Lecture Notes: 002

Introduction to Relativity

Based on Ray d'Inverno, Introducting Einstein's Relativity

March 6, 2024

Contents

CHAPTER 0

Introduction

0.1 Basic information

- Time: 13:15-14:30 (Tu. and Thur.)
- Textbook: **Ray d'Inverno, Introducing Einstein's Relativity** (Oxford) 1993 1st Ed. / 2022 2nd Ed.
- References:

  (1) P.A.M Dirac, General theory of relativity (Princeton) 1975

  (2) S. Weinberg, General relativity and cosmology (John Wiley) 1972

  (3) 't Hooft, Introduction to general relativity (Rinton) 2001

  (4) A. Zee, Einstein gravity in a nutshell (Princeton) 2013

  (5) D. Tong, General relativity, Lecture note (Cambridge) 2019
- Grading policy: HW (100%)
- Login: E-course website (https://ecourse.ccu.edu.tw)
- Office hours: (Tue, Thur)

The textbook covers the following topics:

1. Introduction (GR0)
2. Special relativity (GR1-GR6)
3. Riemannian geometry (GR7-GR15)
4. General relativity (GR16-GR29)
5. Gravitational waves (GR30)
6. Black holes (GR31)
7. Cosmology (optional)

[0] GR0.tex

0.2 Preliminary

- Mechanics and electrodynamics (Marion and Griffiths)
- Calculus (basic)
- Tensor analysis (reviewed in Ch. 5 and 6)
- Differential equations (optional)

0.3 Physical constants

$$
\begin{aligned}
c &= 2.997 \times 10^8 m/s \qquad \text{speed of light} \\
\hbar &= 1.05 \times 10^{-34} J \cdot s \qquad \text{Planck's constant} \\
G &= 6.67 \times 10^{-11} m^3/kg \cdot s \qquad \text{Newton's constant}
\end{aligned}
$$

These constants play different roles in physical theories.

- Classical mechanics (Newton): $c \to \infty, \hbar \to 0, G < \infty$.
- Special theory of relativity (Einstein): $c < \infty, \hbar \to 0, G = 0$
- Quantum mechanics (Schrödinger, Heisenberg): $c \to \infty, \hbar < \infty, G = 0$
- Quantum field theory (Dirac, Pauli, et al): $c < \infty, \hbar < \infty, G = 0$
- General theory of relativity (Einstein): $c < \infty, \hbar \to 0, G < \infty$

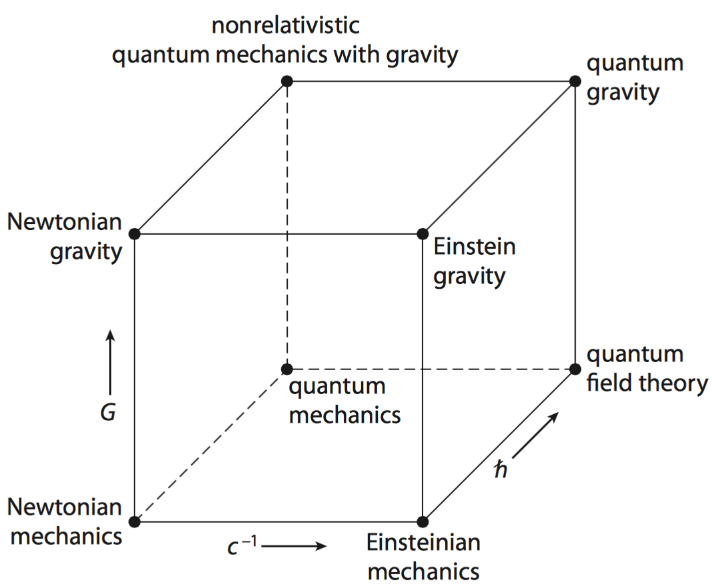

Figure 1: The cube of physics

Remark. Using dimensional analysis we have

$$[c] = \frac{L}{T}, \quad [G] = \frac{L^3}{MT^2}, \quad [\hbar] = \frac{ML^2}{T}$$

from that we can construct three fundamental units

$$\begin{aligned} l_P &= \sqrt{\frac{G\hbar}{c^3}} \simeq 10^{-33}cm \quad \text{(Planck length)} \\ t_P &= \sqrt{\frac{G\hbar}{c^5}} \simeq 10^{-44}sec \quad \text{(Planck time)} \\ m_P &= \sqrt{\frac{c\hbar}{G}} \simeq 10^{19}Gev/c^2 \quad \text{(Planck mass)} \end{aligned}$$

0.4 Fundamental forces

There are four known fundamental forces in nature:

- Electromagnetic force (interaction between matter and photon)
- Weak force (nuclei decay)
- Strong force (interaction between quarks)
- Gravitational force (interaction between massive objects)

The first unification of forces

$$\left.\begin{array}{c}\text{Electric force}\\ \text{Magnetic force}\end{array}\right\}\text{Electromagnetic force (Maxwell)}$$

The quantum electrodynamics (QED) is the quantum version of electromagnetic field interacting with matter. (Feynman-Schwinger-Tomonaga)

$$\left.\begin{array}{c}\text{Quantum electrodynamics (QED)}\\ \text{Weak force}\end{array}\right\}\text{Electroweak force (Weinberg-Salam)}$$

$$\left.\begin{array}{c}\text{Electroweak force}\\ \text{Strong force (QCD)}\end{array}\right\}\text{Standard Model}$$

where QCD stands for quantum chromodynamics.

$$\left.\begin{array}{c}\text{Standard model}\\ \text{Gravitational force}\end{array}\right\}\text{String theory???}$$

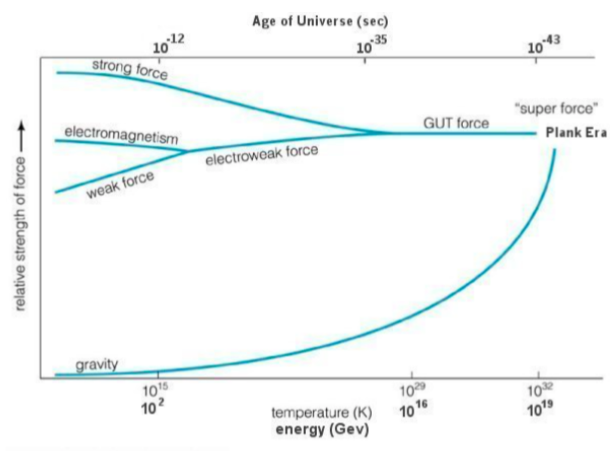

Figure 2: Unification of fundamental forces

0.5 Einstein equation

The Einstein equation which governs the evolution of spacetime and matter is given by

$$\boxed{G^{ab} = \kappa T^{ab}, \quad \kappa = \frac{8\pi G}{c^4}}$$

where

$$\begin{aligned} G^{ab} &= \text{Einstein curvature (geometry)} \\ T^{ab} &= \text{Energy-momentum tensor (matter)} \end{aligned}$$

Remark. By dimensional analysis

$$[G] = L^3/MT^2, \quad [\kappa] = T^2/ML, \quad [T^{00}] = M/LT^2, \quad [G^{00}] = 1/L^2$$

0.6 Minkowski metric

In the lecture we shall use the following convention for the Minkowski metric

$$\eta_{ab} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

CHAPTER 1

Special Relativity I

Einstein invented special theory of relativity (SR) so that Newton theory is just a low velocity limit of SR.

1.1 Brief history

1865 Maxwell showed that EM wave = light.

What is the medium to carry the light wave? Ether postulated

1881 Michelson-Morley experiment: to detect the relative motion between earth and ether. The deviation is very tiny

1895 Some hypotheses are introduced to overcome the contradictions: e.g. Lorentz, Fitzyerald and Poincaré to expalin

(1) contraction of rigid bodies (length contraction)

(2) Slowing down of clocks (time dilation)

They are consistent with observations, but has philosophical defects

1905 Einstein: Lorentz transformation can be derived from two postulates

(1) principle of special relativity

(2) constancy of the speed of light

It turns out that the most important concept is 'simultaneity'"

An **event** can be denoted by a pair of numbers (t, x) where t is an instant in time and x is a point in space. Hence, the motion of an object is a collection of events described by space-time diagram.

[0]GR1.tex

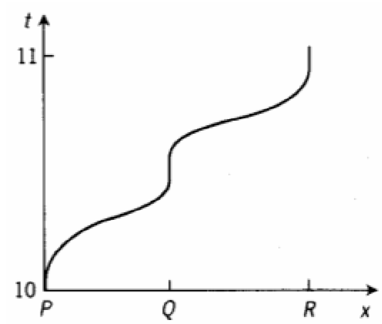

Figure 1.1: Trajectory in $1+1$-dimensional space-time diagram

Observer: a man with a clock and a ruler

Frame of reference: an observer chose a coordinate system

S.R. is just concerned the relationship between different observers seeing the same phenomena

Recall Newton's point of view: free motion = particle is at rest or motion in straight line. Thus, an inertial frame can be described by the co-moving frame of an observer with free motion.

Newton's 1st law: A free particle has $\vec{a}=0$. Inertial frames are connected by Galilean transformation

$$\begin{aligned} x &= x' + vt \\ y &= y' \\ z &= z' \\ t &= t' \quad \text{(time is absolute !)} \end{aligned}$$

Some consequences of the Galilean transformation:

(1) When we measure the length of an object, we shall measure its end points at once, which implies

$$\Delta x = \Delta x', \quad \Delta y = \Delta y', \quad \Delta z = \Delta z', \quad \Delta t = \Delta t',$$

hence the length of an object and time duration of a clock are independent of the motion of the object.

(2) From Galilean transformation, we have additional formula of velocity

$$\dot{x} = \dot{x}' + v, \quad \dot{y} = \dot{y}', \quad \dot{z} = \dot{z}'$$

(3) The acceleration of the object is unchanged

$$\ddot{x} = \ddot{x}', \quad \ddot{y} = \ddot{y}', \quad \ddot{z} = \ddot{z}'$$

In summary:

- Newton: It's a restricted principle of special relativity, and all inertial observers are equivalent as far as dynamical experiments are concerned

- Einstein: Two postulates of S.R.

  PI Principle of special relativity: all inertial observers are equivalent (no pure dynamical experiment)

  PII Constancy of speed of light: the speed of the light is the same in all inertial frame

$$c = 2.997924580 \times 10^8 m/s$$

(we shall set $c=1$ for relativistic unit)

1.2 The k-factor

This is a factor to detect the difference between two inertial frames. Denoting observer A (at rest), observer B (moving away from A with velocity v), and light signal C ($v = c = 1$). Thus, in space-time diagram, A is a vertical line, B is a straight line of slope $v > 1$, and C is a straight line of slope $c = 1$.

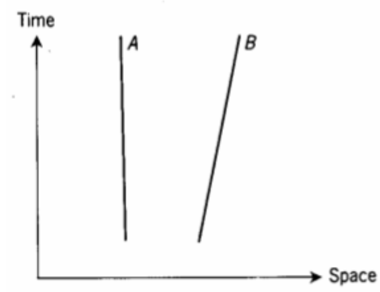

Figure 1.2: The world lines of A and B

Suppose the time interval of flash light in A's clock is T and the time interval of flash light in B's clock is kT. Then the parameter k has to characterize the relative motion between A and B.

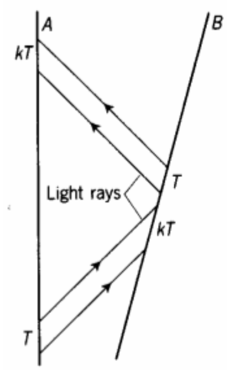

Figure 1.3: The k factor defined by flash light

Remark. We have assumed that the relationship of time and space between A and B is linear.

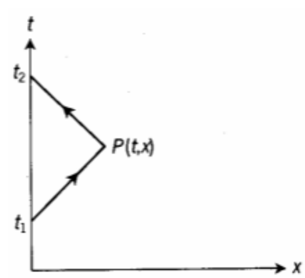

Figure 1.4: Coordinating events

Before exploring the factor k, let's use radar to characterize an event (t, x). If a rest light signal emitted at t_1, reflected by P at (t, x) and received at t_2, then we have

$$x = (t_2 - t_1)/2, \quad t = (t_1 + t_2)/2$$

Consider two observers A and B that are synchronized at O. If the light signal emitted at T bunching back by B at $P = (t, x)$ and received by A again. Based on above, we know that $t_1 = T$ and $t_2 = k^2 T$. Hence

$$t = \frac{1}{2}(k^2 + 1)T, \quad x = \frac{1}{2}(k^2 - 1)T$$

and the relative speed of A and B is

$$v = \frac{x}{t} = \frac{k^2 - 1}{k^2 + 1} < 1$$

or

$$k = \sqrt{\frac{1 + v}{1 - v}} > 1$$

Remark. When $v \to -v$ we have $k \to k^{-1}$. If $v = 0$ then $k = 1$.

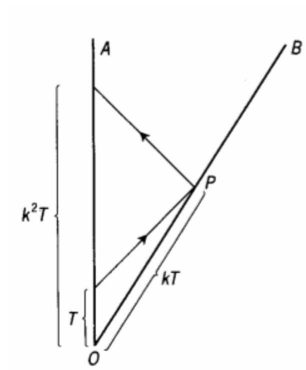

Figure 1.5: The physical meaning of k

Let's denote T as the period of light signal and the corresponding angular frequency is $\omega = 2\pi/T$.

If $k > 1$ (or B is away from A), then

$$\omega_A = \frac{2\pi}{T} > \omega_B = \frac{2\pi}{kT} \quad \text{(red shift)}$$

If $k < 1$ (or B approaches to A), then

$$\omega_A = \frac{2\pi}{T} < \omega_B = \frac{2\pi}{kT} \quad \text{(blue shift)}$$

where ω_A/ω_B is the frequency observed by A/B. This is just the relativistic formula for radar.

Remark. Bondi k-calculus is a method of teaching special relativity popularised by Sir Hermann Bondi.

CHAPTER 2

Special Relativity II

2.1 The k-factor (cont'd)

Finally, let's consider three observers A, B, and C and denote the relative speeds as v_{AB}, v_{BC}, and v_{AC}.

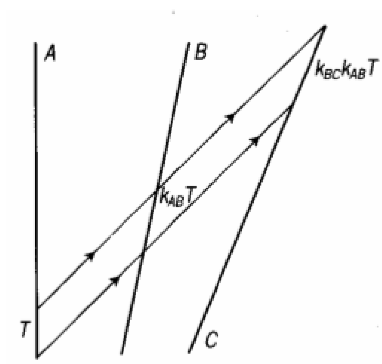

Figure 2.1: Composition of speeds

From the discussion of two observers, we know that a time interval T of light signal observed by A corresponds to $k_{AB}T$ observed by B, and $k_{BC}k_{AB}T = k_{AC}T$ observed by C where

$$k_{AB} = \sqrt{\frac{1+v_{AB}}{1-v_{AB}}}, \quad k_{BC} = \sqrt{\frac{1+v_{BC}}{1-v_{BC}}}, \quad k_{AC} = \sqrt{\frac{1+v_{AC}}{1-v_{AC}}}$$

Hence, $k_{AC} = k_{AB}k_{BC}$ which implies the additional formula of speeds

$$v_{AC} = \frac{v_{AB}+v_{BC}}{1+v_{AB}v_{BC}}.$$

Notice that

(1) As $v_{AB}, v_{AC} << 1$ (low speed limit), the additional formula recovers the Galilean addition law

$$v_{AC} \simeq v_{AB} + v_{BC}$$

(2) If $v_{AB} = 1$, then $v_{AC} = 1$ as well.

[0]GR2.tex

(3) If $v_{AB}, v_{BC} < 1$, then

$$v_{AC} - 1 = \frac{v_{AB} + v_{BC}}{1 + v_{AB}v_{BC}} - 1 = \frac{(v_{AB} - 1)(1 - v_{BC})}{1 + v_{AB}v_{BC}} < 0$$

That means the speed of light $c = 1$ is the upper limit of speed in any inertial frame.

Remark. We may classify particles by their speeds as

(a) massive particle: $v < 1$ e.g. electron, neutron, etc

(b) massless particles: $v = 1$ e.g. photon, graviton, neutrino* (Super K exp.)

(c) faster-than-light particles: $v > 1$ e.g. tachyon

2.2 Simultaneity

Now we come to a more strange issue in theory of special relativity, the concept of simultaneity. Recall in Newton's theory that if two events are simultaneous in an inertial frame then they do as well in another. However, this is not the case in the theory of special relativity. Einstein provided the following "thought experiment" to explain the concept.

Suppose observer B sits in the middle of a train moving with speed v and there are two lamps on the track separated by a distance of the length of the train. When B passing the still observer A in the train station, the lamps are triggered by the ends of the train to lightening.

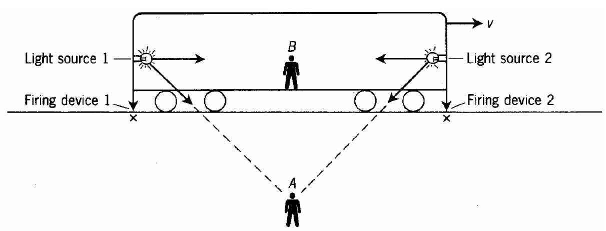

Figure 2.2: Einstein's thought experiment for simultaneity

It turns out that

For observer A, lamp 1 and lamp 2 are turned on simultaneously.

For observer B, lamp 2 was turned on earlier than lamp 1.

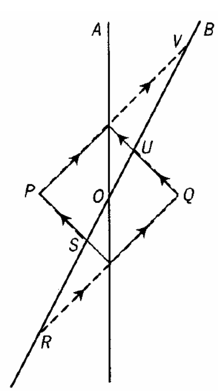

Figure 2.3: World-line of Einstein's thought experiment

To understand this result we plot the world-lines of observers and light emitted by lamps. The signals of lamp 1 and lamp 2 reach to A simultaneously, while the lamp 2 reach earlier than lamp 1 for observer B. Hence, the concept of simultaneity depends on the motion of an observer.

Remark. The world-line of light is called light cone. The light cone separates the space-time into three regions: the absolute future, the absolute past and elsewhere.

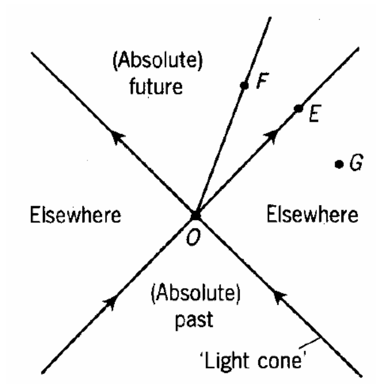

Figure 2.4: Light cone and causality

Point O: present
Point F: absoluly future
Point E: can be communicated by light
Point G: might be relative past, relative future, or simultaneously w.r.t. O
Example. (clock paradox)

2.3 Lorentz transformation

Lorentz transformation can be derived from the postulated of special relativity by many ways. Here we use the k-calculus to do that. Consider two observers A and B so that B moves with speed v w.r.t. A and coincides with A at $t = x = 0$.

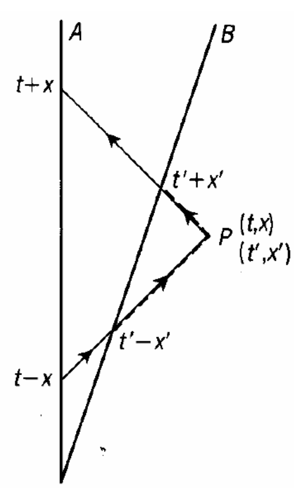

Figure 2.5: Lorentz transformation derived by k-calculus

By radar exploring scheme, the space-time coordinate of the event P is (t,x) for observer A and (t',x') for B where the emitting time t_1 and receiving time t_2 for A are

$$t_1 = t - x, \quad t_2 = t + x$$

and

$$t_1' = t' - x', \quad t_2' = t' + x'.$$

By k-calculus, we have $t_1' = kt_1$ and $t_2 = kt_2'$ which implies

$$t' - x' = k(t - x), \quad t + x = k(t' + x')$$

A simple algebra shows that

$$t' = \frac{t - vx}{\sqrt{1-v^2}}, \quad x' = \frac{x - vt}{\sqrt{1-v^2}}$$

Remark.

(1) (t,x) and (t',x') are space-time coordinates referred to the same events by two different observers. They satisfies the relation

$$(t')^2 - (x')^2 = t^2 - x^2 = \text{constant}$$

(2) We can generalize space-time coordinate (t,x) to the standard one (t,x,y,z) with three spatial dimensions. Since under boost in x-direction, the xz-plane (or $y=0$) coincides with the $x'z'$-plane ($y'=0$) and thus we have $y = ny'$. Now due to the assumption of isotropic property under parity transformation $x \to -x, y \to -y$, A and B interchange their roles, we get $y' = ny = n^2y'$. Thus $n = \pm 1$, however the limit $v \to 0$ fixes $n = +1$, that is $y = y'$. Similarly, we have $z = z'$.

In summary:

Galilean transformation	Lorentz transformation
$t' = t$	$t' = \frac{t-vx}{\sqrt{1-v^2}}$
$x' = x - vt$	$x' = \frac{x-vt}{\sqrt{1-v^2}}$
$y' = y$	$y' = y$
$z' = z$	$z' = z$

- invariant for two events (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) is

$$s^2 = (t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2$$

- infinitesimal distance between events is

$$(ds)^2 = (dt)^2 - (dx)^2 - (dy)^2 - (dz)^2$$

- distance between two spatial points is

$$\sigma^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

CHAPTER 3

The Key Attributes of SR I

3.1 Lorentz transformation

Let's give a standard derivation of the Lorentz transformation based on two postulates:

PI. The principle of special relativity

PII. The constancy of the speed of light

By PI: If a particle is free in S frame then it is free as well in S' frame, i.e.

$$\vec{r} = \vec{r}_0 + \vec{u}t \xrightarrow{L} \vec{r}' = \vec{r}'_0 + \vec{u}'t'$$

Thus the transformation L is linear

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = L \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

where, for the boost in x-axis only, L is given by

$$L = \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} & 0 & 0 \\ -\frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Remark. We shall use the convection $\beta \equiv v/c$ and $\gamma \equiv 1/\sqrt{1-\beta^2}$.

By PII: The speed of light c is a constant in any inertial frame and in any direction. Thus the wave front of light in S is a sphere with radius $R = ct$. Hence

$$I(t, x, y, z) = x^2 + y^2 + z^2 - c^2t^2 = 0.$$

where (t, x, y, z) is the coordinate on the wave front. Since this is the same for S'

$$I' = (x')^2 + (y')^2 + (z')^2 - c^2(t')^2 = 0.$$

[0]GR3.tex

we have, for light front,

$$I = 0 \Longleftrightarrow I' = 0.$$

Due to the fact that S and S' are relative inertial frames, we have

$$I = nI' \quad \text{or} \quad I' = nI$$

which implies $I = n^2 I$ or $n = \pm 1$. Since $I \to I'$ for $v \to 0$ we have $n = +1$. Hence

$$x^2 + y^2 + z^2 - c^2t^2 = (x')^2 + (y')^2 + (z')^2 - c^2(t')^2.$$

For a simple boost in x-direction, it can be check that

$$x^2 - c^2t^2 = (x')^2 - c^2(t')^2.$$

Introducing the imaginary time $T = ict$ where $i = \sqrt{-1}$ the above equation becomes

$$x^2 + T^2 = (x')^2 + (T')^2$$

That is the Lorentz boost between inertial frames looks like a rotation in 2-dimensional plane

$$\begin{pmatrix} x' \\ T' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ T \end{pmatrix}$$

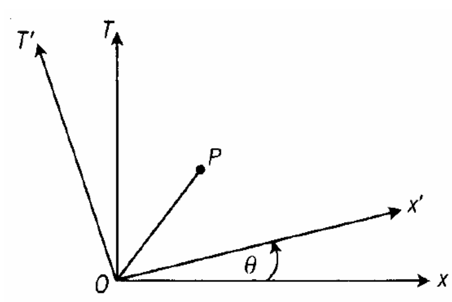

Figure 3.1: Rotation in (x, T) space

Consider an observer seats at O', then the world line of its motion in S is the T'-axis (or $x' = 0$). So

$$x\cos\theta + T\sin\theta = 0$$

or

$$v = \frac{x}{t} = ic\frac{x}{T} = -ic\tan\theta$$

where θ is pure imaginary. Thus

$$\cos\theta = \frac{1}{\sec\theta} = \frac{1}{\sqrt{1+\tan^2\theta}} = \frac{1}{\sqrt{1 - v^2/c^2}} \equiv \gamma$$

The Lorentz transformation now becomes

$$x' = \cos\theta(x + \tan\theta T) = \gamma(x + \frac{iv}{c} \cdot ict) = \gamma(x - vt)$$

and, similarly,

$$t' = \gamma(t - vx/c^2)$$

Let's extract some mathematical properties associated with the Lorentz transformation from the boost (a special L.T.)

- Using imaginary time $T = ict$, a Lorentz boost along the x-axis of speed v can be viewed as a rotation in $x - T$ space with angle θ where $\tan\theta = iv/c$.
- When $v << c$ (low speed limit), the boost transformation recovers the Galilean transformation.
- The Lorentz transformation can be expressed from S as well as S' point of view as

$$\begin{cases} x' = \gamma(x - vt) \\ y' = y \\ z' = z \\ t' = \gamma(t - vx/c^2) \end{cases} \qquad v \to -v \qquad \begin{cases} x = \gamma(x' + vt') \\ y = y' \\ z = z' \\ t = \gamma(t' + vx'/c^2) \end{cases}$$

- The special Lorentz transformation (i.e. boost) forms a (commutative) group

  (a) Identity: $L(v = 0) = I$

  (b) Inverse: $L(-v)$ which satisfies $L(-v)L(v) = I$.

  (c) Closure: $L(v')L(v) = L(v'')$ where v is the relative speed of S and S', v' is the relative speed of S' and S'', and v'' is the relative speed of S and S''. Then

$$\tan\theta'' = \tan(\theta + \theta') = \frac{\tan\theta + \tan\theta'}{1 - \tan\theta\tan\theta'} = \frac{iv/c + iv'/c}{1 + vv'/c^2} = iv''/c$$

  or

$$v'' = \frac{v + v'}{1 + vv'/c^2}$$

  which is just the additional formula for speed.

- Associativity: $L_3(L_2L_1) = (L_3L_2)L_1$
- The infinitesimal interval

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

  is invariant under Lorentz transformation.

In the following, for convenience, we assume that S and S' are connected by a Lorentz boost in x-direction.

3.2 Length contraction

If a rod of length $l_0 = x'_B - x'_A$ is at rest in S' where $x'_{A,B}$ are coordinates of rod's ends. By Lorentz transformation,

$$x'_A = \gamma(x_A - vt_A), \quad x'_B = \gamma(x_B - vt_B)$$

The length of the rod measured in S is defined by $l = x_B - x_A$ with $t_A = t_B$ (i.e. measure the coordinates of ends in S simultaneously). Hence, we have

$$l_0 = \gamma(x_B - x_A) = \gamma l$$

or

$$l = l_0\sqrt{1 - v^2/c^2} < l_0$$

where l_0 is called proper length of the rod.

Remark.

(1) This is not a "illusory" effect but a real effect due to the simultaneity.

(2) The length contraction effect is relative.

(3) There is no length contraction effect in the direction transverse to the Lorentz boost.

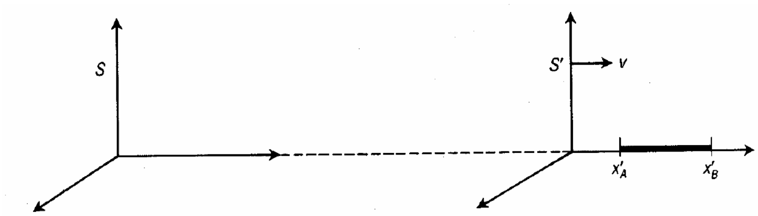

Figure 3.2: Length contraction

3.3 Time dilation

Next, we consider another famous effect called time dilation. If an observer in S' places a clock at x'_A. Then after a time duration T_0 in S', the observer in S will measure the time interval of the clock as

$$T = \Delta t = \gamma(T_0 + v\Delta x'_A/c^2) = \gamma T_0 > T_0$$

where T_0 is called proper time.

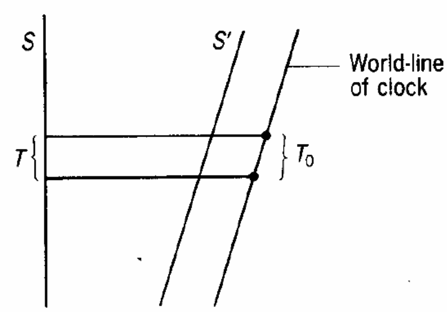

Figure 3.3: Time dilation

For an accelerating clock, the time dilation formula holds for an infinitesimal time interval

$$dt = \gamma d\tau \quad \text{or} \quad d\tau = \gamma^{-1} dt$$

where $d\tau$ is an infinitesimal proper time interval. Then, along the world line of the particle, we have

$$\tau = \int_{t_0}^{t_1} \sqrt{1 - v^2/c^2} dt < \int_{t_0}^{t_1} dt = (t_1 - t_0)$$

Remark.

(1) The time dilation experiment can be carries out by traveling a clock around the earth. (2005, within 4% confirmed the prediction)

(2) The π-meson from atmosphere has a proper life time τ_π. However, the measured lift time $t = L/c > \tau_\pi$ where L is the traveling distance of the π-meson before decaying to other particle.

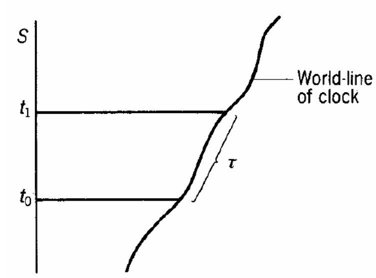

Figure 3.4: Time dilation for an accelerating clock

CHAPTER 4

The Key Attributes of SR II

4.1 Transformation of velocity

Consider two inertial frames S and S' with relative speed v in x-direction. Their space-time coordinates satisfy Lorentz transformation

$$x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma(t - vx/c^2)$$

or infinitesimal form as

$$dx' = \gamma(dx - vdt), \quad dy' = dy, \quad dz' = dz, \quad dt' = \gamma(dt - vdx/c^2).$$

Since the velocity of a particle is defined by

$$(u_1, u_2, u_3) = (\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt})$$

in S and

$$(u_1', u_2', u_3') = (\frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'})$$

in S'.

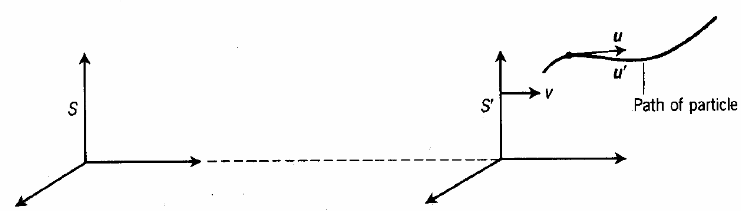

Figure 4.1: Velocity in different inertial frames

Their transformation can be easily derived as

$$\begin{aligned}
u_1' &= \frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma(dt - vdx/c^2)} = \frac{u_1 - v}{1 - u_1 v/c^2} \\
u_2' &= \frac{dy'}{dt'} = \frac{dy}{\gamma(dt - vdx/c^2)} = \frac{u_2}{\gamma(1 - u_1 v/c^2)} \\
u_3' &= \frac{dz'}{dt'} = \frac{u_3}{\gamma(1 - u_1 v/c^2)}
\end{aligned}$$

[0]GR4.tex

where the first equation is just the additional formula shown before. Note that $u_2' \neq u_2$ and $u_3' \neq u_3$ are due to the property of simultaneity for the two frames.

4.2 Space-time diagrams

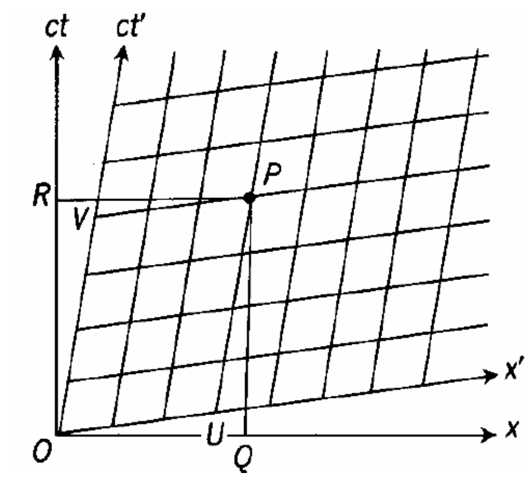

Figure 4.2: Space-time axis for different frames

Label the space-time coordinate of the frame S as (ct, x). Then from Lorentz transformation, we can read out the coordinate axis of S' in S frame as follows

- The x'-axis of S' is defined by $t' = 0$ from which we have

$$0 = t' = \gamma(t - vx/c^2).$$

  Therefore, the x'-axis in S frame is defined by the straight line $ct = (v/c)x$ with slope $v/c < 1$. The equal-time world lines are parallel to this axis.

- The t'-axis of S' is defined by $x' = 0$ from which we have

$$0 = x' = \gamma(x - vt).$$

  Therefore, the t'-axis in S frame is defined by the straight line $ct = (c/v)x$ with slope $c/v > 1$. The equal-position world lines are parallel to this axis.

- It turns out that the Lorentz transformation can be viewed as a folding of the space-time axis toward the light-cone for $v > 0$ and outward for $v < 0$.

How to calibrate the length and time scales between S and S'? Using invariant relation

$$x^2 - c^2t^2 = (x')^2 - c^2(t')^2 = \pm 1.$$

Hence, for $(x')^2 - c^2(t')^2 = +1$, $t' = 0$ implies $x' = +1 = \overline{OA}$. For $(x')^2 - c^2(t')^2 = -1$, $x' = 0$ implies $ct' = +1 = \overline{OB}$. Similarly, for S, the unit scale on ct and x-axes are $\overline{Ob}$ and $\overline{Oa}$ respectively. If an event P has coordinate (OR, OQ) in S and (OV, OU) in S' then, taking the scale into account, the real coordinate are

$$(\frac{OR}{Ob}, \frac{OQ}{Oa}) \quad \text{in } S \quad (\frac{OV}{OB}, \frac{OU}{OA}) \quad \text{in } S'$$

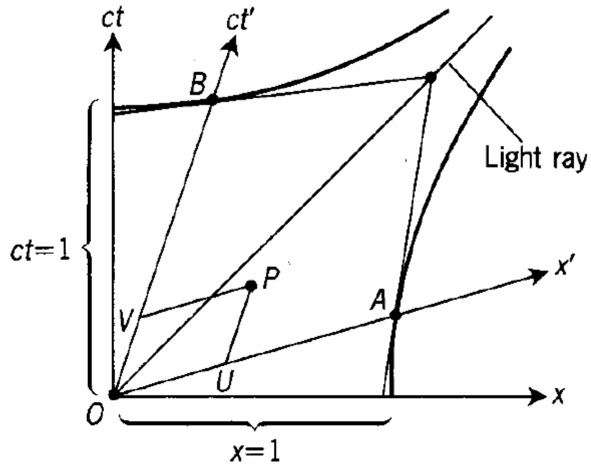

Figure 4.3: Length scales in different inertial frames

Remark.

(1) In terms of imaginary time $T = ict$, the boost can be viewed as a rigid rotation with an imaginary angle θ. However, in real time t, the boost looks like skewing of the coordinate axes.

(2) We shall use the hyperbolae to calibrate the unit distances for all inertial frames.

(3) The length contraction can be explained from the world line of A in S, which intersects at A' on x-axis. Since $\overline{OA} = 1$ in S', but $\overline{OA'} < \overline{Oa} = 1$, we have length contraction indeed.

(4) If $\overline{OB} = 1$ in S', the $\overline{OB'}$ is the time interval observed in S. Since $\overline{OB'} > \overline{Ob} = 1$ we have time dilation. We mention that time dilation is reciprocal between S and S'. If $\overline{Ob} = 1$ in S, then from S' point of view, the observed interval is $\overline{Ob'}$ which is greater than $\overline{OB} = 1$ in S' as expected.

(5) Since $\overline{OA} = 1$ in S', the point A has coordinate $(ct', x') = (0, 1)$. The same point in S has coordinate (ct, x) with

$$x = \gamma(x' + vt') = \gamma, \quad ct = \gamma(t' + vx'/c^2) = \gamma v/c^2$$

Hence, in S

$$\overline{OA} = \sqrt{(ct)^2 + x^2} = \gamma\sqrt{1 + v^2/c^2} = \sqrt{\frac{1 + v^2/c^2}{1 - v^2/c^2}} > 1$$

This is a calibration factor between S and S'.

4.3 Acceleration in SR

Recall the additional formula for velocity

$$u_1 = \frac{u_1' + v}{1 + u_1' v/c^2}$$

or its infinitesimal form

$$du_1 = \frac{(1 - v^2/c^2)du_1'}{(1 + u_1' v/c^2)^2} = \gamma^{-2}\frac{1}{(1 + u_1' v/c^2)^2}du_1'$$

which together with

$$dt = \gamma(1 + u_1' v/c^2)dt'$$

implies the transformation for acceleration

$$a_1 = \frac{du_1}{dt} = \gamma^{-3}\frac{1}{(1+u_1'v/c^2)^3}a_1'$$

Similarly, we have

$$\begin{aligned} a_2 &= \frac{1}{\gamma^2(1+u_1'v/c^2)}a_2' - \frac{vu_2'}{c^2\gamma^2(1+u_1'v/c^2)^3}a_1' \\ a_3 &= \frac{1}{\gamma^2(1+u_1'v/c^2)}a_3' - \frac{vu_3'}{c^2\gamma^2(1+u_1'v/c^2)^3}a_1' \end{aligned}$$

Based on above, it's now clear that acceleration is an absolute concept in SR since

$$\vec{a} \neq 0 \quad \text{in } S \Longleftrightarrow \vec{a}' \neq 0 \quad \text{in } S'$$

However, it is not true in general relativity.

In summary:

Table 4.1: Relative concept in different theories

theory	$\vec{x}$	t	$\vec{v}$	$\vec{a}$
Newtonian	R	A	R	A
SR	R	R	R	A
GR	R	R	R	R

where R = relative and A = absolute.

CHAPTER 5

The Key Attributes of SR III

5.1 Examples

Example. (uniform proper acceleration)

(1) In Newtonian theory, time t is absolute and thus uniform acceleration means $du/dt = a$ for all inertial frame. Then

$$u = u_0 + at, \quad u \to \infty \quad \text{as} \quad t \to \infty$$

which is impossible in SR.

(2) In special relativity, the uniform acceleration means that at any instant, the acceleration a of the particle in the co-moving frame S' is the same, i.e. at t, $v = u(t)$, $u'(t') = 0$ and $du'/dt' = a$. For instance, a rocket whose propeller has a constant emission rate would be uniformly accelerated.

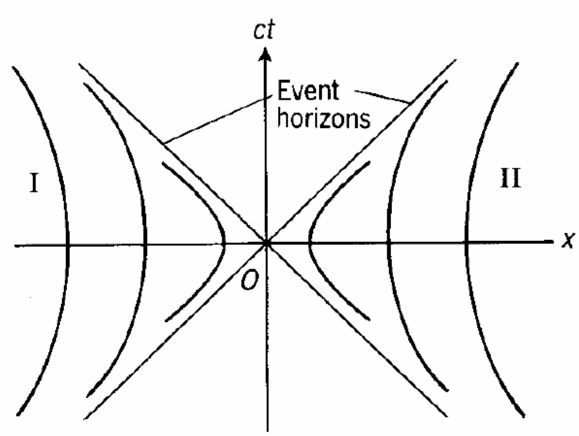

Figure 5.1: World lines of uniform proper acceleration motion

By Lorentz transformation for acceleration, we have

$$\frac{du}{dt} = \frac{1}{\gamma^3(1 + u'v/c^2)^3}\frac{du'}{dt'} = (1 - u^2/c^2)^{3/2}a < a$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$. Solving $u(t)$ we get

$$u(t) = \frac{dx(t)}{dt} = \frac{a(t - t_0)}{\sqrt{1 + a^2(t - t_0)^2/c^2}} < c, \quad (t \to \infty, \quad u \to c)$$

[0]GR5.tex

which implies c is the upper bound of $u(t)$. Integrating once more, we obtain

$$x - x_0 = \frac{c}{a}(c^2 + a^2(t - t_0)^2)^{1/2} - \frac{c^2}{a}$$

or

$$\frac{(x - x_0 + c^2/a)^2}{(c^2/a)^2} - \frac{(ct - ct_0)^2}{(c^2/a)^2} = 1$$

Choosing $x_0 - c^2/a = t_0 = 0$ we get a family of hyperbolae for different proper acceleration a. The more smaller of a, the more flatter of the hyperbola.

Remark.

Since each point on a hyperbola emits light along the 45 degree, hence region I cannot communicate with region II due to event horizon which plays as barrier of information.

Example. (The Doppler effect)

(1) Classical Doppler effect: $(v << c)$

Let S and S' be two inertial frames with relative speed $u_r = \vec{v} \cdot \hat{r}$ where $\vec{v}$ is the relative velocity of S and S' and $\hat{r}$ is the direction $\vec{O'} - \vec{O}$.

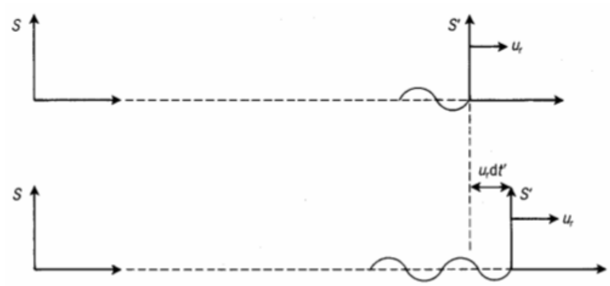

Figure 5.2: Classical Doppler effect

Since time is absolute in Newtonian theory, i.e. $dt = dt'$, if a time period for a light in S' is dt' then the time period observed in S is $\Delta t = dt' + u_r dt'/c$ where the extra term coming from duration traveled by light. In terms of wave length of light, we have

$$\lambda_0 = cdt' \quad (S') \quad \lambda = c\Delta t = cdt'(1 + u_r/c)$$

which implies

$$\frac{\lambda}{\lambda_0} = 1 + \frac{u_r}{c} \begin{cases} > 1 & u_r > 0 \\ < 1 & u_r < 0 \end{cases}$$

This is the so-called classical Doppler effect.

(2) Relativistic Doppler effect:

Due to time dilation, $dt = \gamma dt' > dt'$, the time interval observed in S now becomes

$$\Delta t = \gamma dt' + u_r \gamma dt'/c = \gamma dt'(1 + u_r/c)$$

where $\beta = 1/\sqrt{1 - v^2/c}$ and $u_r = \vec{v} \cdot \hat{r}$. Then

$$\frac{\lambda}{\lambda_0} = \frac{1 + u_r/c}{\sqrt{1 - v^2/c^2}}$$

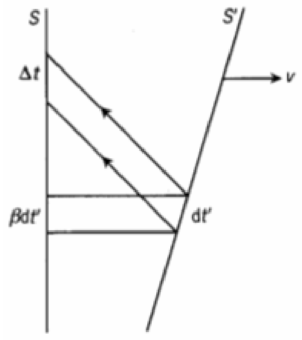

Figure 5.3: Special relativistic Doppler effect ($\beta = 1/\sqrt{1 - v^2/c^2}$)

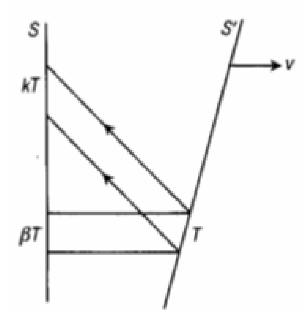

Figure 5.4: Radial Doppler effect ($u_r = v$)

Remark.

(1) If $u_r = v$, (i.e. the source's motion is purely radial) then

$$\frac{\lambda}{\lambda_0} = \sqrt{\frac{1 + v/c}{1 - v/c}} = k\text{-factor}$$

called radial Doppler effect.

(2) If $\vec{v} \perp \hat{r}$, then $u_r = 0$ and

$$\frac{\lambda}{\lambda_0} = \frac{1}{\sqrt{1 - v^2/c^2}}$$

called transverse Doppler effect.

(3) There is no transverse effect in classical Doppler effect.

(4) Mössbauer effect has been confirmed for transverse Doppler effect and hence a verification of time dilation effect.

Example. (Star war problem)

Consider two rockets O and O' with the same length L. They move to each other in opposite direction. The head and bottom coordinates of the rocket O are a and b respectively and b' and a' for O'.

Question: From rocket O point of view, when its head a meets a' (the bottom of O'), b shots a bullet simultaneously to O'. Does the rocket O' survive or not?

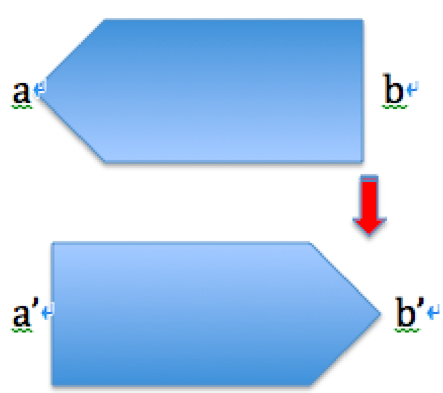

Figure 5.5: Two rockets in relative motion

ANS: From O point of view, the space-time coordinates of both rockets are

$$a_O = (0, 0), \quad a'_O = (0, 0), \quad b_O = (0, L), \quad b'_O = (0, L\sqrt{1 - v^2/c^2})$$

Hence, due to length contraction, O' is safe. Now, let's see what happen if we think of this problem from O' point of view. When a' meets a we have

$$a_{O'} = (0, 0), \quad a'_{O'} = (0, 0)$$

By Lorentz transformation $a_O = (0, 0)$ and by assumption $b_O = (0, L)$. Then, using Lorentz transformation

$$\begin{pmatrix} t'_b \\ x'_b \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v \\ -\gamma v & \gamma \end{pmatrix} \begin{pmatrix} 0 \\ L \end{pmatrix}, \quad \gamma = \frac{1}{\sqrt{1 - v^2}}$$

we get $b_{O'} = (t, x) = (-v\sqrt{\frac{1+v}{1-v}}L, \sqrt{\frac{1+v}{1-v}}L)$. Hence, for O', the bullet at b (in fact, outside of O') shotted earlier than the time when a and a' meet. So O' is safe as well.

Remark. If O' observes the length of O simultaneously, then the length contraction still holds.

Example. (Twin paradox) Suppose M and N are 21 years old twin brother. If N takes a rocket of speed $24c/25$ to make a space traveling and back to the earth after 14 years. When they meet each other again, who is younger?

Let's think of this problem naively. From N point of view, who is $21 + 14 = 35$ years old. From M point of view, due to time dilation, M is $21 + 14/\sqrt{1 - (24/25)^2} = 21 + 50 = 71$ years old. So M said: by time dilation N is younger than me. However, from N point of view, M is moving away from me, so M should be younger than me. Since they do not agree with each other, this is the so-called twin paradox problem.

Let's think of this problem more seriously.

(1) When N moves to point $T = (x', t') = (0, 7)$, then by L.T. $x = \gamma(x' + vt'), \quad t = \gamma(t' + vx'/c^2)$ we get $(x, t) = (24c, 25)$. Hence, from M point of view, $M = 21 + 25 \times 2 = 71$ years old.

(2) From N point of view, when N costs 7 years to reach T, M needs $7 \times 7/25 = 1.96$ to reach point A. So when N backs to C, he thinks that M is $21 + 1.96 \times 2 = 24.92$ years old. Indeed it's wrong since N does not stay at the same inertial system S' over the whole traveling journey, he jumps to S'' at T. Then $t'' = \gamma(t - vx/c^2)$ we have $t''_B - t''_C = -7 = (t_B - t_C) \times (25/7)$ which implies $t_B - t_C = -1.96$. That means the coordinate t_B is $50 - 1.96 = 48.04$. Hence, from N point of view, M is $21 + 48.04 + 1.96 = 71$ years old.

Remark. What's going on for M from $t = 25$ to t_B?

(1) The time intervals T_1, T_2, T_3 are for acceleration and T_0 for uniform boost.

(2) When $T_1, T_2, T_3 \to 0$ (no acceleration interval), the age of M predicted by N is less than that by M himself.

(3) The acceleration, however brief, has effect on N not on M.

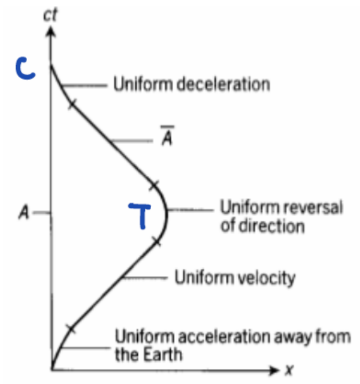

Figure 5.6: Twin paradox

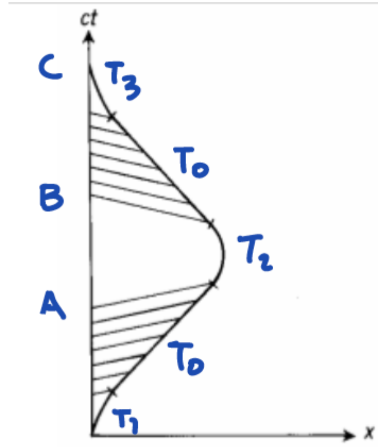

Figure 5.7: Simultaneity lines of N on the outward and return journeys

Summary:

From M point of view: $N_M = 21 + 7 \times 2 = 35$,, $M_M = 21 + 14 \times 25/7 = 71$

From N point of view: $N_N = 21 + 7 \times 2 = 35$, $M_N = 21 + 48.04 + 1.96 = 71$

The point is that M always stays in the same inertial frame, while N changed his inertial frame at turning point.

CHAPTER 6

Relativistic Mechanics

6.1 Newtonian mechanics

From Newton's 2nd law
$$\vec{F} = m\vec{a}$$
where $\vec{F}$ is the external force, m is the inertial mass and $\vec{a}$ is the acceleration which can be viewed as the response to the force $\vec{F}$. In general
$$\vec{F} = \frac{d\vec{p}}{dt}$$
where $\vec{p} = m\vec{v}$ is the linear momentum of the particle. From Newton's 3rd law, we have action-reaction principle
$$\vec{F}_1 = -\vec{F}_2$$

Example. The Newtonian gravitational force acting on m_1 is given by
$$\vec{F} = -\frac{Gm_1m_2}{r^2}\hat{r} = -m_1\nabla\phi$$
where $\vec{r} = \vec{r}_1 - \vec{r}_2$ and the gravitational potential
$$\phi = -\frac{Gm_2}{r}$$
For continuous mass distribution $\rho_i(\vec{r})$, $i = 1, 2$
$$\vec{F} = -G\int \frac{\rho_1(\vec{r})\rho_2(\vec{r}')\hat{\imath}}{|\vec{r}-\vec{r}'|^2}d^3\vec{r}d^3\vec{r}' = \int \rho_1(\vec{r})\vec{U}(\vec{r})d^3\vec{r}$$
where $\vec{U}(\vec{r})$ is the force per unit mass, defined by
$$\vec{U}(\vec{r}) = -G\int \frac{\rho_2(\vec{r}')\hat{\imath}}{|\vec{r}-\vec{r}'|^2}d^3\vec{r}' = -\nabla\phi$$
with scalar potential
$$\phi(\vec{r}) = -G\int \frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}d^3\vec{r}'$$
which satisfies the Poisson equation
$$\nabla^2\phi = G\int \rho(\vec{r}')\nabla^2\frac{1}{|\vec{r}-\vec{r}'|}d^3\vec{r}' = +4\pi G\rho(\vec{r}).$$

[0]GR6.tex

For a system of particles, the 2nd law for each particle reads

$$\vec{F}_i^{ext} = \frac{d\vec{p}_i}{dt}, \quad i = 1, 2, \cdots, n$$

and the total linear momentum is defined by $\vec{P} = \sum_i \vec{p}_i$. If

$$\vec{F}^{ext} = \sum_i \vec{F}_i^{ext} = \frac{d\vec{P}}{dt} = 0$$

then $\vec{P} = \sum_i \vec{p}_i = \text{constant}$ or $\vec{P}(t) = \vec{P}(0)$ for all t.

6.2 Relativistic mass

Recall that, in special relativity, length and time are observer-dependent and thus the mass, being a fundamental property of a particle, is indeed observer-dependent, i.e. $m = m(\vec{u})$ where $\vec{u}$ is the velocity of a particle w.r.t. an inertial frame.

Consider an inelastic scattering:

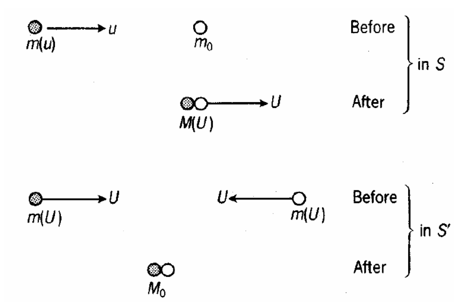

Figure 6.1: Inelastic collision

(1) In S system (Lab system): By mass conservation we have

$$m(u) + m_0 = M(U), \quad m_0 \equiv m(0)$$

and by momentum conservation

$$m(u)u = M(U)U = (m(u) + m_0)U$$

which implies

$$m(u) = \frac{m_0 U}{u - U}.$$

(2) In S' system (Center of mass system): In center of mass system, two particles with equal masses move in opposite direction of speed U. Hence, S' is relative to S by U and additional formula of velocity,

$$u = \frac{U + U}{1 + U^2/c^2}$$

which implies

$$U = \frac{c^2}{u}(1 \pm \sqrt{1 - u^2/c^2}).$$

Since $U \to 0$ as $u \to 0$, only minus sign holds and thus

$$m(u) = \frac{m_0 U}{u - U} = \frac{m_0}{\sqrt{1 - u^2/c^2}} \equiv \gamma m_0$$

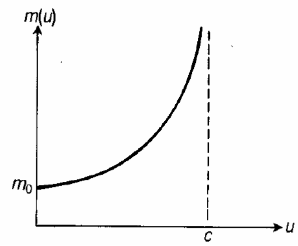

Figure 6.2: Relativistic mass as a function of velocity

6.3 Relativistic energy

How to extract energy (e.g. kinetic energy) from mass formula $m(u) = m_0/\sqrt{1-u^2/c^2}$? Consider low speed limit $u << c$, then

$$m(u) \simeq m_0(1 + \frac{u^2}{2c^2}) + O(u^4/c^4)$$

from which, after multiplying c^2, becomes

$$m(u)c^2 \simeq m_0 c^2 + \frac{1}{2} m_0 u^2 + \text{higher order terms}$$

where the first term can be viewed as the rest mass energy and the second is the kinetic energy. Hence, conservation of relativistic mass reduces to conservation of mass and kinetic energy in Newtonian limit.

Example. Consider the collision of two particles with rest masses m_0 and $\bar{m}_0$ respectively. Before collision, they have speeds v_1 and $\bar{v}_1$, while after collision the speeds become v_2 and $\bar{v}_2$. From conservation of relativistic mass

$$\frac{m_0}{\sqrt{1 - v_1^2/c^2}} + \frac{\bar{m}_0}{\sqrt{1 - \bar{v}_1^2/c^2}} = \frac{m_0}{\sqrt{1 - v_2^2/c^2}} + \frac{\bar{m}_0}{\sqrt{1 - \bar{v}_2^2/c^2}}$$

which, in low speed limit $v_i, \bar{v}_i << c$, becomes

$$\frac{1}{2} m_0 v_1^2 + \frac{1}{2} \bar{m}_0 \bar{v}_1^2 = \frac{1}{2} m_0 v_2^2 + \frac{1}{2} \bar{m}_0 \bar{v}_2^2.$$

This is just the conservation of kinetic energy and hence suggests the relativistic energy $E(u) = m(u)c^2$, i.e. mass and energy are equivalent concepts. In particular, $\Delta E = \Delta m c^2$, i.e. a small mass can produce a large amount of energy!

On the other hand, conservation of relativistic momentum

$$m(v_1)v_1 + \bar{m}(\bar{v}_1)\bar{v}_1 = m(v_2)v_2 + \bar{m}(\bar{v}_2)\bar{v}_2$$

which, in low speed limit, gives

$$m_0 v_1 + \bar{m}_0 \bar{v}_1 = m_0 v_2 + \bar{m}_0 \bar{v}_2.$$

In summary, we have

$$m(u) = \frac{m_0}{\sqrt{1 - u^2/c^2}}, \quad E(u) = \frac{m_0 c^2}{\sqrt{1 - u^2/c^2}}, \quad \vec{p}(u) = \frac{m_0 \vec{u}}{\sqrt{1 - u^2/c^2}},$$

which satisfy

$$\boxed{E^2 = \vec{p}^2 c^2 + m_0^2 c^4}$$

Hence, just likes the 4-coordinate (ct, x, y, z) which satisfies $(ct)^2 - x^2 - y^2 - z^2 - s^2$, here we have energy-momentum 4-vector $p^\mu = (E/c, \vec{p}) = (c(E/c^2), \vec{p})$ which transforms as

$$\begin{cases} \frac{E'}{c^2} &= \gamma(\frac{E}{c^2} - \frac{vp_x}{c^2}) \\ p'_x &= \gamma(p_x - \frac{vE}{c^2}) \\ p'_y &= p_y \\ p'_z &= p_z \end{cases} \overset{v \to -v}{\Longleftrightarrow} \begin{cases} \frac{E}{c^2} &= \gamma(\frac{E'}{c^2} + \frac{vp'_x}{c^2}) \\ p_x &= \gamma(p'_x + \frac{vE'}{c^2}) \\ p_y &= p'_y \\ p_z &= p'_z \end{cases}$$

where v is the relative speed between S and S', and $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$. When $v = u$, $p'_x = 0$, $E' = m_0 c^2$ and thus $E = m(u)c^2$ and $p_x(u) = m_0 u/\sqrt{1 - u^2/c^2}$ as expected.

Example. (Photon). Planck suggested the concept of quanta of light as $E = h\nu$ where h is the Planck's constant and ν is the frequency of light. Since the rest mass of photon is zero and it is impossible to boost to an inertial frame with speed c to observe the rest mass. Hence, $E = pc$ and we can define an effective mass $m_{eff} = E/c^2 = p/c$ for photon. In summary:

$$E = h\nu, \quad \vec{p} = \frac{h\nu}{c}\hat{n}, \quad m_{eff} = \frac{h\nu}{c^2}$$

Example. (Relativistic Doppler effect) We shall show that the Planck's hypothesis can be derived from Doppler effect. Recall $\lambda/\lambda_0 = \sqrt{(1 + v/c)/(1 - v/c)}$. Since $c = \lambda\nu = \lambda_0\nu_0$ we have $\nu_0/\nu = \sqrt{(1 + v/c)/(1 - v/c)}$. However, from Lorentz transformation, we have

$$E = \gamma\left(E_0 - \frac{vE_0}{c}\right) = E_0\sqrt{\frac{1 - v/c}{1 + v/c}} = \frac{E_0\nu}{\nu_0}$$

Hence, $E/\nu = E_0/\nu_0 =$ constant in any inertial frame.

CHAPTER 7

Tensor Algebra I

There are two approaches to tensors

(1) conventional approach based on indices (practically)

(2) index-free based on forms, vector fields, etc. (more abstract)

7.1 Manifold and coordinates

An n-dimensional manifold M is something which locally looks like $\mathbb{R}^n$.

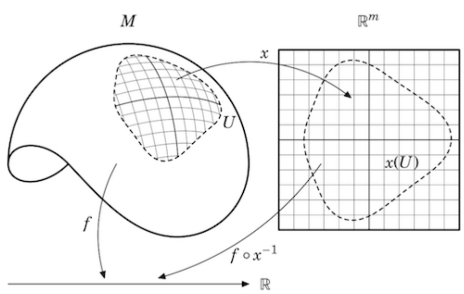

Figure 7.1: A manifold locally likes $\mathbb{R}^n$

Example. A sphere S^2 is a two-dimensional manifold which is locally likes $\mathbb{R}^2$

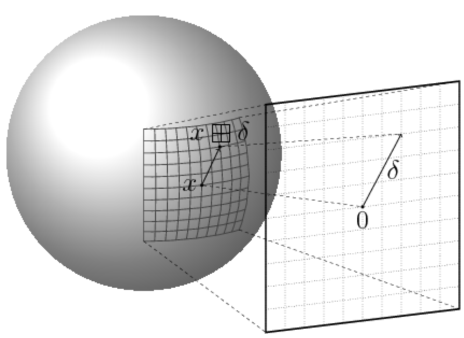

Figure 7.2: S^2 locally likes $\mathbb{R}^2$

[0]GR7.tex

Remark. Note that S^2 is compact however $\mathbb{R}^2$ is non-compact.

Therefore, each point on M can be characterized by a set of n coordinates $(x^1, x^2, \cdots, x^n)$. It might not be possible to cover M by a single non-degenerate coordinate system. Sometimes, it might be possible for a degenerate one.

Example. ($\mathbb{R}^2$) Consider the polar coordinate system on $\mathbb{R}^2$

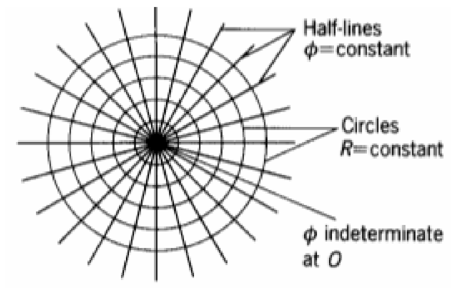

Figure 7.3: Polar coordinate system

In polar coordinate system, ϕ is in-determined at O and thus (r, ϕ) is a degenerate coordinate system, while (x, y) is a non-degenerate one. In general, we have no choice in the matter.

Example. (sphere) Consider a sphere S^2 and (θ, ϕ) is the spherical coordinate system on it.

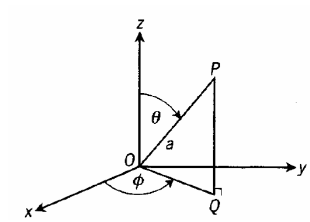

Figure 7.4: Spherical coordinate system (θ : polar angle, ϕ : azimuthal angle)

Since ϕ is in-determined at $\theta = 0, \pi$, we need at least two degenerate systems to cover S^2. One is for Northern hemisphere $(0 \leq \theta < \pi)$, the other for Southern hemisphere $(0 < \theta \leq \pi)$.

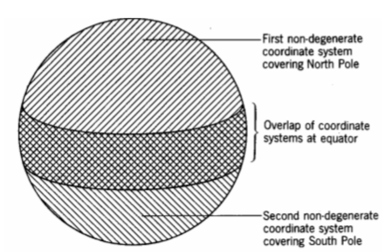

Figure 7.5: Two non-degenerate systems covering S^2

Example. (Stereo-projection) Consider the stereo projection of a unit sphere onto the $z = 0$ plane. Let $P = (x, y, z)$ be a point on the sphere and $P' = (X, Y, 0)$ be the stereo-projection from the North pole $N = (0, 0, 1)$. We may also consider the projection $P'' = (X', Y', 0)$ from the South pole $S = (0, 0, -1)$. All projection are non-singular except the North pole N and the south pole S. Show that

$$(X, Y) = (\frac{x}{1-z}, \frac{y}{1-z}), \quad (X', Y') = (\frac{x}{1+z}, \frac{y}{1+z})$$

and

$$X' = \frac{X}{X^2+Y^2}, \quad Y' = \frac{Y}{X^2+Y^2}$$

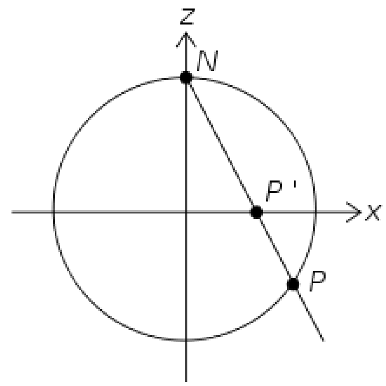

Figure 7.6: Stereographic projection from the North pole

Thus we work with coordinate systems which cover only a portion of the manifold, which we call coordinate patches. In general, we can use a set of patches to cover a manifold M, i.e.

$$\{\text{patch } 1, \text{patch } 2, \cdots\} = \text{atlas}$$

The theory of manifold tells us how to get from one coordinate patch to another by a coordinate transformation in the overlapping region.

Remark. Tensors are geometric quantities defined on a manifold with respect to coordinate transformation.

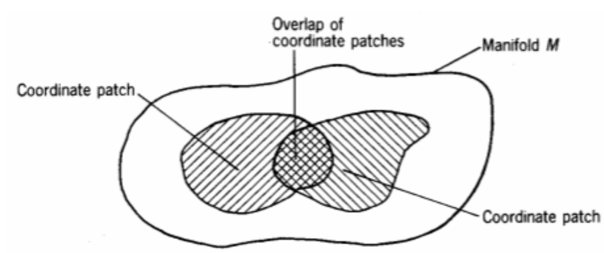

Figure 7.7: Overlapping coordinate patch in a manifold

7.2 Curves and surfaces

A curve on M is given by

$$x^a = x^a(u), \quad a = 1, 2, \cdots, n$$

where $n = \dim M$ and u is a parameter. An m-dimensional surface in M is parametrized by

$$x^a = x^a(u^1, u^2, \cdots, u^m).$$

If $m = n - 1$ such a surface is called a hypersurface, i.e.

$$x^a = x^a(u^1, u^2, \cdots, u^{n-1}), \quad a = 1, 2, \cdots, n$$

For instance, $S^2 \subset \mathbb{R}^3$. Eliminating $u^1, u^2, \cdots, u^{n-1}$ from these n equation gives

$$g(x^1, x^2, \cdots, x^n) = 0.$$

In general, an m-dimensional surface has m degrees of freedom due to $(n-m)$ constraints

$$g^1(x^1, x^2, \cdots, x^n) = 0, \cdots, g^{n-m}(x^1, x^2, \cdots, x^n) = 0$$

7.3 Transformations of coordinates

A point on M can be covered by many different coordinate patches. When we make a statement about tensors for one coordinate system, we also wish it to hold for others. This motivates us to study transformations between different coordinates for a tensor. Consider a change of coordinates (passively)

$$x^a \to x'^a = f^a(x^1, x^2, \cdots, x^n) \equiv x'^a(x), \quad a = 1, 2, \cdots, n$$

Define the transformation matrix $J'^a_b = \partial x'^a / \partial x^b$ and its Jacobian as

$$J' = \left| \frac{\partial x'^a}{\partial x^b} \right| = \det(J'^a_b).$$

By the implicit function theorem, we have $x^a = x^a(x')$ and define $J = |\partial x^a / \partial x'^b|$. Obviously, we have $J = 1/J'$. Indeed, from chain rule of differentiation

$$\frac{\partial x'^a}{\partial x^b} \frac{\partial x^b}{\partial x'^d} = J'^a_b J^b_d = \delta^a_d$$

where δ^a_d is the Kronecker delta symbol defined by $\delta^a_d = 1$ for $a = d$ and $\delta^a_d = 0$ for $a \neq d$. Hence, $\det(J'^a_b J^b_d) = J' J = 1$. The transformation matrix $\partial x'^a / \partial x^b$ can be viewed from infinitesimal differentials. For instance, let $z = f(x, y)$ be a hypersurface in $\mathbb{R}^3$ then

$$dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

In general, from $x'^a = x'^a(x)$ we have

$$dx'^a = \frac{\partial x'^a}{\partial x^1} dx^1 + \cdots + \frac{\partial x'^a}{\partial x^n} dx^n = \sum_{b=1}^{n} \frac{\partial x'^a}{\partial x^b} dx^b.$$

Remark. We shall use Einstein convention to rewrite $\sum_{b=1}^n \frac{\partial x'^a}{\partial x^b} dx^b$ as $\frac{\partial x'^a}{\partial x^b} dx^b$, i.e. sum over repeated indices.

CHAPTER 8

Tensor Algebra II

8.1 Contravariant tensors

From now on, we would like to define geometric quantities according to its transformation properties. Let P and Q be two neighboring points on the manifold M with coordinate x^a and $x^a + dx^a$, respectively. Here, dx^a represents the infinitesimal displacement vector from P to Q in x-coordinate. In another coordinate x', the corresponding displacement becomes

$$dx'^a = \left[\frac{\partial x'^a}{\partial x^b}\right]_P dx^b$$

where the subscript in $[]_P$ means that the transformation matrix is evaluated at point P.

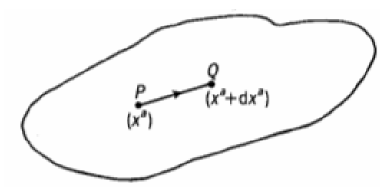

Figure 8.1: The infinitesimal vector attached to P

Definition 8.1.1. *A contravariant vector (rank 1) is a set of quantities X^a defined on $(x^a)_P$ such that under a coordinate transformation $x^a \to x'^a(x)$,*

$$X'^a = \frac{\partial x'^a}{\partial x^b} X^b$$

Example.
(1) $X^a = dx^a$ (infinitesimal displacement)
(2) $X^a =$ tangent vector of a curve $x^a(u)$.

[0]GR8.tex

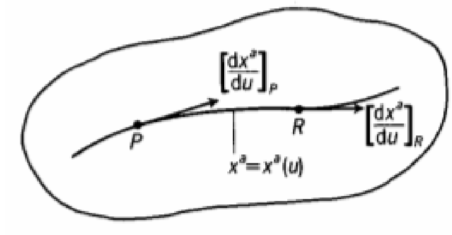

Figure 8.2: Tangent vector on a curve

For any function $f(x(u))$ on the curve we have $df/du = (\partial f/\partial x^a)(dx^a/du)$ or dropping f,

$$\frac{d}{du} = \frac{\partial x^a}{\partial u}\frac{\partial}{\partial x^a} \equiv X^a \frac{\partial}{\partial x^a}$$

where $\partial/\partial x^a$ and X^a can be viewed as basis and component of the tangent vector respectively. In x'-system, $X'^a = dx'^a/du$ which is related to X^a by above linear homogeneous transformation. The higher rank generalization, e.g. contravariant tensor with rank 2, is given by

$$X'^{ab} = \frac{\partial x'^a}{\partial x^c}\frac{\partial x'^b}{\partial x^d} X^{cd}$$

Remark. A scalar ϕ (without indices) is called rank-0 tensor which satisfies

$$\phi'(x') = \phi(x).$$

8.2 Covariant and mixed tensors

Let $\phi(x^a)$ be a real-valued, differentiable scalar function on the manifold M, i.e. $\partial\phi/\partial x^a$ is well-defined. Since $x^a = x^a(x')$ we have

$$\frac{\partial \phi}{\partial x'^b} = \frac{\partial x^a}{\partial x'^b}\frac{\partial \phi}{\partial x^a}$$

Definition 8.2.1. *A covariant vector (rank 1) is a set of quantities X_a defined on $(x^a)_P$ such that under coordinate transformation*

$$X'_a = \frac{\partial x^b}{\partial x'^a} X_b$$

The higher rank covariant tensors can be defined in a similar way as

$$X'_{ab} = \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b} X_{cd}$$

for rank-2 tensor of $(0,2)$ type and

$$X'^a_{bc} = \frac{\partial x'^a}{\partial x^d}\frac{\partial x^e}{\partial x'^b}\frac{\partial x^f}{\partial x'^c} X^d_{ef}$$

for mixed tensor of $(1,2)$ type. In general, we denote a mixed tensor with contravariant rank p and covariant rank q as (p,q) type of the form $X^{a_1\cdots a_p}_{b_1\cdots b_q}$.

In summary, physical laws described by tensor equations are coordinate-independent. For instance, consider a rank-2 equation in x-system as

$$X_{ab} = Y_{ab}$$

where both hand sides are covariant rank-2 tensors. Then, in x'-system

$$X'_{ab} = \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b} X_{cd} = \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b} Y_{cd} = Y'_{ab}$$

Therefore, if we work with tensor equations, they hold in all coordinate systems.

8.3 Tensor fields

Definition 8.3.1. *Over some region in M, we can associate every point a vector, which forms a vector field. Similarly, For every point of the region, we can assign $X^{a\cdots}_{b\cdots}$ for a tensor field. Usually, we assume that $X^{a\cdots}_{b\cdots}$ is differentiable (C^∞) over the region. At every point P of the region, the tensor fields, say contravariant vector field, satisfies the transformation law*

$$X'^a(x') = \left[\frac{\partial x'^a}{\partial x^b}\right]_P X^b(x).$$

From now on, we denote a tensor field as $X^{a\cdots}_{b\cdots}(x)$ and a tensor as $X^{a\cdots}_{b\cdots}(P)$.

8.4 Elementary operations with tensors

In this section we shall discuss some operations on tensors without changing the properties of tensors.

(1) Addition: Given tensors X^a_{bc}, Y^a_{bc}, and Z^a_{bc} of the same type we can define addition of them as

$$X^a_{bc} = Y^a_{bc} + Z^a_{bc}$$

(2) Scalar multiplication: Given a tensor X^a_{bc} we can define scalar multiplication as

$$kX^a_{bc}, \quad k \in \mathbb{R}$$

(3) Symmetrization: A tensor X_{ab} is called symmetric tensor if $X_{ab} = X_{ba}$ which has $(n^2-n)/2+n = n(n+1)/2$ independent components. A tensor X_{ab} is called anti-symmetric (or skew-symmetric) tensor if $X_{ab} = -X_{ba}$ which has $(n^2-n)/2 = n(n-1)/2$ independent components. Given an arbitrary tensor X_{ab}, the symmetric and anti-symmetric parts can be extracted by

$$\begin{aligned} X_{(a,b)} &= \frac{1}{2}(X_{ab} + X_{ba}) \quad \text{(symmetric)} \\ X_{[a,b]} &= \frac{1}{2}(X_{ab} - X_{ba}) \quad \text{(anti-symmetric)} \end{aligned}$$

Hence, $X_{(a,b)} = X_{(b,a)}$, $X_{[a,b]} = -X_{[b,a]}$ and $X_{a,b} = X_{(a,b)} + X_{[a,b]}$. In general, for higher rank tensors,

$$\begin{aligned} X_{(a_1,\cdots,a_r)} &= \frac{1}{r!}(\text{sum over all permutations of } a_i) \quad \text{(symmetric)} \\ X_{[a_1,\cdots,a_r]} &= \frac{1}{r!}(\text{alternating sum of all permutations of } a_i) \quad \text{(anti-symmetric)} \end{aligned}$$

Example. For a rank-3 covariant tensor X_{abc}

$$X_{[abc]} = \frac{1}{6}(X_{abc} + X_{bca} + X_{cab} - X_{bac} - X_{cba} - X_{acb})$$

(4) Tensor product: We can construct higher rank tensors from the lower ones by tensor product, e.g., given $(1,1)$ type tensor Y^a_b and $(0,2)$ type tensor Z_{cd} we can construct a $(1,3)$ type tensor as

$$X^a_{bcd} = Y^a_b Z_{cd} \qquad (1,1)\times(0,2) = (1,3)$$

In general, we have $(p_1,q_1)\times(p_2,q_2) = (p_1+p_2, q_1+q_2)$.

(5) Tensor contraction: We can reduce the rank of a tensor by contracting some indices of the tensor, e.g.

$$X^a_{bcd} \xrightarrow{contraction} X^a_{acd} \equiv Y_{cd}$$

where the index a has been contracted. The $(1,3)$ tensor, after contracting, then becomes a $(0,2)$ tensor.

8.5 Index-free interpretation of contravariant vector fields

A more geometric interpretation of vector fields is the following. We can view (contravariant) vector field as an operator acting on real-valued function, i.e. $Xf = g$ in x-systm where

$$X \equiv X^a \partial_a, \quad \partial_a \equiv \frac{\partial}{\partial x^a}$$

and $Xf = X^a \partial_a f$. The expression X is indeed coordinate-independent since

$$X'^a \partial'_a = X'^a \frac{\partial}{\partial x'^a} = \frac{\partial x'^a}{\partial x^b} X^b \frac{\partial x^c}{\partial x'^a} \frac{\partial}{\partial x^c} = \delta^c_b X^b \partial_c = X^c \partial_c = X$$

Hence, $[\frac{\partial}{\partial x^a}]_P$ is a basis for all tangent vectors at P and any vector X at P can be expressed as

$$X = [X^a]_P \left[\frac{\partial}{\partial x^a}\right]_P$$

All vectors at P form a tangent plane $T_P(M)$.

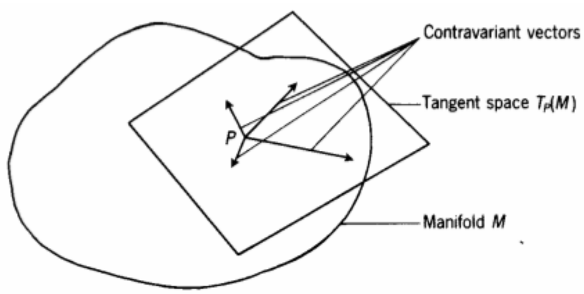

Figure 8.3: The tangent plane $T_p(M)$ at P

Remark. Note that, in general, contravariant vectors X lie on T_P, not on the manifold M. (Exception: If $M = \mathbb{R}^n$, then $T_P(M) = M$)

Given any two vector fields X and Y we can construct a new vector field via commutator or Lie bracket of X and Y by

$$[X, Y] \equiv XY - YX.$$

Let's verify this assertion. Denote $Z = [X, Y]$, then acting on a differential function, we get

$$\begin{aligned}
Zf &= [X, Y]f = XYf - YXf \\
&= XY^a\partial_a f - YX^a\partial_a f \\
&= X^b\partial_b Y^a\partial_a f + X^bY^a\partial_b\partial_a f - Y^b\partial_b X^a\partial_a f - Y^bX^a\partial_b\partial_a f \\
&= [X^b\partial_b Y^a - Y^b\partial_b X^a]\partial_a f \\
&= Z^a\partial_a f.
\end{aligned}$$

where $Z^a = X^b\partial_b Y^a - Y^b\partial_b X^a$.

Example. Show that Z^a is a contravariant vector.

The Lie bracket defined above satisfies the following properties:

(1) Anti-symmetry: $[X, Y] = -[Y, X]$

(2) Leibnitz rule: $[XY, Z] = X[Y, Z] + [X, Z]Y$

(3) Jacobi identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

CHAPTER 9

Tensor Calculus I

9.1 Partial derivative of a tensor

Due to the fact that a manifold can not be covered by a single coordinate system and thus the definition of differentiation of tensors becomes a non-trivial problem. What differential operations are there that are tensorial? Consider ordinary partial derivative of a vector X^a, then

$$\partial'_c X'^a = \partial'_c(\frac{\partial x'^a}{\partial x^b}X^b) = \frac{\partial x'^a}{\partial x^b}\partial'_c X^b + \frac{\partial^2 x'^a}{\partial x'^c \partial x^b}X^b = (\frac{\partial x'^a}{\partial x^b})(\frac{\partial x^d}{\partial x'^c})\partial_d X^b + \frac{\partial^2 x'^a}{\partial x'^c \partial x^b}X^b$$

where the first term looks like a transformation of a $(1,1)$ type tensor, except that the second term is an extra term. The reason is the following. From the definition of derivative at P

$$\lim_{\delta u \to 0} \frac{[X^a]_P - [X^a]_Q}{\delta u}.$$

However,

$$X'^a_P = (\frac{\partial x'^a}{\partial x^b})_P X^b_P, \quad X'^a_Q = (\frac{\partial x'^a}{\partial x^b})_Q X^b_Q$$

and thus the transformation matrix for X^a_P and X^a_Q are evaluated at different points. This implies that the difference $X^a_P - X^a_Q$ is not a tensor.

To construct tensorial differentiation, we have to introduce some auxiliary fields onto manifold. We focus the following constructions:

(1) introducing a contravariant vector field to define Lie derivative

(2) introducing an affine connection to define covariant differentiation

(3) introducing a metric tensor to construct a special affine connection called metric connection which also defines a covariant differentiation.

9.2 Lie derivative

In the following, the coordinate transformation will be performed actively rather than passively.

Let's begin to discuss the Lie derivative.

(1) A congruence of curves define tangent vectors for a curve $x^a(u)$ by $dx^a(u)/du$. If we do so for every points on M, we get a vector field X^a.

(2) Conversely, given a vector field $X^a(x)$ over M, we can define a curve or orbit by $dx^a(u)/du = X^a(x(u))$. By uniqueness and existence of ODE, we obtain orbits $x^a(u)$.

[0]GR9.tex

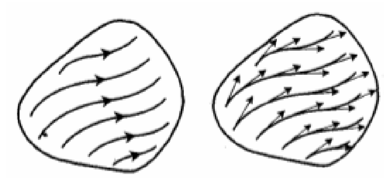

Figure 9.1: Congruence of curves define a vector field

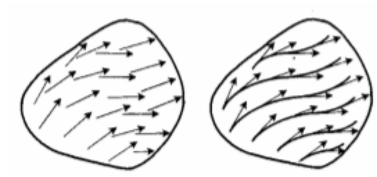

Figure 9.2: A vector field define a congruence of curves

Now, consider a differentiation of a tensor field $T^{a\cdots}_{b\cdots}(x)$ along the direction X^a. Let P and Q be two neighboring points on the curve $x^a(u)$ with coordinates x^a and $x'^a = x^a + \delta u X^a$, respectively. Note that here the coordinate of each point are given in the same x-system. Hence

$$\frac{\partial x'^a}{\partial x^b} = \delta^a_b + \delta u \partial_b X^a$$

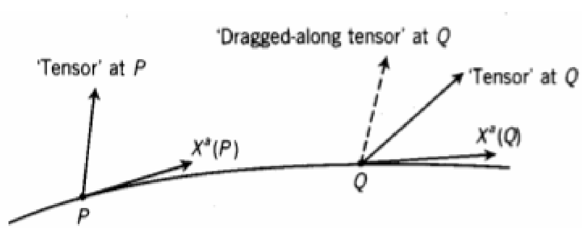

Figure 9.3: Lie derivative of a tensor along X direction

Example. $(2,0)$ tensor T^{ab}. Consider coordinate transformation from P to Q generated by above which induces a drag-along tensor at Q as

$$\begin{aligned} T^{ab}(x) \to T'^{ab}(x') &= (\frac{\partial x'^a}{\partial x^c})(\frac{\partial x'^b}{\partial x^d})T^{cd}(x) \\ &= [\delta^a_c + \delta u \partial_c X^a][\delta^b_d + \delta u \partial_d X^b]T^{cd}(x) \\ &= T^{ab}(x) + \delta u[\partial_c X^a T^{cb} + \partial_d X^b T^{ad}] + O(\delta u^2). \end{aligned}$$

On the other hand, the tensor located at Q is given by

$$T^{ab}(x') = T^{ab}(x^c + \delta u X^c) = T^{ab}(x) + \delta u X^c \partial_c T^{ab} + O(\delta u^2)$$

The Lie derivative of tensor field T^{ab} along X is defined by

$$\begin{aligned} L_X T^{ab} &\equiv \lim_{\delta u \to 0} \frac{T^{ab}(x') - T'^{ab}(x')}{\delta u} \\ &= X^c \partial_c T^{ab} - T^{ac}\partial_c X^b - T^{cb}\partial_c X^a \end{aligned}$$

Example. Check that $L_X T^{ab}$ is also a $(2,0)$ tensor.

Remark. In the following the notation $\overset{*}{=}$ means that equation holds only for a particular coordinate system, e.g. for $X^a \overset{*}{=} \delta_1^a = (1, 0, \cdots)$ then $L_X T^{ab} = \partial_1 T^{ab}$ which looks like the ordinary differentiation.

Lie derivative satisfies the following properties:

(1) It is linear:

$$L_X(\lambda Y^a + \mu Z^a) = \lambda L_X Y^a + \mu L_X Z^a$$

(2) It is Leibniz:

$$L_X(Y^a Z_{bc}) = (L_X Y^a) Z_{bc} + Y^a (L_X Z_{bc})$$

(3) It is type-preserving: If T is a (p, q) type tensor, then $L_X T = S$ is also a (p, q) type tensor.

(4) It commutes with contraction:

$$\delta_a^b L_X T_b^a = L_X T_a^a$$

(5) Given a scalar ϕ, then

$$L_X \phi = X\phi = X^a \partial_a \phi$$

which can be traced back to the definition that

$$L_X \phi = \lim_{\delta u \to 0} \frac{\phi(x') - \phi'(x')}{\delta u} = \lim_{\delta u \to 0} \frac{\phi(x') - \phi(x)}{\delta u} = X^a \partial_a \phi.$$

(6) Lie derivative on a contravariant vector Y^a is given by

$$L_X Y^a = [X, Y]^a = X^b \partial_b Y^a - Y^b \partial_b X^a$$

(7) Lie derivative on a covariant vector Y_a is given by

$$L_X Y_a = X^b \partial_b Y_a + Y_b \partial_a X^b$$

(8) In general, for a general tensor field $T_{b\cdots}^{a\cdots}$, its Lie derivative is given by

$$\begin{aligned} L_X T_{b\cdots}^{a\cdots} &= X^c \partial_c T_{b\cdots}^{a\cdots} - T_{b\cdots}^{c\cdots} \partial_c X^a - \cdots \\ &\quad + T_{c\cdots}^{a\cdots} \partial_b X^c + \cdots \end{aligned}$$

Example. Check (6), (7) from the definition.

9.3 Affine connection and covariant differentiation

Now, we would like to define another tensorial differentiation which does not rely on a curve. Consider a contravariant vector X^a on M The vector field evaluated at P and Q are $X^a(x)$ and $X^a(x + \delta x) = X^a(x) + \delta X^a$, respectively. If we take the difference of them as

$$X^a(x + \delta x) - X^a(x) = \delta x^b \partial_b X^a \equiv \delta X^a$$

then it is easy to show that δX^a is not a vector. The problem is the same as before that we subtract two vectors located at different points.

Introducing a "parallel" vector at Q as

$$X_{\parallel}^a(x + \delta x) = X^a(x) + \bar{\delta} X^a$$

Since $\bar{\delta} X^a = 0$ for $\delta x^a = 0$ or $X^a = 0$ and assume $\bar{\delta} X^a$ is linear in X^a and δx^a, we may write

$$\bar{\delta} X^a = -\Gamma_{bc}^a(x) X^b(x) \delta x^c$$

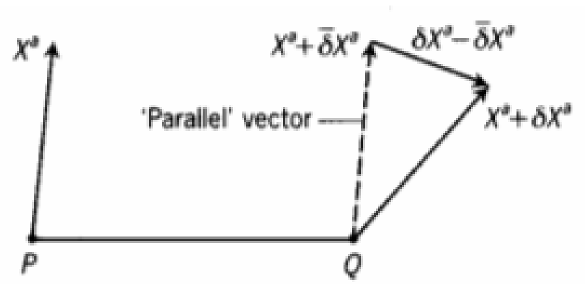

Figure 9.4: Parallel transport of a vector X^a

where $\Gamma^a_{bc}(x)$ contains n^3 functions called affine connection.

Remark. For flat space, $\bar{\delta}X^a = 0$.

Now, define the covariant derivative as

$$\begin{aligned}
\nabla_c X^a &= \lim_{\delta x^c \to 0} \frac{\delta X^a - \bar{\delta} X^a}{\delta x^c} \\
&= \lim_{\delta x^c \to 0} \frac{\delta x^c \partial_c X^a + \Gamma^a_{bc}(x) X^b(x) \delta x^c}{\delta x^c} \\
&= \partial_c X^a + \Gamma^a_{bc}(x) X^b(x)
\end{aligned}$$

Since we claim that $\nabla_c X^a$ is a $(1,1)$ type tensor, i.e.

$$\nabla'_c X'^a = (\frac{\partial x^b}{\partial x'^c})(\frac{\partial x'^a}{\partial x^d})\nabla_b X^d$$

which will determine the transformation rule for the affine connection Γ^a_{bc}.

CHAPTER 10

Tensor Calculus II

10.1 Affine connection (cont'd)

Since we claim that $\nabla_c X^a$ is a $(1,1)$ type tensor, i.e.

$$\nabla'_c X'^a = (\frac{\partial x^b}{\partial x'^c})(\frac{\partial x'^a}{\partial x^d})\nabla_b X^d$$

which will determine the transformation rule for the affine connection Γ^a_{bc}. From left-hand side we have

$$\begin{aligned}
\nabla'_c X'^a &= \partial'_c X'^a + \Gamma'^a_{bc} X'^b \\
&= \partial'_c(\frac{\partial x'^a}{\partial x^b} X^b) + \Gamma'^a_{bc}\frac{\partial x'^b}{\partial x^d} X^d \\
&= \partial'_c(\frac{\partial x'^a}{\partial x^b}) X^b + \frac{\partial x'^a}{\partial x^b}\partial'_c X^b + \Gamma'^a_{bc}\frac{\partial x'^b}{\partial x^d} X^d
\end{aligned}$$

The right-hand side gives

$$(\frac{\partial x^b}{\partial x'^c})(\frac{\partial x'^a}{\partial x^d})\nabla_b X^d = (\frac{\partial x^b}{\partial x'^c})(\frac{\partial x'^a}{\partial x^d})(\partial_b X^d + \Gamma^d_{eb} X^e)$$

Hence

$$\Gamma'^a_{bc}\frac{\partial x'^b}{\partial x^e} = (\frac{\partial x^b}{\partial x'^c})(\frac{\partial x'^a}{\partial x^d})\Gamma^d_{eb} - \partial'_c(\frac{\partial x'^a}{\partial x^e})$$

or

$$\Gamma'^a_{bc} = (\frac{\partial x'^a}{\partial x^d})(\frac{\partial x^e}{\partial x'^b})(\frac{\partial x^f}{\partial x'^c})\Gamma^d_{ef} - \frac{\partial x^e}{\partial x'^b}\partial'_c(\frac{\partial x'^a}{\partial x^e}).$$

Since the last term can be rewritten as

$$\frac{\partial x^e}{\partial x'^b}\partial'_c(\frac{\partial x'^a}{\partial x^e}) = \partial'_c(\frac{\partial x^e}{\partial x'^b}\frac{\partial x'^a}{\partial x^e}) - \frac{\partial^2 x^e}{\partial x'^c \partial x'^b}\frac{\partial x'^a}{\partial x^e} = -\frac{\partial^2 x^e}{\partial x'^c \partial x'^b}\frac{\partial x'^a}{\partial x^e}$$

we have

$$\boxed{\Gamma'^a_{bc} = (\frac{\partial x'^a}{\partial x^d})(\frac{\partial x^e}{\partial x'^b})(\frac{\partial x^f}{\partial x'^c})\Gamma^d_{ef} - \frac{\partial^2 x^e}{\partial x'^c \partial x'^b}\frac{\partial x'^a}{\partial x^e}}$$

Remark.

(1) Γ^a_{bc} is not a $(1,2)$ type tensor due to the inhomogeneous term. We call it affine connection or affinety.

(2) A manifold with a connection prescribed on it is called an affine manifold.

[0] GR10.tex

(3) If ϕ is a scalar, then $\nabla_c\phi = \partial_c\phi$.

We demand that covariant derivative is Leibniz, i.e.

$$\begin{aligned}
\nabla_c(X_a X^a) &= (\nabla_c X_a)X^a + X_a(\nabla_c X^a) \\
&= (\nabla_c X_a)X^a + X_a(\cancel{\partial_c X^a} + \Gamma^a_{bc}X^b) \\
&= (\partial_c X_a)X^a + X_a(\cancel{\partial_c X^a})
\end{aligned}$$

which implies covariant derivative for covariant vector

$$\nabla_c X_a = \partial_c X_a - \Gamma^b_{ac}X_b$$

which is a $(0,2)$ type tensor. Thus for a (p,q) type tensor, the covariant derivative brings it to a $(p,q+1)$ type tensor. In general, we have

$$\begin{aligned}
\nabla_c T^{a\cdots}_{b\cdots} = \partial_c T^{a\cdots}_{b\cdots} &+ \Gamma^a_{dc}T^{d\cdots}_{b\cdots} + \cdots \\
&- \Gamma^d_{bc}T^{a\cdots}_{d\cdots} - \cdots
\end{aligned}$$

Remark. Although Γ^a_{bc} is not a $(1,2)$ tensor, but $\Gamma^a_{bc} - \Gamma^a_{cb}$ is since their inhomogeneous terms are canceled. We call $T^a_{bc} \equiv \Gamma^a_{bc} - \Gamma^a_{cb}$ the torsion tensor which is the anti-symmetric part of Γ^a_{bc}.

Example. (torsion free) For a torsion free case $T^a_{bc} = 0$, or $\Gamma^a_{bc} = \Gamma^a_{cb}$. Then the Lie derivative along X can be expressed as

$$\begin{aligned}
L_X T^{a\cdots}_{b\cdots} &= X^c\partial_c T^{a\cdots}_{b\cdots} - (T^{c\cdots}_{b\cdots}\partial_c X^a + \cdots) + (T^{a\cdots}_{c\cdots}\partial_b X^c + \cdots) \\
&= X^c(\nabla_c T^{a\cdots}_{b\cdots} - \cancel{\Gamma^a_{dc}T^{d\cdots}_{b\cdots}} + \cancel{\Gamma^d_{bc}T^{a\cdots}_{d\cdots}} + \cdots) \\
&\quad -T^{c\cdots}_{b\cdots}(\nabla_c X^a - \cancel{\Gamma^a_{dc}X^d}) - \cdots \\
&\quad +T^{a\cdots}_{c\cdots}(\nabla_b X^c - \cancel{\Gamma^c_{db}X^d}) - \cdots \\
&= X^c\nabla_c T^{a\cdots}_{b\cdots} - (T^{c\cdots}_{b\cdots}\nabla_c X^a + \cdots) + (T^{a\cdots}_{c\cdots}\nabla_b X^c + \cdots)
\end{aligned}$$

where all ordinary derivative ∂_c can be replaced by covariant derivative ∇_c.

Example. For a symmetric connection,

$$L_X Y^a = [X,Y]^a = X^b\partial_b Y^a - Y^b\partial_b X^a = X^b\nabla_b Y^a - Y^b\nabla_b X^a$$

From now on, unless we state, we shall assume that the affine connection is torsion free, i.e. $\Gamma^a_{bc} = \Gamma^a_{cb}$.

10.2 Affine geodesics

Let X be a vector and $\nabla_X T^{a\cdots}_{b\cdots} = X^c\nabla_c T^{a\cdots}_{b\cdots}$ be covariant derivative of the tensor $T^{a\cdots}_{b\cdots}$ along the vector X. If X is a tangent vector of a curve $x^a(u)$, then $X^a = dx^a(u)/du$ and we denote

$$\nabla_X T^{a\cdots}_{b\cdots} \equiv \frac{DT^{a\cdots}_{b\cdots}}{Du}$$

as the absolute derivative of a tensor $T^{a\cdots}_{b\cdots}$ along a curve C which is different from $dT^{a\cdots}_{b\cdots}/du = X^c\partial_c T^{a\cdots}_{b\cdots}$. If the tensor field $T^{a\cdots}_{b\cdots}$ is defined as parallel transport along the curve $C : x^a(u)$, then we have

$$\frac{DT^{a\cdots}_{b\cdots}}{Du} = 0.$$

Given $T^{a\cdots}_{b\cdots}(P)$ we obtain a tensor along the curve C which is everywhere parallel to $T^{a\cdots}_{b\cdots}(P)$.

The so-called affine geodesic is defined by

$$\frac{D}{Du}\left(\frac{dx^a}{du}\right) = \lambda(u)\frac{dx^a}{du} \quad \text{or} \quad \nabla_X X^a = \lambda X^a$$

such that the parallel-transported vector of $X^a(P)$ along the curve C to Q is proportion to the tangent vector $X^a(Q)$, in other word, $X^a_{\parallel}(Q) - X^a(Q) = \lambda X^a(Q)$. Therefore,

$$X^c \nabla_c X^a = X^c(\partial_c X^a + \Gamma^a_{bc} X^b) = \lambda X^a$$

or

$$\frac{d^2 x^a}{du^2} + \Gamma^a_{bc}\frac{dx^b}{du}\frac{dx^c}{du} = \lambda\frac{dx^a}{du}.$$

Remark. Note that only the symmetric part of Γ^a_{bc} contribute the above equation.

If we choose the parameter $u = s$ such that $\lambda(s) = 0$, then $\nabla_X X^a = 0$ or

$$\boxed{\frac{d^2 x^a}{ds^2} + \Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds} = 0}$$

Remark. Obviously, the geodesic equation is invariant under affine transformation $s \to \alpha s + \beta$ for the parameter s.

10.3 The Riemann tensor

For ordinary derivative we have

$$\partial_a \partial_b T = \partial_b \partial_a T.$$

however, for covariant derivative, in general

$$\nabla_c \nabla_d T^{a\cdots}_{b\cdots} \neq \nabla_d \nabla_c T^{a\cdots}_{b\cdots}.$$

where $\nabla_c X^a = \partial_c X^a + \Gamma^a_{bc} X^b$.

Example. Show that,

$$\nabla_c \nabla_d X^a - \nabla_d \nabla_c X^a = R^a{}_{bcd} X^b + (\Gamma^b_{cd} - \Gamma^b_{dc})\nabla_b X^a$$

where $R^a{}_{bcd}$ is the Riemann tensor defined by

$$\boxed{R^a{}_{bcd} \equiv \partial_c \Gamma^a_{bd} - \partial_d \Gamma^a_{bc} + \Gamma^e_{bd}\Gamma^a_{ec} - \Gamma^e_{bc}\Gamma^a_{ed}}$$

For torsion free case, $\Gamma^b_{cd} - \Gamma^b_{dc} = 0$ and

$$\nabla_{[c}\nabla_{d]} X^a \equiv \frac{1}{2}(\nabla_c \nabla_d X^a - \nabla_d \nabla_c X^a) = \frac{1}{2} R^a{}_{bcd} X^b.$$

The Riemann tensor is a $(1,3)$ type tensor which measures the non-commutativity of the covariant derivative. In general,

$$\nabla_{[c}\nabla_{d]} T^{a\cdots}_{b\cdots} \sim \sum RT$$

which suggests that

$$\nabla_{[c}\nabla_{d]} T^{a\cdots}_{b\cdots} = 0 \Longleftrightarrow R^a{}_{bcd} = 0$$

10.4 Geodesic coordinate

So far, the non-flatness effect of a manifold is controlled by the affine connection Γ^a_{bc}. At any point, we can chose a special coordinate called **geodesic coordinate systems** such that

$$[\Gamma^a_{bc}]_P \stackrel{*}{=} 0$$

where P is chosen so that $[x^a]_P \stackrel{*}{=} 0$. Indeed, we can achieve this by following coordinate transformation

$$x^a \to x'^a = x^a + \frac{1}{2} Q^a_{bc} x^b x^c$$

where $Q^a_{bc} = Q^a_{cb}$ are constants and to be determined. Thus

$$\frac{\partial x'^a}{\partial x^d} = \delta^a_d + Q^a_{bd} x^b \quad \Longrightarrow \quad [\frac{\partial x'^a}{\partial x_d}]_P = \delta^a_d$$

and

$$\frac{\partial^2 x'^a}{\partial x^d \partial x^c} = Q^a_{cd} = Q^a_{dc} \quad \Longrightarrow \quad [\frac{\partial^2 x'^a}{\partial x^d \partial x^c}]_P = Q^a_{dc}$$

which shows that

$$\begin{aligned}
[\Gamma'^a_{bc}]_P &= [(\frac{\partial x'^a}{\partial x^d})(\frac{\partial x^f}{\partial x'^c})(\frac{\partial x^e}{\partial x'^b})\Gamma^d_{ef} - \frac{\partial^2 x'^a}{\partial x^d \partial x^e} \frac{\partial x^e}{\partial x'^b} \frac{\partial x^d}{\partial x'^c}]_P \\
&= [\delta^a_d \delta^f_c \delta^e_b \Gamma^d_{ef} - \delta^e_b \delta^d_c Q^a_{de}]_P \\
&= [\Gamma^a_{bc}]_P - Q^a_{bc}.
\end{aligned}$$

Thus, if we chose $Q^a_{bc} = [\Gamma^a_{bc}]_P$ then $[\Gamma'^a_{bc}]_P \stackrel{*}{=} 0$.

Remark.

(1) The affine connection $\Gamma'^a_{bc} \neq 0$ in general.

(2) It is useful to establish a tensorial equation in geodesic coordinate. The general form of the equation can be recovered by replacing ∂_c by ∇_c.

(3) If there exists a special coordinate so that $\Gamma^a_{bc} = 0$ for the whole manifold M, then we have an affine flat manifold.

CHAPTER 11

Tensor Calculus III

11.1 Affine flatness

If we parallel transport a vector from one point to another so that the transformed vector is independent of the path, then the connection is called **integrable**.

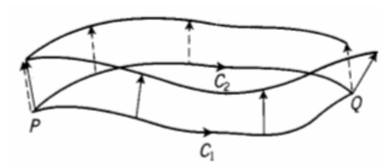

Figure 11.1: Parallel transport round two curves in M

Lemma 11.1.1. *A connection is integrable* $\Longleftrightarrow R^a_{bcd} = 0$.

Proof. (1) " $\Longrightarrow$ ". If affine connection Γ^a_{bc} is integrable, then the parallel transport of $X^a(P)$ to point Q is independent of paths connecting P and Q. We have

$$\frac{DX^a}{Du} = \nabla_X X^a = \frac{dx^c}{du}\nabla_c X^a$$

is independent of dx^c/du. Hence, $\nabla_c X^a = 0$ or

$$\partial_c X^a + \Gamma^a_{bc} X^b = 0.$$

which is a PDE for X^a. Note that X^a might not be the tangent vector, it just a contravariant vector. Since the necessary condition for the existence of solution is $\partial_d \partial_c X^a = \partial_c \partial_d X^a$, we have

$$0 = \nabla_c \nabla_d X^a - \nabla_d \nabla_c X^a = (\partial_c \partial_d - \partial_d \partial_c) X^a + R^a_{bcd} X^b$$

which, due to the fact that X^b is arbitrary, implies $R^a_{bcd} = 0$.

(2) " $\Longleftarrow$ ". Consider an infinitesimal loop formed by 4 vertices x^a, $x^a + \delta x^a$, $x^a + dx^a$, and $x^a + \delta x^a + dx^a$. Let's carry out the parallel transport from x^a to $x^a + \delta x^a + dx^a$.

[0]GR11.tex

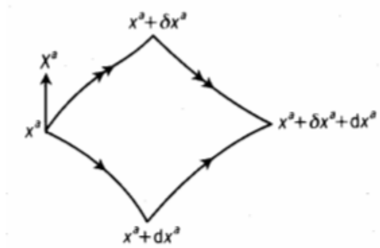

Figure 11.2: Parallel transport of X^a around a infinitesimal loop

(i) $x^a \to x^a + \delta x^a$: Since $X^a(x+\delta x) = X^a(x) + \bar{\delta}X^a(x)$ where

$$\bar{\delta}X^a(x) = -\Gamma^a_{bc}(x)X^b(x)\delta x^c$$

(ii) $x^a + \delta x^a \to x^a + \delta x^a + dx^a$:

$$\begin{aligned}
X^a(x+\delta x+dx) &= X^a(x+\delta x) + \bar{\delta}X^a(x+\delta) \\
&= X^a(x+\delta x) - \Gamma^a_{bc}(x+\delta x)X^b(x+\delta x)dx^c \\
&= X^a(x+\delta x) - (\Gamma^a_{bc} + \partial_d\Gamma^a_{bc}\delta x^d)(X^b - \Gamma^b_{ef}X^e\delta x^f)dx^c \\
&= \cancel{X^a} - \cancel{\Gamma^a_{bc}X^b\delta x^c} - \cancel{\Gamma^a_{bc}X^b dx^c} \\
&\quad -\partial_d\Gamma^a_{bc}X^b\delta x^d dx^c + \Gamma^a_{bc}\Gamma^b_{ef}X^e\delta x^f dx^c
\end{aligned}$$

Similarly,

$$\begin{aligned}
X^a(x+dx+\delta x) &= \cancel{X^a} - \cancel{\Gamma^a_{bc}X^b dx^c} - \cancel{\Gamma^a_{bc}X^b\delta x^c} \\
&\quad -\partial_d\Gamma^a_{bc}X^b dx^d\delta x^c + \Gamma^a_{bc}\Gamma^b_{ef}X^e dx^f\delta x^c
\end{aligned}$$

Hence,

$$\begin{aligned}
\Delta X^a &= X^a(x+\delta x+dx) - X^a(x+dx+\delta x) \\
&= [-\partial_d\Gamma^a_{bc} + \partial_c\Gamma^a_{bd} - \Gamma^a_{ed}\Gamma^e_{bc} + \Gamma^a_{ec}\Gamma^e_{bd}]X^b\delta x^d dx^c \\
&= -R^a_{bcd}X^b\delta x^d dx^c.
\end{aligned}$$

If $R^a_{bcd} = 0$ then $\Delta X^a = 0$ which means the connection is integrable.

$\square$

Remark.
(1) $R^a_{bcd} = -R^a_{bdc}$

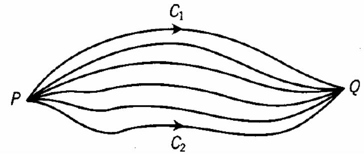

Figure 11.3: Deformation of curves in a simply connected manifold

(2) If the manifold is simply connected (no holes) then the above lemma applied to infinitesimal deformation shows that parallel transport is independent of paths.

Lemma 11.1.2. *A manifold is affine flat $\Longleftrightarrow$ the connection is symmetric and integrable.*

Proof. Since the necessary part " $\Longrightarrow$ " is trivial, we shall focus on the sufficient part " $\Longleftarrow$ ". If the connection Γ^a_{bc} is integrable at P we can choose n linear independent vectors $X_i{}^a$, $i = 1, 2, \cdots, n$ at P and parallel transport them to everywhere due to integrable property. At any point x the basis vector fields $X_i{}^a(x)$ can be viewed as a non-singular matrix, i.e. $\exists\ X^i{}_b$ such that $X^i{}_b X_i{}^a = \delta^a_b$. From parallel transport equation

$$0 = \nabla_b X_i{}^a = \partial_b X_i{}^a + \Gamma^a_{eb} X_i{}^e$$

we have $\Gamma^a_{cb} = -X^i{}_c \partial_b X_i{}^a$ which implies

$$\Gamma^a_{bc} - \Gamma^a_{cb} = X^i{}_c \partial_b X_i{}^a - X^i{}_b \partial_c X_i{}^a = X_i{}^a(\partial_c X^i{}_b - \partial_b X^i{}_c) = 0.$$

Thus $\partial_c X^i{}_b - \partial_b X^i{}_c$ or $X^i{}_b = \partial_b f^i(x)$. Let's perform a special coordinate transformation

$$x^a \to x'^a = f^a(x)$$

then

$$\frac{\partial x'^a}{\partial x^b} = \partial_b f^a = X^a{}_b, \quad \frac{\partial x^a}{\partial x'^b} = X_b{}^a.$$

The transformed connection becomes

$$\begin{aligned} \Gamma'^a_{bc} &= \frac{\partial x'^a}{\partial x^d}\frac{\partial x^e}{\partial x'^b}\frac{\partial x^f}{\partial x'^c}\Gamma^d_{ef} - \frac{\partial x^d}{\partial x'^b}\frac{\partial x^e}{\partial x'^c}\frac{\partial^2 x'^a}{\partial x^d \partial x^e} \\ &= X^a{}_d X_b{}^e X_c{}^f \Gamma^d_{ef} - X_b{}^d X_c{}^e \partial_d X^a{}_e \end{aligned}$$

Multiplying $X_a{}^h$ on both sides, we get

$$X_a{}^h \Gamma'^a_{bc} = X_b{}^e X_c{}^f \Gamma^h_{ef} - X_b{}^d X_c{}^e \Gamma^h_{de} = 0.$$

Since $X_a{}^h$ is non-singular, we have $\Gamma'^a_{bc} = 0$ everywhere which implies the connection is affine flat.

□

Corollary 11.1.3. *A manifold is affine flat $\Longleftrightarrow$ $R^a_{bcd} = 0$.*

Proof. This is just the combination of the two lemmas shown above. □

11.2 The metric

Recall that in a vector space we can introduce an inner product so that the length of a vector can be defined. The vector space with this property is called a vector space with inner product or metric space. Now we can introduce an additional structure on a smooth manifold called metric field. Let $g_{ab}(x)$ be a symmetric, covariant tensor field of $(0, 2)$ type.

Definition 11.2.1. *(Riemannian manifold) A manifold M with a metric tensor which defines the interval between x^a and $x^a + dx^a$ as*

$$(ds)^2 = g_{ab}(x) dx^a dx^b.$$

Remark.
(1) We usually denote $(ds)^2$, the square of the line element, as ds^2.
(2) The metric tensor g_{ab} is also called metric form or first fundamental form.

- Given a vector field X^a, its length or norm is defined by

$$X^2 = g_{ab}(x) X^a X^b$$

where we classify g_{ab} as follows

$$g_{ab} \text{ is } \begin{cases} \text{positive} & X^2 > 0 \qquad \forall X \\ \text{negative} & X^2 < 0 \qquad \forall X \\ \text{indefinite} & \text{otherwise} \end{cases}$$

- The angle between X^a and Y^a with $X^2 Y^2 \neq 0$ is defined by

$$\cos(X, Y) = \frac{g_{ab} X^a Y^b}{\sqrt{|X^2||Y^2|}}$$

- If $g_{ab} X^a Y^b = 0$ then $X^a \perp Y^a$ or X^a is orthogonal to Y^a.
- If $g_{ab} X^a X^b = 0$ then X^a is called null vector.
- The metric g_{ab} is called non-singular if $g = \det g_{ab} \neq 0$.
- The inverse of g_{ab} is defined by $(g_{ab})^{-1} = g^{ab}$ such that $g_{ab} g^{bc} = \delta^c_a$.
- Using g_{ab} and g^{ab} we can lowering and raising tensorial indices by defining

$$T^{\cdots\ \cdots}_{\cdots a \cdots} = g_{ab} T^{\cdots b \cdots}_{\cdots\ \cdots}, \qquad T^{\cdots a \cdots}_{\cdots\ \cdots} = g^{ab} T^{\cdots\ \cdots}_{\cdots b \cdots}$$

Thus we may consider the contravariant tensors and covariant tensors as representation of the same geometric objects, e.g. g_{ab}, δ^b_a, and g^{ab} are different representations of the same geometric object, the metric g.

Remark. Note that we have to caution about the order of indices of a tensor, e.g. ${X_b}^a = g^{ac} X_{bc}$ and ${X^a}_b = g^{ac} X_{cb}$ are different.

CHAPTER 12

Tensor Calculus IV

12.1 Metric geodesic

Let M be a manifold endowed with a metric g_{ab}. Consider a time-like curve $C : x^a = x^a(u)$ in M, then

$$(\frac{ds}{du})^2 = g_{ab}\frac{dx^a}{du}\frac{dx^b}{du}$$

and the length from P_1 to P_2 on C is given by

$$S = \int_{P_1}^{P_2} ds = \int_{P_1}^{P_2} \frac{ds}{du}du = \int_{P_1}^{P_2} \sqrt{g_{ab}\frac{dx^a}{du}\frac{dx^b}{du}}du.$$

The metric geodesic from P_1 to P_2 is a curve defined by the minimizing condition $\delta S = 0$. From Lagrangian mechanics, we can view $\sqrt{g_{ab}\frac{dx^a}{du}\frac{dx^b}{du}}$ as the Lagrangian of the system with Lagrangian

$$L(x, \dot{x}) = \sqrt{g_{ab}\frac{dx^a}{du}\frac{dx^b}{du}}, \quad \dot{x^a} \equiv \frac{dx^a}{du}$$

By Euler-Lagrangian equation, we have

$$\frac{d}{du}\frac{\partial L}{\partial \dot{x}^a} - \frac{\partial L}{\partial x^a} = 0$$

Multiplying $2L$ we rewrite above as

$$\frac{d}{du}\frac{\partial L^2}{\partial \dot{x}^a} - \frac{\partial L^2}{\partial x^a} = 2\frac{dL}{du}\frac{\partial L}{\partial \dot{x}^a}$$

which implies

$$g_{ab}\frac{d^2x^b}{du^2} + \{bc, a\}\frac{dx^b}{du}\frac{dx^c}{du} = \frac{(\frac{d^2s}{du^2})}{\frac{ds}{du}}g_{ab}\frac{dx^b}{du}$$

where

$$\{ab, c\} = \frac{1}{2}(\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab}), \quad \{ab, c\} = \{ba, c\}.$$

Introducing Christoffel symbols of 2nd kind

$$\left\{ \begin{array}{c} a \\ bc \end{array} \right\} = g^{ad}\{bc, d\}$$

[0]GR12.tex

the Euler-Lagrangian equation becomes

$$\frac{d^2x^a}{du^2} + \left\{ \begin{array}{c} a \\ bc \end{array} \right\} \frac{dx^b}{du}\frac{dx^c}{du} = \frac{(\frac{d^2s}{du^2})}{\frac{ds}{du}}\frac{dx^a}{du}.$$

Let's choose a special parametrization $u = \alpha s + \beta$, then $du = \alpha ds$ and

$$\boxed{\frac{d^2x^a}{ds^2} + \left\{ \begin{array}{c} a \\ bc \end{array} \right\} \frac{dx^b}{ds}\frac{dx^c}{ds} = 0}$$

with

$$g_{ab}\frac{dx^a}{ds}\frac{dx^b}{ds} = 1.$$

Remark.

(1) For space-like geodesic:

$$ds^2 = g_{ab}dx^a dx^b < 0$$

we can set $d\sigma^2 = -ds^2 = -g_{ab}dx^a dx^b > 0$ and thus

$$\frac{d^2x^a}{d\sigma^2} + \left\{ \begin{array}{c} a \\ bc \end{array} \right\} \frac{dx^b}{d\sigma}\frac{dx^c}{d\sigma} = 0$$

with

$$g_{ab}\frac{dx^a}{d\sigma}\frac{dx^b}{d\sigma} = -1.$$

(2) For indefinite metric, we can choose the so-called affine parameter $u \neq s$ such that

$$\frac{d^2x^a}{du^2} + \left\{ \begin{array}{c} a \\ bc \end{array} \right\} \frac{dx^b}{du}\frac{dx^c}{du} = 0$$

with

$$g_{ab}\frac{dx^a}{du}\frac{dx^b}{du} = 0$$

called null geodesic which is invariant under $u \to \alpha u + \beta$.

(3) The contravariant and covariant indices can be viewed as components of vectors in non-orthogonal frame (or oblique axis). Suppose $\{e_1, e_2\}$ is an non-orthogonal basis, i.e. $e_1 \cdot e_2 \neq 0$. Then a vector A can be decomposed by e_i as

$$A = A^1 e_1 + A^2 e_2.$$

If we define $A_i = A \cdot e_i$, then

$$\begin{aligned} A_1 &= A \cdot e_1 = A^1 e_1 \cdot e_1 + A^2 e_2 \cdot e_1 \\ A_2 &= A \cdot e_2 = A^1 e_1 \cdot e_2 + A^2 e_2 \cdot e_2 \end{aligned}$$

Writing $A_a = g_{ab}A^b$ where

$$(g_{ab}) = \begin{pmatrix} e_1 \cdot e_1 & e_1 \cdot e_2 \\ e_2 \cdot e_1 & e_2 \cdot e_2 \end{pmatrix}$$

Note that if $e_1 \perp e_2$, then

$$(g_{ab}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and $A_a = A^a$.

12.2 The metric connection

So far, we have two kind of geodesics

(1) affine geodesic: defined by parallel transport as

$$\frac{d^2x^a}{ds^2} + \Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds} = 0$$

(2) metric geodesic: defined by minimizing the length interval as

$$\frac{d^2x^a}{ds^2} + \left\{ \begin{array}{c} a \\ bc \end{array} \right\}\frac{dx^b}{ds}\frac{dx^c}{ds} = 0$$

If we identify $\Gamma^a_{bc} = \left\{ \begin{array}{c} a \\ bc \end{array} \right\}$ and call it metric connection, then the torsion free condition $\Gamma^a_{bc} = \Gamma^a_{cb}$ holds automatically. Therefore

$$\boxed{\Gamma^a_{bc} = \frac{1}{2}g^{ad}(\partial_b g_{cd} + \partial_c g_{bd} - \partial_d g_{bc})}$$

Now, looking at the covariant derivative of g_{ab}, we find

$$\begin{aligned}
\nabla_c g_{ab} &= \partial_c g_{ab} - \Gamma^d_{ac}g_{db} - \Gamma^d_{bc}g_{ad} \\
&= \partial_c g_{ab} - \frac{1}{2}g^{de}(\partial_a g_{ec} + \partial_c g_{ea} - \partial_e g_{ac})g_{db} \\
&\quad -\frac{1}{2}g^{de}(\partial_b g_{ec} + \partial_c g_{eb} - \partial_e g_{bc})g_{da} \\
&= \cancel{\partial_c g_{ab}} - \frac{1}{2}(\cancel{\partial_a g_{bc}} + \cancel{\partial_c g_{ba}} - \cancel{\partial_b g_{ac}}) - \frac{1}{2}(\cancel{\partial_b g_{ac}} + \cancel{\partial_c g_{ab}} - \cancel{\partial_a g_{bc}}) \\
&= 0
\end{aligned}$$

which implies that the metric g_{ab} likes a scalar under parallel transport. Conversely, if $\nabla_c g_{ab} = 0$ then

$$\partial_c g_{ab} = \Gamma^d_{ac}g_{db} + \Gamma^d_{bc}g_{ad}$$

which implies

$$\Gamma^a_{bc} = \frac{1}{2}g^{ad}(\partial_b g_{cd} + \partial_c g_{bd} - \partial_d g_{bc})$$

Example. Show that above assertion is true.

Theorem 12.2.1. *Let ∇_c be the covariant derivative with respect to the connection Γ^a_{bc}. Then*

$$\nabla_c g_{ab} = 0 \quad \Longleftrightarrow \quad \Gamma^a_{bc} = \frac{1}{2}g^{ad}(\partial_b g_{cd} + \partial_c g_{bd} - \partial_d g_{bc})$$

Hence g_{ab} commute with ∇_c. In particular,

$$\nabla_c(g^{ad}g_{db}) = \nabla_c\delta^a_b = \partial_c\delta^a_b + \Gamma^a_{dc}\delta^d_b - \Gamma^d_{bc}\delta^a_d = 0$$

which implies $\nabla_c g^{ad} = 0$ as well.

12.3 Metric flatness

Recall two definitions of flatness for a manifold.

(1) affine flatness: $\exists$ a special coordinate system such that $\Gamma^a_{bc} = 0$ everywhere. In fact, we have

$$\text{affine flatness} \quad \Longleftrightarrow \quad R^a_{bcd} = 0$$

(2) metric flatness: $\exists$ a special coordinate system such that $g_{ab} = const.$ or

$$(g_{ab}) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

Theorem 12.3.1. *(metric flatness)*

$$\textit{A metric is flat} \quad \Longleftrightarrow \quad R^a_{bcd} = 0$$

Proof. (1) " $\Longrightarrow$ ". If $g_{ab} = const.$ then the metric connection $\Gamma^a_{bc} = 0$ and $R^a_{bcd} = 0$.

(2) " $\Longleftarrow$ ". If $R^a_{bcd} = 0$ then we have affine flatness, i.e. $\Gamma^a_{bc} = 0$ everywhere. Thus

$$\nabla_c g_{ab} = 0 = \partial_c g_{ab}$$

which implies $g_{ab} = const.$ or metric flatness.

$\square$

Hence, under identification $\Gamma^a_{bc} = \left\{ \begin{matrix} a \\ bc \end{matrix} \right\}$

$$\underbrace{\text{metric flatness}}_{g_{ab}=const.} = \underbrace{\text{affine flatness}}_{\Gamma^a_{bc}=0}$$

CHAPTER 13

Tensor Calculus V

13.1 The curvature tensor

Recall the Riemann-Christoffel tensor

$$R^a{}_{bcd} = \partial_c \Gamma^a_{bd} - \partial_d \Gamma^a_{bc} + \Gamma^e_{bd}\Gamma^a_{ec} - \Gamma^e_{bc}\Gamma^a_{ed}$$

where

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad}(\partial_b g_{cd} + \partial_c g_{bd} - \partial_d g_{bc})$$

Several remarks are in order.

(1) Riemann tensor is 2nd order derivative in g, i.e.

$$R \sim \partial\Gamma + \Gamma\Gamma \sim g\partial^2 g + (\partial g)(\partial g)$$

(2) Anti-symmetry: By definition

$$R^a{}_{bcd} = -R^a{}_{bdc}$$

(3) Cyclic property:

$$R^a{}_{bcd} + R^a{}_{dbc} + R^a{}_{cdb} = 0$$

It can verified by definition of $R^a{}_{bcd}$ and $\Gamma^a_{bc} = \Gamma^a_{cb}$.

(4) Cyclic property in terms of covariant tensor:

$$R_{abcd} + R_{adbc} + R_{acdb} = 0$$

where a, b, c, d are all different.

(5) Symmetry property of covariant Riemann tensor:

$$\begin{aligned}
R_{abcd} &= -R_{bacd} \quad (a \leftrightarrow b) \\
R_{abcd} &= -R_{abdc} \quad (c \leftrightarrow d) \\
R_{abcd} &= -R_{cdab} \quad ((ab) \leftrightarrow (cd))
\end{aligned}$$

[0]GR13.tex

Let's justify above properties in geodesic coordinate.

$$
\begin{aligned}
R_{abcd} &= g_{ae}R^e{}_{bcd} \\
&= g_{ae}(\partial_c\Gamma^e_{bd} - \partial_d\Gamma^e_{bc} + \Gamma\Gamma - \Gamma\Gamma) \\
&= \partial_c(g_{ae}\Gamma^e_{bd}) - \partial_d(g_{ae}\Gamma^e_{bc}) - \partial_c g_{ae}\Gamma - \partial_d g_{ae}\Gamma + \Gamma\Gamma - \Gamma\Gamma \\
&\stackrel{*}{=} \frac{1}{2}[\partial_c(g_{ae}g^{ef}(\partial_b g_{fd} + \partial_d g_{fb} - \partial_f g_{bd})) - \partial_d(g_{ae}g^{ef}(\partial_b g_{fc} + \partial_c g_{fb} - \partial_f g_{bc}))] \\
&= \frac{1}{2}[\partial_c(\partial_b g_{ad} + \cancel{\partial_d g_{ab}} - \partial_a g_{bd}) - \partial_d(\partial_b g_{ac} + \cancel{\partial_c g_{ab}} - \partial_a g_{bc})] \\
&= \frac{1}{2}[g_{ad,bc} - g_{bd,ac} - g_{ac,bd} + g_{bc,ad}] \\
&= \frac{1}{2}[g_{d[a,b]c} - g_{c[a,b]d}] \\
&= \frac{1}{2}[g_{a[d,c]b} - g_{b[d,c]a}]
\end{aligned}
$$

where we have used the geodesic coordinate ($\Gamma \stackrel{*}{=} 0$) to reach the fourth equality.

In summary, for $R^a{}_{bcd}$ we have

$$R_{abcd} = -R_{abdc} = -R_{bacd} = R_{cdab}$$

and

$$R_{abcd} + R_{adbc} + R_{acdb} = 0$$

Proposition 13.1.1. *R_{abcd} satisfies the Bianchi identity*

$$\nabla_a R_{debc} + \nabla_c R_{deab} + \nabla_b R_{deca} = 0$$

Proof. Since this is a tensorial equation. If we can verify that it vanishes in a coordinate systems, then it vanishes in all coordinate systems. Let's work in geodesic coordinate system where $\Gamma^a_{bc} = 0$. Hence

$$
\begin{aligned}
& \nabla_a R_{debc} + \nabla_c R_{deab} + \nabla_b R_{deca} \\
={}& (\partial_a R_{debc} + \partial_c R_{deab} + \partial_b R_{deca}) \\
={}& \partial_a(\cancel{\partial_b\Gamma_{dec}} - \cancel{\partial_c\Gamma_{deb}}) + \partial_c(\cancel{\partial_a\Gamma_{deb}} - \cancel{\partial_b\Gamma_{dea}}) + \partial_b(\cancel{\partial_c\Gamma_{dea}} - \cancel{\partial_a\Gamma_{dec}}) \\
={}& 0
\end{aligned}
$$

which implies that the Bianchi identity holds in any coordinate system.

□

Next, we shall introduce two important geometric quantities: Ricci tensor and Ricci scalar.

- Ricci tensor: By contracting the 1st and 3rd indices of the Riemann tensor, we have a $(0,2)$ type symmetric tensor
$$R_{ab} = R^c{}_{acb} = g^{cd}R_{dacb} = g^{cd}R_{cbda} = R_{ba}$$
called Ricci tensor.

- Ricci scalar: By contracting indices of Ricci tensor we get a scalar
$$R = g^{ab}R_{ab} = R^a{}_a$$
called Ricci scalar.

- Einstein tensor: Using Ricci tensor and Ricci scalar we can define Einstein tensor

$$\boxed{G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R}$$

which satisfies the contracted Bianchi identity

$$\nabla_b G_a{}^b = 0.$$

Indeed, from Bianchi identity, we have

$$\begin{aligned}
0 &= g^{db}(\nabla_a R_{debc} + \nabla_c R_{deab} + \nabla_b R_{deca}) = \nabla_a R_{ec} - \nabla_c R_{ea} + \nabla_b R^b{}_{eca} \\
&= g^{ce}(\nabla_a R_{ec} - \nabla_c R_{ea} + \nabla_b R^b{}_{eca}) = \nabla_a R - \nabla_c R^c{}_a - \nabla_b R^b{}_a \\
&= -2\nabla_c(R^c{}_a - \frac{1}{2}\delta^c_a R) = -\nabla_c G^c{}_a
\end{aligned}$$

Finally, we would like to know how many independent components of $R^a{}_{bcd}$ in a 4-dimensional manifold. Let's answer this question as follows:

(1) For an anti-symmetric tensor $A_{\mu\nu} = -A_{\nu\mu}$, where $\mu, \nu = 0, 1, 2, 3$ there are $C^4_2 = 6$ independent components, namely

$$(\mu\nu) = (01), (02), (03), (12), (1,3), (2,3)$$

We may denote $(\mu\nu) = A$, $A = 1, 2, \cdots, 6$.

(2) For a symmetric tensor $T_{AB} = T_{BA}$, $A, B = 1, 2, \cdots, 6$, we have $6(6+1)/2 = 21$ independent components.

(3) Finally, for $R_{a(bcd,cyclic)} = 0$ where a, b, c, d are all different, there is only one equation.

Hence, the total independent components in Riemann tensor R_{abcd} is $(21 - 1) = 20$.

Example. ($D = n$) Let's generalize to an n-dimensional manifold with $n \geq 4$. Following above discussion, the total independent components for $T_{AB} = T_{BA}$ is

$$m = \frac{\frac{n(n-1)}{2}(\frac{n(n-1)}{2} + 1)}{2} = \frac{n(n-1)(n^2 - n + 2)}{8}$$

where $A, B = 1, \cdots, n(n-1)/2$. On the other hand, the equation $R_{\lambda(\mu\nu\rho,cyclic)} = 0$ caontains $C^n_4 = n(n-1)(n-2)(n-3)/4$ independent equations. Thus, the total number N_n of independent components of Riemann tensor $R_{\lambda\mu\nu\rho}$ is

$$N_n = m - C^n_4 = \frac{n^2(n^2 - 1)}{12}.$$

In particular, for $n = 4$, $N_4 = 16 \times 15/12 = 20$.

13.2 The Weyl tensor

The Weyl tensor (or conformal tensor) is defined in $n \geq 3$ as

$$\begin{aligned}
C_{abcd} = R_{abcd} &+ \frac{1}{n-2}(g_{ad}R_{cb} + g_{bc}R_{da} - g_{ac}R_{db} - g_{bd}R_{ca}) \\
&+ \frac{1}{(n-1)(n-2)}(g_{ac}g_{db} - g_{ad}g_{cb})R
\end{aligned}$$

Example. ($n = 3, 4$)

$$\begin{aligned}
n = 3 \qquad & C_{abcd} = R_{abcd} + (\cdots) + \frac{1}{2}(\cdots)R \\
n = 4 \qquad & C_{abcd} = R_{abcd} + \frac{1}{2}(\cdots) + \frac{1}{6}(\cdots)R
\end{aligned}$$

Table 13.1: Curvature tensor in 4 dimensions or less

dimension (n)	N_n	independent components
1	0	$R_{abcd} = 0$
2	1	R (Ricci scalar)
3	6	R_{ab} (Ricci tensor)
4	20=10+10	R_{ab} (Ricci tensor), C_{abcd} (Weyl tensor)

Remark.
(1) The Weyl tensor has the same symmetries as Riemann tensor R_{abcd}, i.e.

$$C_{abcd} = -C_{abdc} = -C_{bacd} = C_{cdab}$$

and

$$C_{abcd} + C_{adbc} + C_{acdb} = 0$$

(2) The Weyl tensor satisfies an additional symmetry

$$C^{a}{}_{bad} \equiv 0.$$

which together with above shows that it is trace free, i.e. vanishes for any pair of contraction. It can be proved as follows

$$\begin{aligned}
C^{a}{}_{bad} &= g^{ac} C_{abcd} \\
&= g^{ac}\left[R_{abcd} + \frac{1}{2}(g_{ad}R_{cb} + g_{bc}R_{da} - g_{ac}R_{db} - g_{bd}R_{ca}) + \frac{1}{6}(g_{ac}g_{db} - g_{ad}g_{cb})R\right] \\
&= R_{bd} + \frac{1}{2}(R_{db} + R_{db} - 4R_{db} - g_{bd}R) + \frac{1}{6}(4g_{bd}R - g_{bd}R) \\
&= 0.
\end{aligned}$$

Hence, the independent components of Weyl tensor C_{abcd} is $20 - 4(4+1)/2 = 10$.

Remark. (Weyl tensor and conformal invariant) Two metrics $g_{ab}(x)$ and $\bar{g}_{ab}(x)$ are conformally related, if

$$\bar{g}_{ab}(x) = \Omega^2(x) g_{ab}(x)$$

where $\Omega(x)$ is a non-zero differentiable function. Note that conformally related metrics define the same angle between two vectors, i.e.

$$\cos(X,Y) = \overline{\cos(X,Y)}.$$

Any quantity which satisfies $\overline{T^{a\cdots}_{b\cdots}} = T^{a\cdots}_{b\cdots}$ is called conformally invariant.

Example. Show that $\bar{C}^{a}{}_{bcd} = C^{a}{}_{bcd}$, (but g_{ab}, Γ^a_{bc}, $R^{a}{}_{bcd}$ are not)

ANS: Check $\bar{C}^{a}{}_{bcd} = \bar{g}^{ae}\bar{C}_{ebcd}$.

Remark.
(1) If $g_{ab}(x) = \Omega^2(x)\eta_{ab}$ then $g_{ab}(x)$ is conformally flat.
(2) Two important theorems which do not prove here.

(i) The metric g_{ab} is conformally flat $\Longleftrightarrow$ $C_{abcd} = 0$ everywhere

(ii) Any 2-d Riemannian manifold is conformally flat i.e. $C_{abcd} = 0$