy. The problem is resolved in quantum gravity and string theory because they reproduce correctly the entropy.

Lecture 20

gio 14 dic 2023 14:30

7 Time evolution

One may study time evolution through the Hamiltonian formalism in which one picks the notion of time. In this way one loses some of the invariance of the coordinate changes of the theory. This problem is shared with other relativistic theory. In Cosmology there is a natural choice of coordinates and time because the Universe is not empty.

7.1 Constraints

Picking a time coordinate t, the corresponding entry in the metric is negative $g_{00} < 0$. Without (further) loss of generality, one may decompose the metric in blocks

$$g_{\mu\nu} = \begin{bmatrix} g_{00} & g_{0i} \\ g_{i0} & g_{ij} \end{bmatrix}$$

It is useful to use a different parametrization, written in terms of the line element

$$\mathrm{d}s^2 = -N^2\,\mathrm{d}t^2 + \gamma_{ij}(\mathrm{d}x^i + N^i\,\mathrm{d}t)(\mathrm{d}x^j + N^j\,\mathrm{d}t) \quad (= g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu)$$

where $\gamma_{ij} = g_{ij}$ and N, N^i are the lapse and shift variable.

There is a notable one-form one can introduce

$$n = n_\mu\,\mathrm{d}x^\mu = N\,\mathrm{d}t\,, \quad n^2 = g_{\mu\nu}n^\mu n^\nu = g^{\mu\nu}n_\mu n_\nu = -1$$

which leads to the vector field

$$n = n^\mu\,\partial_\mu = -\frac{1}{N}(\partial_t - N^i\,\partial_i)$$

One introduces the projector along the space directions

$$P_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu$$

This is called the first fundamental form. It is complementary to n:

$$P_{\mu\nu}n^\nu = n_\mu + n_\mu n_\nu n^\nu = n_\mu(1-1) = 0$$

A projector is idempotent:

$$P^\mu{}_\nu P^\nu{}_\rho = (\delta^\mu{}_\nu + n^\mu n_\nu)(\delta^\nu{}_\rho + n^\nu n_\rho) = \delta^\mu{}_\rho + 2n^\mu n_\rho + n^\mu n_\nu n^\nu n_\rho = \delta^\mu{}_\rho + n^\mu n_\rho = P^\mu{}_\rho$$

Its eigenvalues are either 1 or 0 for the coordinates that are projected and the ones that vanish. This projector is similar to the following projector of Quantum Mechanics

$$1 - |\psi\rangle\langle\psi|$$

which removes $|\psi\rangle$. [r]

The second fundamental form is

$$K_{\mu\nu} = \frac{1}{2}L_n P_{\mu\nu}$$

The Lie derivative with respect to n can be thought of as a time derivative. Utilizing the parametrization above, one can explicitly compute the Lie derivative

$$K_{ij} = \frac{1}{N}\left[-\frac{1}{2}\dot{\gamma}_{ij} + \nabla^3_{(i}N_{j)}\right]$$

where ∇^3 is the covariant derivative with respect to γ_{ij}. One can see the time derivative of the space metric γ_{ij}. [r]

One may now study the dynamics. The hamiltonian formalism is useful for the quantization of a theory. In quantum field theory it works well, but not as well in General Relativity. One has already used the Lagrangian

$$S_{\mathrm{EH}} \propto \int \mathrm{d}^4x\,\sqrt{-g}R = \int \mathrm{d}t\,\mathrm{d}^3x\,N\sqrt{\gamma}(R_3 - K^{ij}K_{ij} - K^2)\,, \quad K = \gamma^{ij}K_{ij}$$

where R_3 is the Ricci scalar for γ_{ij}. This expression is too simple. The main variable in General Relativity is the metric which has been divided into blocks. The metric γ_{ij} appears with its derivatives K_{ij} (similar to kinetic terms), while the Ricci scalar contains only the spatial derivatives, so it is more of a potential. However, [r] N does not appear with derivatives and N_i has only spatial derivatives: they have no spatial derivatives. As such they are Lagrangian multipliers because the equations of motion are just constraints.

Varying the action with respect to N, one gets

$$R_3 = K^{ij}K_{ij} - K^2$$

This is a first-order equation [r]. One has seen a first-order equation with Schwarzschild. The equations of motion for N_i are

$$\nabla^3_i K^{ij} = (\nabla^3)^j K$$

One may compute the Hamiltonian. The canonical momenta are

$$\pi = \delta_{\dot{N}}\mathcal{L}\,, \quad \pi_i = \delta_{\dot{N}^i}\mathcal{L}\,, \quad \pi_{ij} = \delta_{\dot{\gamma}^{ij}}\mathcal{L} = -\sqrt{\gamma}(K_{ij} - \gamma_{ij}K)$$

One has already seen that

$$\pi = \pi_i = 0$$

The Hamiltonian is then

$$H = \int \mathrm{d}^3x\,(\pi\dot{N} + \pi^i\dot{N}_i + \pi^{ij}\dot{\gamma}_{ij} - L) = \int \mathrm{d}^3x\,[\pi\dot{N} + \pi^i\dot{N}_i + N\sqrt{\gamma}(K^{ij}K_{ij} - K^2 - R_3) + N_i\nabla^3_j\pi^{ij}]$$

Computing the time derivative of the constraints, one gets secondary constraints

$$0 = \dot{\pi} = \{H, \pi\} = \delta_N H = \sqrt{\gamma}(K^{ij}K_{ij} - K^2 - R_3)$$

from which follows the equations of motions of N. Similarly

$$0 = \dot{\pi}^j = \nabla^3_i\pi^{ij}$$

gives the equations of motion for N_i. From these it can be seen why they are constraints.

[r] In General Relativity, the Hamiltonian is zero on-shell. This is related to the fact that it is difficult to define a notion of potential energy.

[r] One of the first attempts of quantizing General Relativity was the one of Wheeler [r].

The first and second fundamental forms can be found geometrically through the Gauss–Codazzi equations [r].

7.2 Cosmology

The following discussion involves no matter [r]. At very large scales, the universe presents many symmetries

- homogeneity,

- isotropy.

The distribution of galaxies is even, the matter is homogeneous. No direction is preferred over any other, space is isotropic.

One has learnt how to express symmetries in General Relativity. Homogeneity is linked to spatial translations and isotropy is related to spatial rotations: there are a total of six Killing vectors. Symmetries are manifested through the existence of Killing vectors of space (not time). At very large scales, space is flat. Three-dimensional spaces with six Killing vectors are very constrained.

In general, in d dimensions, space can have at most $\frac{d(d+1)}{2}$ Killing vectors. This is the same number of Killing vectors as in $\mathbb{R}^d$: d translations and $\frac{1}{2}d(d-1)$ rotations (which is the number of possible two-dimensional planes, or one can think about the entries of an anti-symmetrical matrix? [r]). When a space has all its Killing vectors it is called maximally symmetric.

Maximal symmetry is not a property of flat space only, but also spherical and hyperbolic space. Maximally symmetric spaces are classified as

- flat space $\mathbb{R}^3$,
- spherical space
$$S^3 \subset \mathbb{R}^4\,, \quad \mathrm{d}s^2 = \mathrm{d}\psi^2 + \sin^2\psi\,\mathrm{d}s^2_{S^2}$$
where one has
$$x_1 = r\sin\theta\cos\phi\sin\psi\,, \quad x_2 = r\sin\theta\sin\phi\sin\psi\,, \quad x^3 = r\cos\theta\sin\psi\,, \quad x^4 = r\cos\psi$$
The rotations in $\mathbb{R}^4$ are six. Even though there are other symmetries (like translations), these leave the sphere invariant. So the 3-sphere has six Killing vectors.
- Hyperbolic space H^3
$$\mathrm{d}s^2_{H^3} = \mathrm{d}\psi^2 + \sinh^2\psi\,\mathrm{d}s^2_{S^2}$$
Hyperbolic space can be realized as a subspace of Minkowski space $\mathbb{R}^{1,3}$:
$$-x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$$
There are three spatial rotations and three boosts, the Lorentz generators leave the hyperboloid invariant.
- Discrete quotients of the spaces above. For $\mathbb{R}^3$ one may have the torus T^3, the torus with a line $T^2 \times \mathbb{R}$, or $\mathbb{R}^2 \times S^1$ the plane times a circle, etc. The sphere is constrained. Hyperbolic space can be made compact (like a torus with more holes).

All maximally symmetric spaces have the property
$$R_{ijkl} = \alpha(g_{ik}g_{jl} - g_{il}g_{jk})$$
where α is a parameter. In this way one can do the computation for all the spaces above. The parameter α can be changed by rescaling [r] the space and the metric, and made into
$$\alpha = k = \begin{cases} 1\,, & S^3 \\ 0\,, & \mathbb{R}^3 \\ -1\,, & H^3 \end{cases}$$

Lecture 21

lun 18 dic 2023 14:30

The metric γ_{ij} is maximally symmetric. Because of isotropy $N^i = 0$, also because of homogeneity $N^2(t)$ is a function of time only. Reparametrizing time, one has
$$t \to t' = t'(t)\,, \quad \mathrm{d}t' = \mathrm{d}_t t'\,\mathrm{d}t$$

One can choose t' to get any arbitrary function, in particular $\mathrm{d}_t t' = \frac{1}{N}$ [r] so N can be set to 1. Maximally symmetric space are very constrained and there are only the three types above. Considering a rescale a of the space, one has

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\,\mathrm{d}s_3^2\,, \quad \mathrm{d}s_3^2 = g_{ij}^3\,\mathrm{d}x^i\,\mathrm{d}x^j$$

This is the line element of the Friedmann–Lemaître–Robinson–Walker (FLRW) metric

$$g_{00} = -1\,, \quad g_{ij} = a^2 g_{ij}^3\,, \quad g_{0i} = 0$$

The Christoffel symbols are

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

Therefore

$$\Gamma^0_{00} = \frac{1}{2} g^{00}(\partial_0 g_{00}) = 0\,, \quad \Gamma^0_{0i} = \frac{1}{2} g^{00}(\partial_0 g_{0i} + \partial_i g_{00} - \partial_0 g_{0i}) = 0\,, \quad \Gamma^i_{00} = 0$$

also

$$\Gamma^0_{ij} = \frac{1}{2} g^{00}(\partial_i g_{j0} + \partial_j g_{i0} - \partial_0 g_{ij}) = \frac{1}{2}\partial_0(a^2 g_{ij}^3) = a\dot{a} g_{ij}^3$$

[r] where negative sign? similarly

$$\Gamma^i_{j0} = \frac{1}{2} g^{ik}(\partial_j g_{0k} + \partial_0 g_{jk} - \partial_k g_{j0}) = \frac{1}{2} a^{-2} g_3^{ik}\,\partial_0(a^2 g_{jk}^3) = \frac{\dot{a}}{a}\delta^i{}_j$$

likewise

$$\Gamma^i_{jk} = (\Gamma_3)^i_{jk}$$

The Ricci tensor is

$$R_{00} = -3\frac{\ddot{a}}{a}\,, \quad R_{ij} = R_{ij}^3 + (2\dot{a}^2 + a\ddot{a}) g_{ij}^3$$

where one has

$$R_{ijkl}^3 = k(g_{ik}^3 g_{jl}^3 - g_{il}^3 g_{jk}^3) \implies R_{jl}^3 = g_3^{ik} R_{ijkl}^3 = k(3g_{jl}^3 - g_{jl}^3) = 2k g_{jl}^3$$

If the universe were empty, the calculation would be done and the equation above would be zero.

To find the stress-energy tensor, one can impose onto it the same symmetries of the universe as above:

$$L_k T_{\mu\nu} = 0 \implies T_{0i} = 0\,, \quad T_{ij} = p a^2 g_{ij}^3 = p\gamma_{ij}\,, \quad T_{00} \equiv \rho$$

The Lie derivative with respect to the six Killing vectors should be null. The first implication comes from isotropy. The second implication is given by the fact that the only tensor that respects all the symmetries is just the metric. Where ρ is an energy density and p is a pressure density. [r] One may summarize these three statements into a single equation

$$T_{\mu\nu} = \rho n_\mu n_\nu + p P_{\mu\nu}$$

whose trace may be calculated in two ways

$$T = g^{00} T_{00} + g^{ij} T_{ij} = -\rho + 3p\,, \quad T = g^{\mu\nu} T_{\mu\nu} = -\rho - p + 4p = -\rho + 3p$$

in the second one uses $P_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu$.

One can solve the Einstein's field equations

$$R_{\mu\nu} = 8\pi G\left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T\right)$$

The time-time component is

$$-3\frac{\ddot{a}}{a} = 8\pi G\left[\rho - \frac{1}{2}(-1)(-\rho + 3p)\right] = 4\pi G(\rho + 3p) \implies \frac{\ddot{a}}{a} = -\frac{4}{3}\pi G(\rho + 3p)$$

The spatial components are

$$
\begin{aligned}
(2k + 2\dot{a}^2 + a\ddot{a})g^3_{ij} &= 8\pi G\left[pa^2 g^3_{ij} - \frac{1}{2}a^2 g^3_{ij}(-\rho + 3p)\right] \\
2k + 2\dot{a}^2 + a\ddot{a} &= 8\pi G a^2 \frac{\rho - p}{2} \\
\frac{k + \dot{a}^2}{a^2} + \frac{\ddot{a}}{2a} &= 4\pi G(\rho - p) \\
\frac{k + \dot{a}^2}{a^2} &= \frac{8}{3}\pi G\rho
\end{aligned}
$$

The second derivative is redundant. In fact, considering the following consequence of the Einstein's equations

$$
0 = \nabla_\mu T^{\mu\nu} = \partial_\nu T^{\mu\nu} + \Gamma^\mu_{\mu\rho} T^{\rho\nu} + \Gamma^\nu_{\mu\rho} T^{\mu\rho}
$$

For $\nu = 0$, one has

$$
0 = \partial_0 T^{00} + \Gamma^\mu_{\mu 0} T^{00} + \Gamma^0_{ij} = \partial_0 \rho + \Gamma^\mu_{\mu 0}\rho + \Gamma^0_{ij} = \dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p)
$$

where

$$
\Gamma^\mu_{\mu\nu} = \partial_\nu \log\sqrt{-g}\,, \quad \sqrt{-g} = a^3\sqrt{g_3}
$$

The two components above and this last equation are called Friedmann equations. The first equation can be obtain from the other two which are first-order.

There are three parameters in the equations. For ordinary matter, one can approximate the pressure $p = 0$: galaxies are mostly empty and matter does not collide. For massless particles, for radiation, the pressure is not zero. For electromagnetism one has $T = 0$ due to scale invariance and therefore $p = \frac{1}{3}\rho$. One can unite the two cases by introducing a parameter w:

$$
p = w\rho\,, \quad w = 0, \frac{1}{3}
$$

The Friedmann equations become

$$
\dot{\rho} + 3\frac{\dot{a}}{a}(1 + w)\rho = 0 \implies 0 = \frac{\dot{\rho}}{\rho} + 3(1 + w)\frac{\dot{a}}{a} = \partial_0[\log\rho + 3(1 + w)\log a) = \partial_0 \log[\rho a^{3(1+w)}]
$$

therefore

$$
\rho = \rho_0 a^{-3(1+w)}
$$

For matter one has

$$
\rho_{\mathrm{m}} = \rho_{\mathrm{m},0} a^{-3}
$$

The density scales like the volume. For radiation one has

$$
\rho_{\mathrm{r}} = \rho_{\mathrm{r},0} a^{-4}
$$

[r] The total density is

$$
\rho = \rho_{\mathrm{m}} + \rho_{\mathrm{r}} = \rho_{\mathrm{m},0} a^{-3} + \rho_{\mathrm{r},0} a^{-4}
$$

From the equation of motions for the spatial components one can solve for a. This parameter is not a constant. It has been observed that distant objects travel away from each other: the universe is expanding. [r] The Hubble constant is

$$
H_0 = \left(\frac{\dot{a}}{a}\right)_{t=0}
$$

The parameter a can be fitted to experimental points but the fit fails and one finds that the equations of motion are wrong. Going back [r] the Einstein-Hilbert action can be written as

$$
S_{\mathrm{EH}} = \frac{1}{16\pi G}\int \mathrm{d}^4x\,\sqrt{-g}(R - 2\Lambda)
$$

where Λ is the cosmological constant. The field equations become

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \implies R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G (T_{\mu\nu} + T^{\Lambda}_{\mu\nu})$$

[r] where one has

$$T^{\Lambda}_{\mu\nu} = \frac{\Lambda}{8\pi G} g_{\mu\nu} \implies \rho + p = 0 \implies w = -1 \implies \rho_\Lambda = \text{const.}$$

From which

$$\rho = \rho_{\mathrm{m},0} a^{-3} + \rho_{\mathrm{r},0} a^{-4} + \rho_\Lambda$$

This time the fit works. [r] If $\rho_0 \equiv \rho$ today, then

$$\frac{\rho_\Lambda}{\rho_0} \approx 0.7\,, \quad \frac{\rho_\mathrm{m}}{\rho_0} \approx 0.3\,, \quad \frac{\rho_\mathrm{r}}{\rho_0} \approx 10^{-5}$$

At a certain time in the past, the Big Bang at around 13.8 billion years ago, one had $a = 0$. However, at this point one shall not trust the equations since $a = 0$ because one has a singularity. Penrose and Hawking proved that the singularity is not an artifact of this simplified case, but is inevitable with just very mild assumptions. At the beginning, the cosmological constant was very large: inflation. Also the fit predicts an accelerated expansion. [r]

The cosmological density remains constant while the other two densities dilute with time.

Lecture 22

gio 21 dic 2023 14:30

Excursus — cosmological constant. In quantum field theory, one uses a theory that works at any energy scale and derive simplified models, effective theories, that work at particular energies. There is an expectation that in such a theory one should find all allowed terms in a lagrangian. For General Relativity, the only relevant terms at low energies are the ones above, greater powers of the Ricci scalar are irrelevant (apart from R^2). The cosmological constant is then R^0.

In most quantum theories, energy is quantized and the ground state does not have zero energy. In quantum field theories, the zero-point energy is not zero, every particle has non-zero energy. It is energy differences that matter. In General Relativity, the zero-point energy is relevant because it gives rise to the cosmological constant. Trying to evaluate the zero-point energy of the vacuum may give the natural constant. The energy is infinite because there are infinitely many creation operators, one for each momentum. One may argue that quantum field theories are not the correct descriptions at high energies, so one introduces a cutoff. This cutoff may be at the scale where quantum gravity applies. The mass scale one can build out of the constants of the theory is the Planck mass

$$m_\mathrm{p} \sim 10^{19}\,\mathrm{eV}$$

[r] The sum of zero-point energy of the vacuum up to the cutoff is m_p^4 [r]. The observations reveal $(1\,\mathrm{meV})^4$ [r]. This mismatch is the cosmological constant problem.

One idea that if the cosmological were much larger than what it is, the accelerated expansion would have occurred earlier and faster than it did. With a few orders of magnitude more, no structures would have been possible and no one would have been able to observe this universe: this is the anthropic principle. This idea is controversial because if the principle applies to habitable planets, it does not really apply to universes, because there is more than one of the former, but perhaps only one of the latter.

There is also the idea that the expansion is different outside the observable universe. In some models, when inflation ended, there was a phase transition bubble where the cosmological constant is lower than outside. These ideas cannot be observed, they are non-verifiable.

Other ideas is that the cutoff is lower than Planck's mass, but the problems is not solved.

Shape of the universe. Observations revealed that

$$\frac{k}{a^2} \ll 1$$

but one does not know if $k = 0$ or a^2 is very large. The value of the constant k does not matter for the evolution of the universe. [r]

Effect of the cosmological constant. One can study the effect of the cosmological constant setting to zero all other contributions to the density ρ. The second Friedmann equation becomes

$$\frac{k+\dot{a}^2}{a^2} = \frac{8\pi G}{3}\rho_\Lambda = \frac{\Lambda}{3}$$

For $k=0$, the solution is

$$\frac{\dot{a}^2}{a^2} = \frac{\Lambda}{3}\,, \quad a = a_0 \exp\left[\sqrt{\frac{\Lambda}{3}}t\right] = a_0 \mathrm{e}^{\frac{t}{L}}\,, \quad L = \sqrt{\frac{3}{\Lambda}}$$

where L is the typical length scale of the geometry. From this one sees that the expansion is exponential. For $k=1$ the equation is

$$1+\dot{a}^2 = \frac{\Lambda}{3}a^2$$

Letting $\Lambda = 3$, one has

$$a^2 - \dot{a}^2 = 1 \implies a = a_0 \cosh\frac{t}{2}$$

Fro $k=-1$ one has

$$-1+\dot{a}^2 = \frac{\Lambda}{3}a^2 \implies a = a_0 \sinh\frac{t}{2}$$

For a distant future, all three have an exponential behaviour and are all the same solution of the Einstein's field equations: they lead to de Sitter space-time.

7.3 de Sitter spaces

A hyperboloid can be embedded in a higher dimensional Minkowski space

$$\{-X_0^2 + X_1^2 + X_2^2 + X_3^2 + X_4^2 = L^2\} \subset \mathcal{M}^5$$

One can parametrize hyperbolic space using hyperbolic functions

$$X_0 = L\sinh\frac{t}{L}\,, \quad X_i = L\cosh\frac{t}{L}\,\hat{X}_i\,, \quad i = 1,2,3,4\,, \quad \sum_i \hat{X}_i^2 = 1$$

The last sum imply a three-sphere S^3. The metric is

$$\mathrm{d}s^2 = -\mathrm{d}X_0^2 + \mathrm{d}X_1^2 + \mathrm{d}X_2^2 + \mathrm{d}X_3^2 + \mathrm{d}X_4^2 = -\mathrm{d}t^2 + \cosh^2\frac{t}{L}\,\mathrm{d}s^2_{S^3}$$

The hyperboloid is made of slices that are three-spheres. The metric is FLRW with $k=1$ and $a = \cosh\frac{t}{L}$.

The hyperboloid can be parametrized in another manner

$$x_0^2 - x_1^2 - x_2^2 - x_3^2 = L^2\sinh^2\frac{t}{L}\,, \quad x_4 = L\cosh\frac{t}{L}$$

The metric is then

$$\mathrm{d}s^2_{\mathrm{dS}} = -\mathrm{d}X_0^2 + \mathrm{d}X_1^2 + \mathrm{d}X_2^2 + \mathrm{d}X_3^2 + \mathrm{d}X_4^2 = -\mathrm{d}t^2 + \sinh^2\frac{t}{L}\,\mathrm{d}s^2_{H^3}$$

which is FLRW with $k=-1$ and $a = \sinh\frac{t}{L}$.

Another parametrization is

$$X_0 \pm X_4 = Lu_\pm\,, \quad X_i = Lu_+ x_i \implies \mathrm{d}s^2_{\mathrm{dS}} = -\mathrm{d}t^2 + \mathrm{e}^{\frac{2t}{L}}\,\mathrm{d}s^2_{\mathbb{R}^3}$$

which is FLRW with $k=0$ and $a = \mathrm{e}^{\frac{t}{L}}$.

All the solutions are the same de Sitter space.

Symmetries. de Sitter space have many properties and symmetries, more than the six used for space-time: the symmetries of $\mathbb{R}^5$ that leave invariant the hyperboloid are the Lorentz transformation of $\mathbb{R}^5$. By construction, it follows that a de Sitter space has the symmetries of Minkowski space $\mathrm{SO}(1,4)$. The Killing vectors are six rotations and four boosts for a total of ten vectors. This is the same as Minkowski $\mathcal{M}^4$ [r]. Since de Sitter space has the same number of Killing vectors as flat space, then it is maximally symmetric.

The Riemann tensor is

$$R_{\mu\nu\rho\sigma} = \frac{\Lambda}{3}(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho})$$

One can solve Einstein equations starting from [r].

Penrose diagrams. One wants to map the range of coordinates into a compact set to be able to use a Penrose diagram. One introduces a new time coordinate η

$$\cosh\frac{t}{L} = \frac{1}{\cos\eta}$$

The line element is then

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \cosh^2\frac{t}{L}\,\mathrm{d}s^2_{S^3} = \frac{1}{\cos^2\eta}(-\mathrm{d}\eta^2 + \mathrm{d}s^2_{S^3}) = \frac{1}{\cos^2\eta}(-\mathrm{d}\eta^2 + \mathrm{d}\rho^2 + \sin^2\rho\,\mathrm{d}s^2_{S^2})$$

In the space $\rho\eta$, the Penrose diagram is a square. The lateral sides are not infinity but the poles of the horizontal slices which are three-spheres. The upper side is the infinite future and the lower side is the infinite past.

The parametrization is

$$X_0 = L\sin\psi\sinh T\,, \quad X_4 = L\sin\psi\cosh T\,,$$

with

$$X_1^2 + X_2^2 + X_3^2 = L^2\cos^2\psi\,, \quad -X_0^2 + X_4^2 = L^2\sin^2\psi\,, \quad X_i = L\cos\psi\,x_i\,, \quad \sum_i x_i^2 = 1$$

Then the metric is

$$\mathrm{d}s^2_{\mathrm{dS}} = L^2(\mathrm{d}\psi^2 - \sin^2\psi\,\mathrm{d}T^2 + \cos^2\psi\,\mathrm{d}s^2_{S^2})$$

which is static and ∂_T is a Killing vector. Setting $t = LT$ and $r = L\cos\psi$ one has

$$\mathrm{d}s^2_{\mathrm{dS}} = -\left[1 - \frac{r^2}{L^2}\right]\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{1 - \frac{r^2}{L^2}} + r^2\,\mathrm{d}s^2_{S^2}$$

This is similar to Schwarzschild's. One can study same properties. The gravitational potential is

$$V_{\mathrm{grav}} = m\sqrt{-g_{tt}} = m\sqrt{1 - \frac{r^2}{L^2}}$$

It is flat for $r = 0$ and bends to zero at $r = L$. This time, the regular space is inside the cosmological horizon: beyond a certain distance one can no longer see beyond. A particle in this potential goes to increasing radii: this is the expansion of the universe, the potential is driving everything away from everything else. This space is maximally symmetric, every point is equivalent to any other: there is no common center. Every observer has an horizon, but it depends on the observer itself, not on the geometry: this is similar to the apparent horizon of an accelerated observer.

This coordinate system only covers part of the de Sitter space, the left triangle under the main diagonal, over the anti-diagonal. The only coordinate system that covers everything is the one of FLRW with $k = 1$. The two diagonals are the cosmological horizon: a static object inside the former region will eventually cross the horizon.

Lecture 23

lun 08 gen 2024 14:30

8 Gravitational waves

In the beginning, gravitational waves were thought to be just an artifact of gauge invariance.

Electromagnetism. One may start recalling the propagating degrees of freedom of electromagnetism in flat space. The equations of motion are

$$F_{\mu\nu} = 2\,\partial_{[\mu}A_{\nu]} \implies \partial_\mu F^{\mu\nu} = 0$$

and

$$\partial_{[\mu}F_{\nu\rho]} = 0 \implies 0 = \partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = \Box\, A_\nu - \partial_\nu \partial^\mu A_\mu$$

This last equation can be simplified by fixing a gauge. An infinitesimal gauge transformation is a linear transformation given by

$$\delta A_\mu = \partial_\mu \lambda$$

One possible choice is the Lorenz gauge $\partial_\mu A^\mu = 0$ for which one needs to find λ such that

$$\delta \partial^\mu A_\mu = \Box\, \lambda \implies \Box\, \lambda = -\partial^\mu A_\mu$$

Its solution involves a Fourier transform

$$-k^2\lambda(k) = \mathrm{i}k^\mu A_\mu(k) \implies \lambda(k) = -\frac{\mathrm{i}}{k^2}k^\mu A_\mu(k)$$

There are other subtleties, but this is the idea. Transforming back to position space, one gets the solution through Green's function.

Therefore the second equation of motion for electromagnetism is

$$\Box\, A_\nu = 0$$

The solution is not unique because one can add an arbitrary function that has zero d'Alembertian. The residual gauge invariance is such that

$$\Box\, \lambda = 0$$

In momentum space, the equation of motion above is

$$k^2 A_\nu = 0$$

but the gauge transformation is

$$\delta A_\nu = \mathrm{i}k_\nu \lambda(k)$$

Therefore, the equation of motion has support only on the mass shell where $k^2 = 0$. [r] the longitudinal component of A_ν along k_ν can be gauged away: there are two independent polarizations. The two removed are due to

$$k^\mu A_\mu = 0\,, \quad k^2 A_\nu = 0$$

Gravity. Unlike electromagnetism, gravity is a non-linear theory and one uses the weak-field approximation (already seen when deriving the equations of motion)

$$g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}\,, \quad h_{\mu\nu} \ll 1$$

[r] As previously seen, in vacuum, the equations of motion are

$$R_{\mu\nu} = 0 \implies -\frac{1}{2}\Box\, h_{\mu\nu} + \partial_{(\mu}\partial^\rho h_{\nu)\rho} - \frac{1}{2}\partial_\mu\partial_\nu h = 0\,, \quad h \equiv \eta^{\mu\nu}h_{\mu\nu}$$

The metric h is thought of as a field and the metric is η, so the d'Alembertian is

$$\Box = \eta^{\mu\nu}\,\partial_\mu\partial_\nu\,, \quad \partial^\rho = \eta^{\mu\rho}\partial_\rho$$

One truncates the expansion at first order. Gravity is a gauge theory and the gauge transformations are the coordinate changes (which are diffeomorphisms). Under an infinitesimal coordinate change, the metric transforms as

$$\delta g_{\mu\nu} = (L_\xi g)_{\mu\nu} = \xi^\rho\,\partial_\rho g_{\mu\nu} + \partial_\mu \xi^\rho\, g_{\rho\nu} + \partial_\nu \xi^\rho\, g_{\mu\rho}$$

where ξ is a vector field intended as a generator of the coordinate change. The change of the infinitesimal parameter is [r]

$$\delta h_{\mu\nu} = \partial_\mu \xi_\nu + \partial_\nu \xi_\mu = 2\,\partial_{(\mu}\xi_{\nu)}$$

This is the gauge transformation. Considering again the equations of motion for h above, in particular the last two terms, one notices the following gauge transformation

$$\delta(\partial^\mu h_{\mu\nu} - \frac{1}{2}\,\partial_\nu h) = \partial^\mu(\partial_\mu \xi_\nu + \partial_\nu \xi_\mu) - \partial_\nu \partial^\mu \xi_\mu = \Box\,\xi_\nu\,, \quad \delta h = 2\,\partial^\mu \xi_\mu$$

After solving the equation above ? [r], one sets

$$\partial^\rho h_{\rho\nu} - \frac{1}{2}\,\partial_\nu h = 0 \implies \Box\, h_{\mu\nu} = 0$$

There are further gauge invariances that can be used keeping the gauge fixed. The residual gauge invariance is

$$\Box\,\xi_\mu = 0$$

Until one fixes the gauge invariance, one may not know if the solution is physical.

Transforming to momentum space, and considering a monochromatic wave, one has

$$h_{\mu\nu} = \mathrm{Re}[h_{\mu\nu}(k)\mathrm{e}^{\mathrm{i}k^\mu x_\mu}] \implies k^2 h_{\mu\nu}(k) = 0\,, \quad k^\rho h_{\rho\nu} = \frac{1}{2} k_\nu h$$

One also has

$$\delta h = 2k^\mu \xi_\mu$$

given a vector, it is not difficult to find another vector with arbitrary inner product; so one set the trace to zero $h = 0$. Therefore

$$k^\rho h_{\rho\nu} = 0$$

The residual gauge invariances are

$$k^\mu \xi_\mu = 0\,, \quad k^2 \xi_\mu(k) = 0$$

The first condition is required to keep $h = 0$.

One breaks the Lorentz invariance of flat space by choosing a timelike vector u^μ such that

$$u^\mu k_\mu = 1$$

It follows

$$-\mathrm{i}\delta(u^\mu h_{\mu\nu}) = u^\mu(k_\mu \xi_\nu + k_\nu \xi_\mu) = \xi_\nu + u^\mu k_\nu \xi_\mu = \xi_\mu M^\mu{}_\nu\,, \quad M^\mu{}_\nu = \delta^\mu{}_\nu + u^\mu k_\nu$$

The map associated to the matrix M is surjective and as such one can obtain anything [r]. To show this, one inverts the matrix

$$(M^{-1})^\mu{}_\nu = \delta^\mu{}_\nu - \frac{1}{2} u^\mu k_\nu$$

[r] The quantity $u^\mu h_{\mu\nu}$ can be set to zero. The residual gauge invariance is

$$\xi_\mu M^\mu{}_\nu = 0 \implies \xi_\mu = 0$$

[r] Therefore, one arrives at the transverse traceless (TT) gauge

$$k^2 h_{\mu\nu} = 0\,, \quad h = 0\,, \quad k^\mu h_{\mu\nu} = 0\,, \quad u^\mu h_{\mu\nu} = 0$$

Example. Consider

$$u^\mu = (1, 0, 0, 0)\,, \quad k^\mu = (1, 1, 0, 0)$$

The four conditions above imply

$$h_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & h_+ & h_\times \\ 0 & 0 & h_\times & -h_+ \end{bmatrix}$$

There are two polarization, just like in electromagnetism. Gravitational waves exist.

Physical meaning. One can study the effect of gravitational waves on matter, in particular how curvature affects two nearby particles. Consider a line of free-falling particles: in space-time their geodesics are parallel. Consider $T = \partial_\tau$ the vector field pointing along time, while $d = \partial_s$ pointing along a space direction. Their Lie bracket is

$$[T, d] = [\partial_\tau, \partial_s] = 0$$

Since the particles are in free-fall then

$$T^\mu \nabla_\mu T^\nu = 0$$

The expression of the Lie bracket is

$$0 = [T, d] = T^\mu \, \partial_\mu d^\nu - d^\mu \, \partial_\mu T^\nu = T^\mu \nabla_\mu d^\nu - d^\mu \nabla_\mu T^\nu$$

Remembering the notation for the covariant derivative in a particular direction

$$T^\mu \nabla_\mu \equiv \nabla_T$$

Then one has

$$\nabla_T d^\nu = (\nabla_\mu T^\nu) d^\mu$$

This is similar to a first order equation for d in terms of time T. A second order equation is

$$\begin{aligned}
\nabla_T \nabla_T d^\nu &= \nabla_T (d^\mu \nabla_\mu T^\nu) = [\nabla_T, d^\mu \nabla_\mu] T^\nu + d^\mu \nabla_\mu (\nabla_T T^\nu) = [\nabla_T, d^\mu \nabla_\mu] T^\nu + 0 \\
&= [T^\rho \nabla_\rho, d^\mu \nabla_\mu] T^\nu = (T^\rho [\nabla_\rho, d^\mu \nabla_\mu] + [T^\rho, d^\mu \nabla_\mu]) T^\nu \\
&= \{T^\rho (\nabla_\rho d^\mu) \nabla_\mu + T^\rho d^\mu [\nabla_\rho, \nabla_\mu] + d^\mu (-\nabla_\mu T^\rho)\} T^\nu \\
&= T^\rho d^\mu [\nabla_\rho, \nabla_\mu] T^\nu = R_{\rho\mu}{}^\nu{}_\sigma T^\sigma T^\rho d^\mu
\end{aligned}$$

In the first line, the second addendum is zero thanks to the geodesic equation. The Riemann tensor describes how curvature changes over time.

One can study how a particle moves when a gravitational waves goes through. Taking

$$T = (1, 0, 0, 0) \,, \quad \ddot{d}_i = R_{0ji0} d^j$$

The Riemann tensor is

$$R_{\mu\nu\rho\sigma} = \partial_\nu \partial_{[\rho} h_{\sigma]\mu} - \partial_\mu \partial_{[\rho} h_{\sigma]\nu} \implies R_{i00j} = \frac{1}{2} \ddot{h}_{ij}$$

Therefore, for small h_{ij}, one has

$$\ddot{d}_i = \frac{1}{2} \ddot{h}_{ij} d_j \implies d_i \sim d_i^0 + \frac{1}{2} h_{ij} d_j^0$$

Depending on h_{ij}, the motion changes. Consider the example above. Consider the polarization h_+: for a displacement in the x direction, nothing happens; for a displacement in the y direction, the particle oscillates along y. Similarly for z, but with opposite oscillation. Consider the polarization $h_\times$: for displacement in y a particle moves in z, while a displacement in z one moves along y with opposite oscillation. The two polarization resemble a $+$ and a $\times$ when looking at a ring of particles in the plane yz.

Lecture 24

gio 11 gen 2024 14:30

9 Spinors

No exam.

Special relativity. In Special Relativity can introduce the Dirac matrices that obey anti-commutation relations
$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$$
defining a Clifford algebra. One may write a commutation relation
$$\gamma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu] \implies [\gamma_{\mu\nu}, \gamma^{\rho\sigma}] = -8\delta_{[\mu}{}^{[\rho}\gamma_{\nu]}{}^{\sigma]}$$
The set of the Dirac matrices close under the commutation: it is the same commutation relations of the Lorentz algebra and generators $J_{\mu\nu}$ (when considering $-\frac{1}{2}\gamma_{\mu\nu}$). The matrices are a representation of the Lorentz algebra. The space on which the matrices act transforms naturally under Lorentz transformation.

The vector space on which the matrices act has minimum dimension of 4 and, in that dimension, the choice of the matrix is equivalent up to a change of basis. The space is called space of spinors ψ so the representation is the spinor representation. The infinitesimal action of the matrices is
$$\delta\psi = -\frac{1}{4}\psi^{\mu\nu}\gamma_{\mu\nu}\psi$$
The extra factor of $\frac{1}{2}$ comes from the fact that λ is anti-symmetric and accounts for the repeating entries. One considers a linear combination of possible transformations through the infinitesimal parameter $\psi^{\mu\nu}$. The finite transformation is given by
$$\Lambda^\mu{}_\nu = \exp[\lambda]^\mu{}_\nu \implies \psi' = \exp\left[-\frac{1}{4}\lambda^{\mu\nu}\gamma_{\mu\nu}\right]\psi$$
One can utilize Feynman's slash notation when treating multiple gamma matrices
$$\not{\lambda} \equiv -\frac{1}{2}\lambda_{\mu\nu}\gamma^{\mu\nu}$$
The factor $\frac{1}{2}$ describes the fact that a rotation in physical space corresponds to a half rotation in spinor space.

In a vector space one would like an inner product. The product $\psi^\dagger\psi$ is not Lorentz invariant. In fact, the Dirac matrices cannot be all hermitian because
$$(\gamma^0)^2 = -1 \implies (\gamma^0)^\dagger = -\gamma^0$$
The square eigenvalues has to be -1 and as such the eigenvalues are $\pm$i so it is anti-hermitian. The other three gamma matrices can be hermitian since
$$(\gamma^i)^2 = 1 \implies (\gamma^i)^\dagger = \gamma^i$$
Therefore
$$\gamma_\mu^\dagger\gamma_0 = -\gamma_0\gamma_\mu$$
The variation of the product above is then
$$\delta(\psi^\dagger\psi) = -\frac{1}{2}\psi^\dagger(\not{\lambda} + \not{\lambda}^\dagger)\psi$$
The parameter $\not{\lambda}$ has no definite (anti-)hermicity, so the product is not zero in general and is not an invariant. However, the property above suggests a way to correctly define the inner product
$$\bar{\psi} = \psi^\dagger\gamma_0 \implies \bar{\psi}\psi$$

General relativity. Starting from the anti-commutation relation, one has
$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$$
One can reduce the discussion to the case of Special Relativity: one defines an orthonormal basis of vector fields (a vierbein, in other dimensions it is called vielbein). The basis vector fields are
$$E_a\,, \quad a = 0, 1, 2, 3\,, \quad E_a = E_a^\mu\partial_\mu$$

The index a is called flat index, while μ is a curved index (its position is important). The inner product of two of them is

$$E_a \cdot E_b = E_a^\mu E_b^\nu g_{\mu\nu} = \eta_{ab}$$

From this it follows

$$\{E_a^\mu \gamma_\mu, E_b^\nu \gamma_\nu\} = 2\eta_{ab}$$

One then defines

$$\gamma_a \equiv E_a^\mu \gamma_\mu$$

and the anti-commutator above reduces to Special Relativity's Clifford algebra. Each matrix E_a^μ is invertible

$$E_a^\mu e_\mu^b = \delta_a{}^b \,, \quad e^a = e_\mu^a \, \mathrm{d}x^\mu \,, \quad e_\mu^a = g_{\mu\nu} E^{a\nu}$$

The inverse is both a left and right inverse and it is a one-form. Their inner product is then

$$e^a \cdot e^b = e_\mu^a e_\nu^b g^{\mu\nu} = \eta^{ab}$$

One also has

$$e_\mu^a e_\nu^b \eta_{ab} = g_{\mu\nu}$$

Therefore, the gamma matrices are

$$\gamma_\mu = \gamma_a e_\mu^a$$

The theory of spinors in Special Relativity is based on the Lorentz group $\mathrm{SO}^+(3,1)$, but in General Relativity the analogue of the group does not exist. [r] The change of coordinate act with the Jacobian matrix on the vectors. In this section one has developed a theory where spinors transform under the Lorentz group but not matrices in general. [r]

The matrices E_a and their inverse are not unique

$$E_a \to \Lambda_a{}^b E_b$$

where Λ is a local Lorentz transformation (because it is point-dependent). The transformed matrices are also a vierbein.

Spinorial covariant derivative. Consider the following bilinear

$$\bar{\psi}_1 \gamma_\mu \psi_2$$

One would like to make sense of it in General Relativity: it must be a vector field. The Dirac matrices transforms under a vierbein

$$\gamma_\mu = \gamma_a e_\mu^a$$

[r] One makes ψ also depend on the local Lorentz transformation $\Lambda^a{}_b = \exp(\lambda)^a{}_b$:

$$\psi \to \exp\left[-\frac{1}{2}\not{\lambda}\right]\psi$$

Then the bilinear above is invariant under local Lorentz transformations.

Now one studies how the derivative transforms. It does not transform as the spinor but as

$$\partial_\mu \left[\exp\left(-\frac{1}{2}\not{\lambda}\right)\psi\right]$$

One introduces a new covariant derivative

$$D_\mu\psi \to \exp\left[-\frac{1}{2}\not{\lambda}\right] D_\mu\psi$$

The covariant derivative of the veirbein

$$\nabla_\mu E_a = -\omega_{\mu a}{}^b E_b$$

can be expanded on a basis of vector fields for a fixed μ and a. One has

$$E_a = E_a^\nu \, \partial_\nu \implies \nabla_\mu E_a^\nu = \partial_\mu E_a^\nu + \Gamma^\nu_{\mu\rho} E_a^\rho = -\omega_{\mu a}{}^b E_b$$

One may define a new type of derivative such that

$$\partial_\mu E_a^\nu + \Gamma^\nu_{\mu\rho} E_a^\rho + \omega_{\mu a}{}^b E_b = 0$$

For the inverse one has

$$\nabla_\mu e^a = -\omega^a_{\mu b} e^b \,, \quad \partial_\mu e^a_\nu - \Gamma^\rho_{\mu\nu} e^a_\rho + \omega^a_{\mu b} e_\nu b = 0$$

The covariant derivative of a product is

$$0 = \nabla_\mu \eta^{ab} = \nabla_\mu (e^a \cdot e^b) = -\omega^a_{\mu c} e^c e^b - e^a \omega^b_{\mu c} e^c = -\omega^a_{\mu c} \eta^{cb} - \omega^b_{\mu c} \eta^{ac} = -\omega^{ab}_\mu - \omega^{ba}_\mu$$

since the flat metric can raise and lower the indices. Therefore

$$\omega^{ab}_\mu = -\omega^{ba}_\mu$$

For infinitesimal local Lorentz transformations one has

$$\delta e^a = \lambda^a{}_b e^b \implies \delta \omega^{ab}_\mu = -\partial_\mu \lambda^{ab} + [\lambda, \omega_\mu]^{ab} \,, \quad [\lambda, \omega_\mu]^{ab} = \lambda^a{}_c \omega^{cb}_\mu - \omega^{ac}_\mu \lambda_c{}^b$$

So ω like a gauge potential but with an extra term [r]. Therefore, the spinorial covariant derivative is

$$D_\mu \psi \equiv \partial_\mu \psi + \frac{1}{4} \omega^{ab}_\mu \gamma_{ab} \psi \,, \quad \delta(D_\mu \psi) = -\frac{1}{2} \not{\lambda} D_\mu \psi$$

This derivative deals with the issue of local Lorentz transformations. In this way, one can promote the action of Special Relativity to General Relativity.

Structure equations. The spin connections ω are related to the Riemann tensor. Antisymmetrizing [r] in $\mu\nu$ one has

$$\partial_{[\mu} e^a_{\nu]} + \omega^{ab}_{[\mu} e_{\nu]b} = 0$$

In this way one can find the spin connections. The relation to the Riemann tensor is the following. Consider

$$\begin{aligned}
\frac{1}{2} [\nabla_\mu, \nabla_\nu] E^a &= \nabla_{[\mu} \nabla_{\nu]} E^a = \nabla_{[\mu} (-\omega^{ab}_{\nu]} E_b) = -\nabla_{[\mu} \omega^{ab}_{\nu]} E_b - \omega^{ab}_{[\nu} (-\omega^c_{\mu]b} E_c) \\
&= -(\nabla_{[\mu} \omega^{ab}_{\nu]} + \omega^{ac}_{[\mu} \omega^c_{\nu]b}) E_c
\end{aligned}$$

Therefore one has

$$[\nabla_\mu, \nabla_\nu] E^{a\rho} = R_{\mu\nu}{}^\rho{}_\sigma E^{a\sigma} \implies \frac{1}{2} E^{a\rho} E^{a\sigma} R_{\rho\sigma\mu\nu} \equiv R^{ab}{}_{\mu\nu} = \partial_{[\mu} \omega^{ab}_{\nu]} + \omega^{ac}_{[\mu} \omega^b_{\nu]c}$$

The analogue of the covariant derivative is

$$[D_\mu, D_\nu] \psi = \frac{1}{4} R^{ab}{}_{\mu\nu} \gamma_{ab} \psi$$

The Lie–Kosmann derivative is

$$L_v \psi = v^\mu D_\mu \psi + \frac{1}{4} \partial_{\not{\mu}} v_\nu \gamma^{\mu\nu} \eta$$