Publishers’ page
I dedicate.....
Preface

...This material is prepared based on my lecture note for General Relativity course (PHYS 4800) at Middle Tennessee State University.
## Contents

**Preface** vii

1. The spacetime of special relativity 11
   1.1 Inertial frames, Galilean and Lorentz transformations . . . . . . . . 11
   1.2 Four dimensional rotation vs Minkowski space time coordinate trans-
       formation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
   1.3 The interval and the lightcone . . . . . . . . . . . . . . . . . . . . . 21
   1.4 Length contraction and time dilation . . . . . . . . . . . . . . . . . 23
   1.5 Invariant hyperbolae . . . . . . . . . . . . . . . . . . . . . . . . . . . 26
   1.6 Minkowski spacetime element, Particle worldline, and proper time . 28
   1.7 The Doppler effect . . . . . . . . . . . . . . . . . . . . . . . . . . . 30
   1.8 Addition and acceleration in special relativity . . . . . . . . . . . . 32

2. Manifolds 35
   2.1 What is a Manifold? . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
   2.2 Curves and surfaces in a Manifold . . . . . . . . . . . . . . . . . . . . . . 36
   2.3 Coordinate transformations and summation convention . . . . . . . . . 37
   2.4 The Riemannian geometry . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41
   2.5 Intrinsic and extrinsic geometry and the metric . . . . . . . . . . . . . 42
x

Book Title
Chapter 1

The spacetime of special relativity

In this chapter we review the basic notions underlying in the Newtonian and special relativistic viewpoints of space and time.

1.1 Inertial frames, Galilean and Lorentz transformations

If we want to describe an event that occurred somewhere in this country, this continent, planet, solar system, galaxy, or universe we must be able to tell exactly where (space) it occurred and where (time). This means we must be able to tell the time at which the event occurred and the exact position of this event in this space that we know (three dimensional space) as measured from some origin $O$ on some reference frame $S$ or from another origin on another reference frame $S'$ which may be oriented and/or moving with respect to $S$ in an arbitrary manner.

These reference frames could be our home, our car that we are driving on the highway, Orion spacecraft traveling in the deep outer space (in the near future), or the international space station orbiting our planet. Suppose an event in space described by the coordinates $(x, y, z)$ is recorded at time $t$ by an observer at $O$ on reference $S$, as shown in Fig. 1.1. Using a four dimensional space we may express this event as $(t, x, y, z)$. Let’s say this same event for an observer $O'$ on reference $S$ is recorded at a different time $t'$ at a different point in space $(t', x', y', z')$.

Now the question is how these two observation are related to one another if one of the reference frame is moving with some velocity $v$. Well before we relate $(t', x', y', z')$ and $(t, x, y, z)$ it is very important to know about the reference frames. In this course we will consider inertial reference frames. Inertial reference frames are reference frames moving with a constant velocity (constant magnitude and direction) with respect to one another. These are none accelerating frame of references

\[
\frac{d^2x}{dt^2} = \frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} = 0
\]

and

\[
\frac{d^2x'}{dt^2} = \frac{d^2y'}{dt^2} = \frac{d^2z'}{dt^2} = 0.
\]
Consider the inertial reference frame \( S' \) in Fig (1.1). It is moving with a constant velocity \( v \) in the positive \( x \) direction

\[
\vec{v} = v\hat{x}
\]

relative to \( O \). Obviously, reference \( S \) is moving with the same magnitude of velocity but in the opposite direction for an observer \( O' \)

\[
\vec{v} = -v\hat{x}
\]

(1.1)

Taking into account this velocity, we can at least expect the \( x \) position of the event for an observer \( O' (x') \) would depend on both \( x \) and \( t \). We make a linear relationship assumption

\[
x' = Dt + Ex
\]

(1.2)

Let’s make a similar assumption for time \( t' \) too, why not?

\[
t' = At + Bx
\]

(1.3)

For \( y' \) and \( z' \), obviously

\[
y' = y, z' = z.
\]

(1.4)

Since we require that for \( x' = 0 \),

\[
x = vt
\]

we have

\[
x' = Dt + Ex \Rightarrow 0 = Dt + Evt \Rightarrow D = -Ev
\]

(1.5)

For \( x = 0 \), we also require

\[
x' = -vt'
\]
which leads to

\[ x' = Dt + Ex \Rightarrow -vt' = Dt \Rightarrow D = -\frac{v't'}{t}. \] (1.6)

Substituting \( x = 0 \) into

\[ t' = At + Bx \] (1.7)

we find

\[ t' = At \] (1.8)

so that

\[ D = -\frac{v't'}{t} = -Av. \] (1.9)

From Eqs. (1.5) and (1.9) we see that

\[ E = A, D = -Av \]

so that an inertial reference frames we may write

\begin{align*}
  t' &= At + Bx \quad (1.10) \\
  x' &= A(x - vt) \quad (1.11) \\
  y' &= y \quad (1.12) \\
  z' &= z \quad (1.13)
\end{align*}

**Galilean Transformation**: in the Galilean transformation time is absolute that mean \( t \) is independent of the reference frame we have

\[ A = 1, B = 0 \]

which leads to

\begin{align*}
  t' &= t \quad (1.14) \\
  x' &= x - vt \quad (1.15) \\
  y' &= y \quad (1.16) \\
  z' &= z \quad (1.17)
\end{align*}

If take the time derivative of the \( x \) coordinate, we find

\[ \frac{dx'}{dt'} = \frac{dx}{dt} - v \Rightarrow u'_x = u_x - v \] (1.18)

which is an equation that can be obtained from a common sense of relative velocities.

If you differentiate the velocity,

\[ \frac{du'_x}{dt'} = \frac{du_x}{dt} \] (1.19)
since the velocity of the reference frames is constant. This means the acceleration is the same in both \( S \) and \( S' \).

**Homework Problem 1:** You heard in the News that there are two bad events happened somewhere in this planet (as usual). Suppose event one occurred at time \( t_1 \) at a point in space \((x_1, y_1, z_1)\), which we may describe using spacetime coordinates \((t_1, x_1, y_1, z_1)\), as recorded by an observer on an inertial reference frame \( S \). The second event occurred at a later time \( t_2 \) at another point in space \((x_2, y_2, z_2)\) as recorded by the same observer. Show that the time difference

\[
\Delta t = t_2 - t_1
\]

and the quantity

\[
(\Delta r)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2
\]

are separately invariant under any Galilean transformation. Note that

\[
\Delta x = x_2 - x_1, \Delta y = y_2 - y_1, \Delta z = z_2 - z_1
\]

You must show that

\[
\Delta t' = \Delta t, (\Delta r')^2 = (\Delta r)^2
\]

**Lorentz transformation:** In the special theory of relativity Einstein abandoned the postulate of an absolute time and replaced it by the postulated that the speed of light \( c \) is the same in all inertial frame. Next we will derive the Lorentz transformation using Einstein postulate of the speed of light. To this end, we may rewrite Eqs. (1.10)-(1.13)

\[
ct' = c(At + Bx)
\]

\[
x' = A(x - vt)
\]

\[
y' = y
\]

\[
z' = z
\]

For a photon in each frame, both \( S \) and \( S' \), we must have

\[
(ct)^2 = (\Delta r)^2 \Rightarrow (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 = 0
\]

\[
(c'\Delta t')^2 = (\Delta r')^2 \Rightarrow (c\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 = 0
\]

Using Eqs. (1.23)-(1.26), we may write

\[
(c\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 = c^2 (A\Delta t + B\Delta x)^2
\]

\[
- A^2 (\Delta x - v\Delta t)^2 - (\Delta y)^2 - (\Delta z)^2 = 0
\]
\[ c^2 \left(A^2 (\Delta t)^2 + B^2 (\Delta x)^2 + 2AB (\Delta t) (\Delta x)\right) - A^2 \left((\Delta x)^2 + v^2 (\Delta t)^2 - 2v (\Delta x) (\Delta t)\right) - (\Delta y)^2 - (\Delta z)^2 = 0 \]

\[ \Rightarrow \] \[ c^2 A^2 (\Delta t)^2 + c^2 B^2 (\Delta x)^2 + 2c^2 AB (\Delta t) (\Delta x) - \left(A^2 (\Delta x)^2 + A^2 v^2 (\Delta t)^2 - 2A^2 v (\Delta x) (\Delta t)\right) - (\Delta y)^2 - (\Delta z)^2 = 0 \]

\[ \Rightarrow \] \[ c^2 A^2 (\Delta t)^2 + c^2 B^2 (\Delta x)^2 + 2c^2 AB (\Delta t) (\Delta x) - A^2 (\Delta x)^2 - A^2 v^2 (\Delta t)^2 + 2A^2 v (\Delta x) (\Delta t) - (\Delta y)^2 - (\Delta z)^2 = 0 \]

\[ \Rightarrow [c^2 A^2 - A^2 v^2] (\Delta t)^2 + [c^2 B^2 - A^2] (\Delta x)^2 + 2A [c^2 B + Av] (\Delta x) (\Delta t) - (\Delta y)^2 - (\Delta z)^2 = 0 \] \hspace{1cm} (1.29)

Now referring to Eq. (1.27), we must have
\[ c^2 A^2 - A^2 v^2 = c^2 \] \Rightarrow \[ A = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma \] \hspace{1cm} (1.30)

\[ c^2 B + Av = 0 \] \Rightarrow \[ B = -\frac{v}{c^2} A = -\frac{\beta}{c} A = -\frac{\gamma \beta}{c} \] \hspace{1cm} (1.31)

where
\[ \beta = \frac{v}{c}, \] \hspace{1cm} (1.32)

\[ \gamma = \frac{1}{\sqrt{1 - \beta^2}}. \] \hspace{1cm} (1.33)

Substituting Eqs. (1.30) and (1.31) into Eqs. 1.23 and 1.24, the Lorentz transformation is given by
\[ ct' = c (At + Bx) = c \left(\gamma t - \frac{\gamma \beta}{c} x\right) \Rightarrow ct' = \gamma (ct - \beta x), \] \hspace{1cm} (1.34)

\[ x' = \gamma (x - vt) \Rightarrow x' = \gamma (x - c\beta t), \] \hspace{1cm} (1.35)

\[ y' = y, \] \hspace{1cm} (1.36)

\[ z' = z. \] \hspace{1cm} (1.37)

The Lorentz transformation is also known as \textit{the boost} in the \textit{x}-direction. For the case \( v \ll c \), we have
\[ \beta \approx 0, \gamma \approx 1 \]
the Lorentz transformation reduces to

\[
\begin{align*}
t' &= t, \\
x' &= x - vt, \\
y' &= y, \\
z' &= z.
\end{align*}
\]  

(1.38)  

(1.39)  

(1.40)  

(1.41)

which agrees with the Galilean transformation.

**Homework Problem 2:** Consider the two bad events in problem 1 described by the spacetime coordinates \((t_1, x_1, y_1, z_1)\) and \((t_2, x_2, y_2, z_2)\). Show that the interval between these two events squared

\[
(\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2
\]  

(1.42)

is invariant under the Lorentz transformation squared.

The interval in Eq. (1.42) shows that space and time are united in a four dimensional continuum called spacetime. The geometry which is characterized by Eq. (1.42) is known as **Minkowski geometry**. The spacetime of special relativity is not Euclidean because of the minus sign. It is often called pseudo Euclidean.

*Note: The inverse transformation whether it is Galilean or Lorentz, the equations are given by replacing \(v\) by \(-v\).*

\[
\begin{align*}
ct &= \gamma (ct' + \beta x'), \\
x &= \gamma (x' + c\beta t'), \\
y &= y', \\
z &= z'.
\end{align*}
\]  

(1.43)  

(1.44)  

(1.45)  

(1.46)

Eqs. (1.43)- (1.46) are Lorentz transformation from the \(S'\) frame to \(S\).

### 1.2 Four dimensional rotation vs Minkowski space time coordinate transformation

Let’s consider the geometry that we are familiar with—**the Euclidean geometry** in Cartesian coordinates and add a fourth dimension—time. We call this coordinate system \(S\). In this coordinate system a point can be described by the coordinates \((t, x, y, z)\) as now let’s rotate \(S\) about the x axis by an angle \(\theta\) in a clockwise direction to form another coordinate system (reference frame) \(S'\). In this rotated coordinate system the coordinate is given by \((t', x', y', z')\). Applying what you studied in The-
Fig. 1.1 Rotation about the x-axis in a clockwise direction by an angle $\theta$.

Curves, you can easily show that

\begin{align}
ct &= ct', \quad (1.47) \\
x' &= x, \quad (1.48) \\
y' &= y \cos (\theta) - z \sin (\theta) \quad (1.49) \\
z' &= y \sin (\theta) + z \cos (\theta) \quad (1.50)
\end{align}

or in a Matrix form

\[
\begin{bmatrix}
ct' \\
x' \\
y' \\
z'
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 \cos (\theta) & -\sin (\theta) & 0 \\
0 & 0 \sin (\theta) & \cos (\theta) & 0
\end{bmatrix} \begin{bmatrix}
ct \\
x \\
y \\
z
\end{bmatrix} \quad (1.51)
\]

We recall the "boost" (Lorentz transformation) from Eqs. (1.34)- (1.35),

\begin{align}
ct' &= \gamma (ct - \beta x), \quad (1.52) \\
x' &= \gamma (x - c\beta t), \quad (1.53) \\
y' &= y, \quad (1.54) \\
z' &= z. \quad (1.55)
\end{align}

that can be put in, using matrices, the form

\[
\begin{bmatrix}
ct \\
x' \\
y' \\
z'
\end{bmatrix} = \begin{bmatrix}
\gamma & -\gamma \beta & 0 & 0 \\
-\gamma \beta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
ct \\
x \\
y \\
z
\end{bmatrix} \quad (1.56)
\]
The similarity between the boost an ordinary rotation can be seen if we introduce the parameter called the rapidity parameter \( \psi \) defined by

\[
\psi = \tanh^{-1} (\beta)
\]

(1.57)

where

\[
\beta = \frac{v}{c} = \begin{cases} 
1 & v = c \\
0 & v = 0
\end{cases}
\]

(1.58)

Using Eq. (1.58), we note that

\[
\psi = \tanh^{-1} (\beta) = \begin{cases} 
\infty, & \beta = 1 \\
0, & \beta = 0
\end{cases}
\]

(1.59)

From the relation

\[
cosh^2 (\psi) - \sinh^2 (\psi) = 1 \Rightarrow 1 - \tanh^2 (\psi) = \frac{1}{\cosh^2 (\psi)} \Rightarrow \cosh (\psi) = \frac{1}{\sqrt{1 - \tanh^2 (\psi)}}
\]

we can write

\[
\gamma = \frac{1}{\sqrt{1 - \beta^2}} = \cosh (\psi).
\]

and

\[
\gamma \beta = \cosh (\psi) \tanh (\psi) = \sinh (\psi).
\]

Therefore, we can express the "Boost", using the hyperbolic functions, as

\[
\begin{bmatrix}
ct' \\
x' \\
y' \\
z'
\end{bmatrix} = \begin{bmatrix}
\cosh (\psi) & - \sinh (\psi) & 0 & 0 \\
- \sinh (\psi) & \cosh (\psi) & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
ct \\
x \\
y \\
z
\end{bmatrix}.
\]

(1.60)

Now comparing Eqs. (1.51) and (1.60), we can see that the boost has essentially has the same structure as coordinate rotation except the trigonometric functions are replaced by hyperbolic function. Therefore, the physics of relativity (for an inertial frames of reference) is a coordinate transformation in the Minkowski spacetime! We can easily see this by considering the \( S \) and \( S' \) frame of references in the standard configuration (i.e. \( S' \) is moving with a constant velocity \( v \) in the positive \( x \)-direction). Let's omit the \( y \) and \( z \) coordinates an event described by the coordinates \( (ct, x) \) in an inertial reference frame \( S \) and by coordinates \( (ct', x') \) in the \( S' \) reference frame. Let's construct \( S' \) frame by rotating \( ct \) axis in a clockwise direction and the \( x \) axis in a counterclockwise direction by an angle \( \psi \), which is given by Eq.
Fig. 1.2 Rotation in Minkowski space.

(1.59), as shown in Fig. 1.2. Now if we resolve $ct$ as well as $x$ into components along $ct'$ and $x'$, we can easily see that

$$
ct' = ct \cosh(\psi) - x \sinh(\psi),
$$
$$
x' = -ct \sinh(\psi) + x \cosh(\psi).
$$

Using $y = y$ and $z' = z$ for the boost, we have

$$
ct' = ct \cosh(\psi) - x \sinh(\psi),
$$
$$
x' = -ct \sinh(\psi) + x \cosh(\psi),
$$
$$
y' = y,
$$
$$
z' = z.
$$
Fig. 1.3 An inertial frame $S'$ moving with a constant velocity $\vec{v}$ in arbitrary direction (but constant). The origin is off by $\vec{R}$ with respect to $S$ and the axes are rotated at the initial time $t = t' = 0$.

or using matrices

$$
\begin{pmatrix}
  c t' \\
  x' \\
  y' \\
  z'
\end{pmatrix} =
\begin{pmatrix}
  \cosh(\psi) & -\sinh(\psi) & 0 & 0 \\
  -\sinh(\psi) & \cosh(\psi) & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  c t \\
  x \\
  y \\
  z
\end{pmatrix},
$$

(1.61)

which is the Lorentz transformation.

Though generally, the inertial reference frame, $S'$ moves with a constant velocity, $\vec{v}$, in an arbitrary direction and the origin could be off with respect to the origin of the inertial reference frame $S$ at the initial time $t = t' = 0$, and also the axes could be rotated (see Fig. , we will consider only the "boost" (the $S'$ reference frame in Fig. 1.1). This is because the displacement and the rotation does not introduce any new physics. It is the boost only, which is also known as standard configuration, that brings a new physics due to the relative motion. Now let’s look how we can get the inertial reference frame $S'$ from $S$ by decomposing the transformation. First
translate the origin O by \( \vec{R} \). This makes the origin of S and S’ coincide. Rotate the x-axis by \( \alpha \). This lines up the velocity \( \vec{v} \) with the x-axis. A final rotation lines up the inertial frame \( S' \) with \( S \). That makes the \( S' \) the standard configuration shown in Fig. 1.1.

**Homework Problem 3:** Using the Lorentz transformation

\[
\begin{bmatrix}
ct' \\
x' \\
y' \\
z'
\end{bmatrix} = \begin{bmatrix}
\cosh(\psi) & -\sinh(\psi) & 0 & 0 \\
-\sinh(\psi) & \cosh(\psi) & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
ct \\
x \\
y \\
z
\end{bmatrix}, \quad (1.62)
\]

Show that the interval squared between these two bad events in problem 1 is invariant.

### 1.3 The interval and the lightcone

We have proved that the interval

\[
(\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \quad (1.63)
\]

in the Minkowski space time is invariant under the boost. From the expression in Eq. (1.63), we can also see that it could be positive, negative or zero. This sign which is defined as

\[
\begin{cases}
(\Delta s)^2 > 0 & \text{timelike} \\
(\Delta s)^2 = 0 & \text{lightlike} \\
(\Delta s)^2 < 0 & \text{spacelike}
\end{cases}
\]

**Timelike:** we can find an inertial frame in which the two events occur at the same spatial coordinate. **Spacelike:** we can find an inertial reference frame on which the two events occur at the same time coordinate.

Before we see what is the lightcone is let’s consider a cone in the Euclidean geometry. Let’s consider the cone shown in Fig. A point on the surface of the cone with coordinates \((x, y, z)\) can be expressed as

\[
x = \left( \frac{h - u}{h} \right) r \cos(\theta), \quad (1.64)
\]

\[
y = \left( \frac{h - u}{h} \right) r \sin(\theta), \quad (1.65)
\]

\[
z = z. \quad (1.66)
\]

for \( z \in [0, h] \) and \( \theta \in [0, 2\pi) \). The opening angle \( \theta \) of a right cone is the vertex angle made by a cross section through the apex and center of the base. For a cone of
height $h$ and radius $r$, it is given by

$$\vartheta = 2 \sin^{-1} \left[ \frac{r}{h} \right]$$

(1.67)

Then the equation of a cone in Cartesian coordinates can be obtained by adding the squares

$$x^2 + y^2 + z^2 = \left( \frac{h - z_0}{h} \right)^2 r^2 + z^2$$

Introducing the constant

$$c = \frac{r}{h}$$

we may write

$$\frac{x^2 + y^2}{c^2} = (h - z)^2$$

Let’s consider the Minkowski spacetime with axes $(ct, x, y)$ with the $z$-axis omitted. Then the interval

$$(\Delta s)^2 = (c \Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

(1.68)
becomes
\[(\Delta s)^2 = (c\Delta t) - (\Delta x)^2 - (\Delta y)^2\]  \hspace{1cm} (1.69)

For a lightlike \((\Delta s)^2 = 0\), we find
\[\frac{(\Delta x)^2 + (\Delta y)^2}{c^2} = (\Delta t)^2\]  \hspace{1cm} (1.70)

which is equation of a cone and it is known as the lightcone. We note that
\[(\Delta s)^2 > 0 \hspace{0.5cm} \text{Outside the light cone (timelike)}\]
\[(\Delta s)^2 < 0 \hspace{0.5cm} \text{Inside the lightcone (spacelike)}\]
\[(\Delta s)^2 = 0 \hspace{0.5cm} \text{On the lightlike (lightlike)}\]

Now let’s omit both \(y\) and \(z\) axis and consider four different events, Event 1, Event 2, Event 3, and Event 4 described by two coordinates \((ct_1, x_1),(ct_2, x_2),(ct_3, x_3)\), and \((ct_4, x_4)\).

The physical interpretation of spacelike \((\Delta s)^2 < 0\) and timelike \((\Delta s)^2 > 0\) is what we stated earlier. When the interval is Timelike, we can find an inertial frame in which the two events occurs on the same spacial coordinate; and when it is Spacelike, we can find an inertial reference frame on which the two events occur at the same time coordinate. In order to understand this let’s consider the boost \(S'\) with both the \(y'\) and \(z'\) coordinates omitted. In Fig. 1.5 we can see that in the reference frame \(S'\) Event 1 and Event 2 which are timelike and occurred at different time and different place in reference frame, we can see that on the boost frame both have the same spacial coordinates. This meant we found an inertial reference frame (the boost) on which the two events occurred at the same point in space. On the otherhand Event 1 and Event 4 which are spacelike events on the \(S\) inertial reference frame, on the \(S'\) frame both events have the same time coordinates. This means we are able to find an inertial reference frame on which Event 1 and Event 4 occurred at the same time.

### 1.4 Length contraction and time dilation

Let’s assume that the boost \(S\) be the Orion spacecraft traveling in the deep outer space with a constant velocity, \(v\), away from our planet (Inertial reference frame \(S\)) as shown in Fig. ?? . Suppose this spacecraft carries a couple leaving the this planet behind. The couple are at rest in \(S'\) standing with their head pointing in the positive \(x\) direction. The woman measures the man’s height (proper length) and found it to be \(l_0\), that we may express it as
\[l_0 = x'_2 - x'_1\]  \hspace{1cm} (1.72)
Fig. 1.4  Spacetime diagram illustrating the lightcone of Event 1 in relation of three other events occurred inside, on, and outside the cone.

Using the Lorentz transformation

\[ x' = \gamma (x - c\beta t), \quad (1.73) \]

we have

\[ l_0 = \gamma (x_2 - c\beta t_2) - \gamma (x_1 - c\beta t_1) \quad (1.74) \]

\[ l_0 = \gamma [(x_2 - x_1) - c\beta (t_2 - t_1)] \quad (1.75) \]

Suppose the man took the reading of \( x_2 \) and \( x_1 \) at the same time

\[ t_2 = t_1 = t \]
then we find

\[ l_0 = \gamma l \Rightarrow l = l_0 \sqrt{1 - \frac{u^2}{c^2}}, \]  

(1.76)

where

\[ l = x_2 - x_1, \]

is the height of the man measured by an observer (\textit{may be an X-girl friend...}) on the earth (the \( S \) inertial reference frame). Obviously, the observer on earth measures the Man’s contracted height (\textit{length}) by a factor of \( \gamma \).

Suppose after some time (as measured in the couple’s inertial reference frame) during their travel in space the man delivered their firstborn baby. As it is measured by the man with a clock aboard the spacecraft and at rest, the beginning of the labor and the arrival of their firstborn baby girl marked by her first cry are separated by a time interval \( T_0 \). Suppose we call the beginning of the labor Event 1 and the arrival
of the baby Event 2, for an observer on Earth ("the X" on the S inertial reference frame), these two events are recorded at

\[ t_1 = \gamma \left( t'_1 + vx'_1 \right), \]

\[ t_2 = \gamma \left( t'_2 + vx'_2 \right) = \gamma \left( t'_1 + T_0 + vx'_2 \right), \]

(1.77)

where we used inverse transformation in Eq. (1.43) where \( v \) is replace by \( -v \). The time interval, \( T \), between these two events as measured by a clock on \( S \) is given by

\[ T = t_2 - t_1 = \gamma \left( T_0 + v (x'_2 - x'_1) \right), \]

(1.78)

Suppose the man on \( S' \) recorded the two events at the same place

\[ x'_2 = x'_1 = x \]

then

\[ T = \gamma T_0 = \frac{T_0}{\sqrt{1 - \frac{v^2}{c^2}}} \]

(1.80)

The result in Eq. (1.80) shows that the moving clock ticks more slowly by a factor of \( \sqrt{1 - \frac{v^2}{c^2}} \).

**Homework Problem 4**: Suppose the couples on the Orion spacecraft celebrated their daughter sweet sixteen birthday as measured by a clock on board the spacecraft (\( S' \)). Suppose she is about 1.6m tall as measured by her her parents. Assume the spacecraft is traveling with constant velocity \( v = 0.8c \), where \( c \) is the speed of light in vacuum.

(a) What would be the age of the boy as measured by "the X" on earth \( S \) (\( S \) inertial reference frame).

(b) How tall is the girl as measured by "the X" on earth, (\( S \) inertial reference frame).

### 1.5 Invariant hyperbolae

Let’s consider the Minkowski spacetime

\[ (\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \]

(1.81)

and for \( (\Delta y)^2 = (\Delta z)^2 = 0 \), we have

\[ (\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 \]

(1.82)

We can calibrate the length scale the \( ct \) and \( x \) axis Minkowski spacetime so that one may write for spacelike or timelike in both \( S \) and \( S' \) inertial reference frames

\[ (c\Delta t)^2 - (\Delta x)^2 = \pm 1 \]

(1.83)
Now if we plot these equations (using Mathematica) we find what is shown in Fig. 1.6. If we call the point on the $(ct = 0, x = 0), (ct = 0, x = 1)$, and $(ct = 1, x = 0)$, and be $O$, $A$, and $B$ respectively. $OA$ represents a unit length and $OB$ a unit time on the $S$ frame (Fig. 1.7). We know that the interval is invariant in the Minkowski spacetime. This means if we rotate the $ct$ axis clockwise and $x$ axis counterclockwise by $\psi = \sinh^{-1} \left[ v/c \right]$, we find the boost (the $S'$ inertial reference frame) with coordinates $ct'$ and $x'$ on which the interval remains invariant

\[(c\Delta t')^2 - (\Delta x')^2 = \pm 1 \tag{1.84}\]

We can easily see length contraction and time dilation from this invariance and Fig. 1.7

Consider a proper length on the $S'$ inertial reference frame which is $OC$. Now draw a line of constant $x'$ through point $C$ that indicates a unit length on the $S'$ $(x' = 1)$. As we can see from the diagram in Fig. 1.7, this line intersects with the $x$-axis on the $S$ inertial reference frame at a point where $x < 1$. In a similar manner let’s consider a unit $OD$ on the $S'$. Draw a line through point D with constant $ct'$ $(ct' = 1)$. As you can see this line intersects with $ct$ axis at a point less than a unit $(ct' < 1)$. This indicates time dilation on the $S'$ reference frame.
Consider two events (Event 1 and Event 2) as shown in the Fig. (a). The line joining these two events in the Minkowski spacetime is known as the worldline. It could be a straight or wiggly line. If these two events are separated by an infinitesimally interval, $ds$, which is known as the line element of the Minkowski space time and the two events are related in this spacetime by $(t, x, y, z)$ and $(t+dt, x+dx, y+dy, z+dz)$, we can write

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad (1.85)$$

The invariant interval between Event 1 and Event 2 along an arbitrary path (straight or wiggly) is given by

$$\Delta s = \int_{1}^{2} ds = \int_{1}^{2} \sqrt{c^2 dt^2 - dx^2 - dy^2 - dz^2} \quad (1.86)$$
From our discussion in the previous sections we recall

\[
\begin{cases}
(\Delta s)^2 > 0 \text{ timelike} \\
(\Delta s)^2 = 0 \text{ lightlike} \\
(\Delta s)^2 < 0 \text{ spacelike}
\end{cases}
\]

and the invariant interval in the Minkowski spacetime would become

\[
\Delta s = \int_1^2 ds = \begin{cases} 
\text{real for time like} \\
\text{0 for lightlike} \\
\text{imaginary for spacelike}
\end{cases}
\]

This fact that we get an imaginary for \(\Delta s\) in the case of spacelike, it means the worldline must lie within the lightcone as shown in Fig. 1.8. This is also required by relativistic mechanics as it prohibits the acceleration of a massive particle to speeds greater than or equal to the speed of light \(c\).

![Fig. 1.8](image)

Fig. 1.8 The worldline of a photon (solid line) and a massive particle (broken line).

The proper time \(\tau\) is used to express the worldline of a massive particle. It is defined as

\[
c^2 d\tau^2 = ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2
\]

(1.87)

From

\[
T = \gamma T_0 = \frac{T_0}{\sqrt{1 - \frac{v^2}{c^2}}}
\]

(1.88)
we may write the proper time as
\[ d\tau = \sqrt{1 - \frac{v^2}{c^2}} dt = \frac{dt}{\gamma_v}. \] (1.89)

Integrating the proper time between two events Event 1 and Event 2, we find
\[ \Delta \tau = \int_1^{t_2} \sqrt{1 - \frac{v^2}{c^2}} dt. \] (1.90)

This shows that if the particle is at rest, \( v = 0 \), we find
\[ \Delta \tau = \Delta t. \] (1.91)

It means the proper time \( \tau \) is just the coordinate time measured by clocks at rest with \( S \). If at any instant in the history of the particle we introduce an instantaneous rest frame \( S' \) such that the particle is momentarily at rest in the \( S' \), then the proper time \( \tau \) is simply the time recorded by the clock that moves along with particle. Therefore, the proper time is an invariantly defined quantity.

1.7 The Doppler effect

Let’s consider an observer \( O \) on an inertial reference frames \( S \). A second observer \( O' \) holding a laser pointer pointed along the positive \( x \) axis on the boost (the \( S' \) inertial reference frame moving along the positive \( x \) direction). Suppose the observer \( O' \) witnessed the two successive wavecrest (which we call Event 1 and Event 2) of a photon emitted from the laser pointer shown in Fig. 1 at \((t', x')\) and \((t' + \Delta t', x' + \Delta x')\). Let’s assume that observer \( O \) on \( S \) has received these two successive wavecrest at \((t, x)\) and \((t + \Delta t, x)\) (which we refer as Event 3 and Event 4). From Eq. (1.90) the proper time (the time measured by the observer \( O' \), will then be
\[ \Delta \tau_{21} = \sqrt{1 - \frac{v^2}{c^2}} \Delta t'. \] (1.92)

On the other hand the observer on the rest inertial reference frame \( S \) for which \( v = 0 \), Eq. (1.90) give us a proper time
\[ \Delta \tau_{43} = \Delta t. \] (1.93)

Along the world line joining Event 1 to Event 3 and Event 2 to Event 4, the interval is zero since we are considering emission of a photon (Lightlike)
\[ c^2 d\tau^2 = ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = 0. \] (1.94)

For \( dy = dz = 0 \), we have
\[ c^2 dt^2 - dx^2 = 0. \] (1.95)
so that for the worldline joining event Event 1 and Event 3, we can write
\[ \int_{1}^{3} cdt = - \int_{1}^{3} dx \Rightarrow \int_{t'}^{t} cdt = - \int_{x'}^{x} dx. \quad (1.96) \]

The minus sign, as it can be seen from the Fig., is because photon is traveling in the negative x-direction. Similarly, for the worldline connecting Event 2 and Event 4, we can write
\[ \int_{2}^{4} cdt = - \int_{2}^{4} dx \Rightarrow \int_{t'+\Delta t}^{t} cdt = - \int_{x'+\Delta x'}^{x} dx. \quad (1.97) \]

which we may rewrite as
\[ \int_{t}^{t'} cdt + \int_{t}^{t'} cdt + \int_{t'}^{t + \Delta t} cdt = - \int_{x}^{x'} dx - \int_{x'}^{x + \Delta x'} dx. \quad (1.98) \]

or
\[ \int_{t}^{t'} cdt + \int_{t}^{t + \Delta t} cdt - \int_{t'}^{t + \Delta t} cdt = - \int_{x}^{x'} dx - \int_{x'}^{x + \Delta x'} dx. \quad (1.99) \]

Using the result in Eq. (1.96), we find
\[ \int_{t}^{t + \Delta t} cdt - \int_{t'}^{t + \Delta t} cdt = - \int_{x'}^{x + \Delta x'} dx \quad (1.100) \]

that leads to
\[ c(\Delta t - \Delta t') = \Delta x'. \quad (1.101) \]

From Eq. (1.101) we find the relation between the proper time for an observer \( O \) and \( O' \)
\[ \Delta t = \left( 1 + \frac{1}{c} \Delta x' \right) \Delta t' = \left( 1 + \frac{v}{c} \right) \Delta t'. \quad (1.102) \]

Now combining the Eqs (1.92), (1.93) and (1.102), we have
\[ \Delta \tau_{21} = \sqrt{1 - \frac{v^2}{c^2}} \Delta t'. \quad (1.103) \]

On the other hand the observer on the rest inertial reference frame \( S \) for which \( v = 0 \), Eq. (1.90) give us a proper time
\[ \frac{\Delta \tau_{43}}{\Delta \tau_{21}} = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}} \Delta t'} = \frac{(1 + \frac{v}{c}) \Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{(1 + \frac{v}{c})}{(1 + \frac{v}{c}) \left( 1 - \frac{v}{c} \right)}. \quad (1.104) \]
Thus the time interval between two successive wavecrests as observed by an observer on \( S \) (\( \Delta \tau_{43} \)) and an observer on \( S' \) (\( \Delta \tau_{21} \)) are related by

\[
\frac{\Delta \tau_{43}}{\Delta \tau_{21}} = \sqrt{\frac{1 + \frac{\nu}{c}}{1 - \frac{\nu}{c}}}.
\]  

(1.105)

Note that this times are the period of the emitted photon that can be related to the corresponding frequencies for an observer \( O \) (\( \nu \)) and and \( O' \) (\( \nu' \)) by

\[
\Delta \tau_{43} = \frac{1}{\nu}, \Delta \tau_{21} = \frac{1}{\nu'}
\]

we find

\[
\frac{\nu}{\nu'} = \sqrt{\frac{1 - \frac{\nu}{c}}{1 + \frac{\nu}{c}}}.
\]  

(1.106)

This is the familiar \textit{Doppler-effect formula}.

### 1.8 Addition and acceleration in special relativity

Consider a particle that its position is described by the coordinates \((x, y, z)\) are function of time \( t \) on the \( S \) frame. The speed of this particle \((u_x, u_y, u_z)\) on this frame can then be expressed as

\[
u_x = \frac{dx}{dt'}, u_y = \frac{dy}{dt'}, u_z = \frac{dz}{dt'}.
\]  

(1.107)

Using the inverse Lorentz transformation

\[
te = \gamma (ct' + \beta x'),
\]

(1.108)

\[
x = \gamma (x' + c\beta t'),
\]

(1.109)

\[
y' = y,
\]

(1.110)

\[
z' = z.
\]

(1.111)

we can write

\[
c dt = \gamma (cdt' + \beta dx'),
\]

(1.112)

\[
dx = \gamma (dx' + c\beta dt'),
\]

(1.113)

\[
dy = dy',
\]

(1.114)

\[
dz = dz',
\]

(1.115)
so that

\[
\begin{align*}
\frac{dx}{cdt} &= \frac{dx'}{ct'} + \beta \frac{dt'}{ct'}, \\
\frac{dy}{cdt} &= \frac{dy'}{ct'}, \\
\frac{dz}{cdt} &= \frac{dz'}{ct'}.
\end{align*}
\] (1.116)

This can be rewritten as

\[
\begin{align*}
\frac{dx}{dt} &= \frac{dx'}{ct'} + \beta \frac{dt'}{ct'}, \\
\frac{dy}{dt} &= \frac{dy'}{ct'}, \\
\frac{dz}{dt} &= \frac{dz'}{ct'}.
\end{align*}
\] (1.117)

\[
\begin{align*}
\frac{dx'}{dt'} &= \frac{u'}{c^2} + \beta \frac{u'}{c^2}, \\
\frac{dy'}{dt'} &= \frac{y'}{c^2}, \\
\frac{dz'}{dt'} &= \frac{z'}{c^2}.
\end{align*}
\] (1.118)

The inverse transformation can be obtained by replacing \( v \) with \(-v\)

\[
\begin{align*}
u_x' &= \frac{u_x' - v}{1 - \frac{v^2}{c^2} u_x'}, \\
u_y' &= \frac{u_y' \sqrt{1 - \frac{v^2}{c^2}}}{1 - \frac{u_x'^2}{c^2}}, \\
u_z' &= \frac{u_z' \sqrt{1 - \frac{v^2}{c^2}}}{1 - \frac{u_x'^2}{c^2}}.
\end{align*}
\] (1.119)

**Homework Problem 5**: Consider three inertial reference frames \( S, S', \) and \( S'' \). Suppose \( S' \) is related to \( S \) by a boost of speed \( v \) in the \( x \) direction and that \( S'' \) is related to \( S' \) by a boost of speed \( u' \) in the \( x' \)-direction. Using the rapidity parameter defined as

\[
\psi_v = \tanh^{-1} \left( \frac{v}{c} \right), \quad \psi_{u'} = \tanh^{-1} \left( \frac{u'}{c} \right).
\] (1.120)

show that

(a)

\[
\begin{align*}
t'' &= ct \cosh (\psi_v + \psi_{u'}) - x \sinh (\psi_v + \psi_{u'}), \\
x' &= -ct \sinh (\psi_v + \psi_{u'}) + x \cosh (\psi_v + \psi_{u'}), \\
y' &= y, \\
z' &= z.
\end{align*}
\]
(b) \[
u = ct \tanh (\psi_v + \psi_{v'}) = c \frac{\tanh \psi_v + \tanh \psi_{v'}}{1 + \tanh \psi_v \tanh \psi_{v'}} = \frac{u' + v}{1 + u'v/c^2}
\]
Chapter 2

Manifolds

2.1 What is a Manifold?

Consider a ridged meterstick pinned to the positive z-axis on the \( x - y - z \) Cartesian coordinate system as shown in Fig. 2.1. Suppose this coordinate system is at rest at initial time \( t = 0 \) and suddenly begins to rotate. It is free to rotate about the x-axis, y-axis, or z-axis. Let’s say we want to describe the angular position of the center of mass of the meter stick over a period of time, \( \tau = 10s \) with a time interval of 2s. How many independent parameters that depend on time do we need to describe the angular position of the center of mass of the meterstick relative to its initial position at \( t = 0 \)? Well the answer is simple. We need three independent parameters the Euler angles, which I write as \((\alpha^1, \alpha^2, \alpha^3)\) which describes the rotation about the x-, y-, and z-axes at a given instant of time. Then over the 10 second interval we have a set that consist of 5 points

\[
\{ [\alpha^1 (2), \alpha^2 (2), \alpha^3 (2)], [\alpha^1 (4), \alpha^2 (4), \alpha^3 (4)], [\alpha^1 (6), \alpha^2 (6), \alpha^3 (6)], \\
[\alpha^1 (8), \alpha^2 (8), \alpha^3 (8)], [\alpha^1 (10), \alpha^2 (10), \alpha^3 (10)] \}
\]

We can make the time interval infinitesimal to continuously describe the position of the center of mass of the meterstick. The resulting set of points form a Manifold of dimension three.

Let’s consider another example of a Manifold. In classical mechanics you have study the phase space. In this space you can describe the state of a particle over a period of time using the three coordinates of space locating the position of the particle and the three coordinate of speed (or momentum) describing how fast the particle is moving at a given instant of time. Then over the 10 second interval we have a set that consist of 5 points

\[
\{ [x^1 (t), y^2 (t), x^3 (t), x^4 (t), x^5 (t), x^6 (t)] \}
\]

So when you describe the state of the particle say from \( t = 0 \) to \( t = t_0 \) you can use infinitesimal time interval so that you will have a set of points that can be parameterized continuously in terms of \((x^1 (t), x^2 (t), ..., x^6 (t))\). These set of points form a Manifold of dimension six.
Therefore, a manifold is any set that can be continuously parameterized. The number of independent parameters is the dimension of the Manifold. A Manifold is continuous and differentiable.

A manifold is Continuous: if you pick any point, $p$, on the Manifold and you can find another points whose coordinates differ infinitesimally from the point $p$.

A manifold is differentiable: if you pick any point, $p$, on the Manifold and you can find a scalar field $\phi$ that is differentiable at that point $p$.

Coordinates of a Manifold $M$: a point in an N-Dimensional Manifold is represented by the coordinates $(x^1, x^2, x^3, ..., x^N)$ which we represent by $x^a$ where it is understood that $a = 1, 2, 3...N$.

Degeneracy in a Manifold: sometimes it may not be possible to cover the whole manifold with only one none-degenerate coordinate system. Example is a plane in polar coordinate system $(\rho, \varphi)$. A plane is a two dimension Manifold. (called $R^2$). A plane in polar coordinates has a degeneracy at the origin since $\varphi$ is indeterminate at the origin.

Coordinate patches: these are coordinate systems that covers a portion of the Manifold where we have degeneracy. For example the surface of a sphere is a two dimensional Manifold (called $S^2$). It can be described by two independent parameters $(\theta, \varphi)$ except at two points on the Manifold. These are the north and south pole where $\varphi$ is indeterminate (or there is degeneracy). There is no a coordinate system that covers the entire sphere without running into these two degenerate points. In this case the smallest number of patches we need is two.

Atlas: an atlas is a set of coordinate patches that coves the whole Manifold.

2.2 Curves and surfaces in a Manifold

Both curves and surfaces on a Manifold are defined parametrically. That means we use some common parameters. For example, in the phase space that we saw earlier, the curves can be defined in terms of the time parameter, $t$. Generally, we use a parameter $u$ to define a curve.

A curve: a curve in a Manifold of dimension $N$ is define by a parametric equation

$$x^a = x^a(u), \text{ where } a = 1, 2, 3...N.$$  \hfill (2.1)

A surface: a surface in a Manifold of dimension $N$ (which also referred as a submanifold) has $M$ degrees of freedom which is always less than the dimension of the Manifold ($M < N$) and therefore it depends on $M$ parameter that we represent by $(u^1, u^2, u^3, ..., u^M)$ and is define by the parametric equation

$$x^a = x^a(u^1, u^2, u^3, ..., u^M), \text{ where } a = 1, 2, 3...N.$$  \hfill (2.2)

Hypersurface: a surface of dimension $M$ in a Manifold of dimension $N$ with $M = N - 1$. In this case the $N - 1$ parameters can be eliminated from the $N$ equations
and you can find one equation

\[ f \left( x^1, x^2, x^3, \ldots, x^N \right) = 0 \]  

(2.3)

**Example 2.1** Let’s consider the 3-D Euclidean Manifold. A sphere is Hypersurface since \( M = 2 \). A point on a sphere is defined by

\[ x^2 + y^2 + z^2 = a^2 \]  

(2.4)

where \( a \) is the radius of the sphere. We note that in this case the surface of the sphere is a Hypersurface that can be defined by the equation

\[ f \left( x^1, x^2, x^3 \right) = (x^1)^2 + (x^2)^2 + (x^3)^2 - a^2 = 0, \]  

(2.5)

where we used \((x^1, x^2, x^3)\) for \((x, y, z)\).

Therefore a point is restricted to lie in a hypersurface (submanifold with \( N - 1 \) dimensional a Manifold of dimension \( N \)), then the points coordinate must satisfy Eq. (2.3). We come up with a similar generalization to this for a point that belong to any surface with dimension \( M \) in a Manifold of dimension \( N \) (\( M < N \)).

\[
\begin{align*}
  f_1 \left( x^1, x^2, x^3, \ldots, x^N \right) & = 0, \\
  f_2 \left( x^1, x^2, x^3, \ldots, x^N \right) & = 0, \\
  f_1 \left( x^1, x^2, x^3, \ldots, x^N \right) & = 0, \\
  \vdots & \nonumber \\
  \vdots & \nonumber \\
  f_{N-M} \left( x^1, x^2, x^3, \ldots, x^N \right) & = 0
\end{align*}
\]

(2.6)

### 2.3 Coordinate transformations and summation convention

Let’s consider the 3-D Euclidean Manifold. A point in this Manifold can be represented using Cartesian coordinates \((x, y, z)\) which we may represent by \((x^1, x^2, x^3)\). This same point can also be represented using spherical coordinates \((r, \theta, \varphi)\) which similarly represent it by \((x^1, x^2, x^3)\). Now the question is how we relate the Cartesian coordinates with the spherical coordinates or vice versa. Noting that

\[
\begin{align*}
  r & \rightarrow r \left( x, y, z \right), \text{ or } x'^1 \rightarrow x'^1 \left( x^1, x^2, x^3 \right) \\
  \theta & \rightarrow \theta \left( x, y, z \right), \text{ or } x'^2 \rightarrow x'^2 \left( x^1, x^2, x^3 \right) \\
  \varphi & \rightarrow \varphi \left( x, y, z \right), \text{ or } x'^3 \rightarrow x'^3 \left( x^1, x^2, x^3 \right)
\end{align*}
\]

(2.7)
or

\[ x \rightarrow x(r, \theta, \varphi), \text{ or } x^1 \rightarrow x^1(x^1, x^2, x^3) \]  
\[ y \rightarrow y(r, \theta, \varphi), \text{ or } x^2 \rightarrow x^2(x^1, x^2, x^3) \]  
\[ z \rightarrow z(r, \theta, \varphi), \text{ or } x^3 \rightarrow x^3(x^1, x^2, x^3) \]  

Suppose we have a function, \( f \), which can be expressed as \( f(x^1, x^2, x^3) \) and also as \( f(x^1, x^2, x^3) \), then we recall from Theoretical Physics I,

\[
\frac{\partial f}{\partial x^1} = \frac{\partial f}{\partial x^1} \frac{\partial x^1}{\partial x^1} + \frac{\partial f}{\partial x^2} \frac{\partial x^2}{\partial x^1} + \frac{\partial f}{\partial x^3} \frac{\partial x^3}{\partial x^1} \tag{2.9}
\]

\[
\frac{\partial f}{\partial x^2} = \frac{\partial f}{\partial x^1} \frac{\partial x^1}{\partial x^2} + \frac{\partial f}{\partial x^2} \frac{\partial x^2}{\partial x^2} + \frac{\partial f}{\partial x^3} \frac{\partial x^3}{\partial x^2} \tag{2.10}
\]

\[
\frac{\partial f}{\partial x^3} = \frac{\partial f}{\partial x^1} \frac{\partial x^1}{\partial x^3} + \frac{\partial f}{\partial x^2} \frac{\partial x^2}{\partial x^3} + \frac{\partial f}{\partial x^3} \frac{\partial x^3}{\partial x^3} \tag{2.11}
\]

which we can express in a Matrix form as

\[
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} \tag{2.12}
\]

For the inverse case we can easily show

\[
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} \tag{2.13}
\]

so that

\[
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} \tag{2.14}
\]

can be written as

\[
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial f}{\partial x^1} \\
\frac{\partial f}{\partial x^2} \\
\frac{\partial f}{\partial x^3}
\end{bmatrix} \tag{2.15}
\]
There follows that
\[
\begin{pmatrix}
\frac{\partial x^1}{\partial x^0} & \frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^0} & \frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^0} & \frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{pmatrix}
\begin{pmatrix}
\frac{\partial x^0}{\partial x^0} & \frac{\partial x^0}{\partial x^1} & \frac{\partial x^0}{\partial x^2} & \frac{\partial x^0}{\partial x^3} \\
\frac{\partial x^1}{\partial x^0} & \frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^0} & \frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^0} & \frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{pmatrix}
= 1
\]  
(2.16)

This means the matrix
\[
A^{-1} =
\begin{pmatrix}
\frac{\partial x^0}{\partial x^1} & \frac{\partial x^0}{\partial x^2} & \frac{\partial x^0}{\partial x^3} \\
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{pmatrix}
\]  
(2.17)

must be the inverse matrix for
\[
A =
\begin{pmatrix}
\frac{\partial x^1}{\partial x^0} & \frac{\partial x^2}{\partial x^0} & \frac{\partial x^3}{\partial x^0} \\
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^2}{\partial x^1} & \frac{\partial x^3}{\partial x^1} \\
\frac{\partial x^1}{\partial x^2} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^3}{\partial x^2} \\
\frac{\partial x^1}{\partial x^3} & \frac{\partial x^2}{\partial x^3} & \frac{\partial x^3}{\partial x^3}
\end{pmatrix}
\]  
(2.18)

so that
\[
A^{-1}A = AA^{-1} = 1
\]  
(2.19)

We note that the transpose of the Matrix A, \(A^T\) is given by
\[
A^T =
\begin{pmatrix}
\frac{\partial x^0}{\partial x^1} & \frac{\partial x^0}{\partial x^2} & \frac{\partial x^0}{\partial x^3} \\
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{pmatrix}
\]  
(2.20)

Similarly for the inverse matrix, the transpose matrix which we express as
\[
\frac{\partial x^a}{\partial x^b} =
\begin{pmatrix}
\frac{\partial x^0}{\partial x^0} & \frac{\partial x^0}{\partial x^1} & \frac{\partial x^0}{\partial x^2} & \frac{\partial x^0}{\partial x^3} \\
\frac{\partial x^1}{\partial x^0} & \frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^0} & \frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^0} & \frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{pmatrix}
\]  
(2.21)

is the transformation matrix that transforms the coordinates \((x^1, x^2, x^3)\) to \((x'^1, x'^2, x'^3)\).

For a Manifold of dimension \(N\), the transformation matrix is given by
\[
\frac{\partial x^a}{\partial x^b} =
\begin{pmatrix}
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \cdots & \frac{\partial x^1}{\partial x^N} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \cdots & \frac{\partial x^2}{\partial x^N} \\
& \ddots & & \vdots \\
\frac{\partial x^N}{\partial x^1} & \frac{\partial x^N}{\partial x^2} & \cdots & \frac{\partial x^N}{\partial x^N}
\end{pmatrix}
\]  
(2.22)
As we recall from theoretical physics, a Matrix is invertible provided the determinant is different from zero. Therefore, the inverse transformation is possible provided the determinant of the transformation matrix which is known as the Jacobian, $J$ different from zero

$$J = \det \left[ \frac{\partial x^a}{\partial x^b} \right] = \left| \begin{array}{cccc} \frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \cdots & \frac{\partial x^1}{\partial x^N} \\ \frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \cdots & \frac{\partial x^2}{\partial x^N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x^N}{\partial x^1} & \frac{\partial x^N}{\partial x^2} & \cdots & \frac{\partial x^N}{\partial x^N} \end{array} \right|$$ (2.23)

The inverse transformation Matrix can be written as

$$\frac{\partial x^a}{\partial x^b} = \left[ \begin{array}{cccc} \frac{\partial x^1}{\partial x^a} & \frac{\partial x^1}{\partial x^b} & \cdots & \frac{\partial x^1}{\partial x^N} \\ \frac{\partial x^2}{\partial x^a} & \frac{\partial x^2}{\partial x^b} & \cdots & \frac{\partial x^2}{\partial x^N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x^N}{\partial x^a} & \frac{\partial x^N}{\partial x^b} & \cdots & \frac{\partial x^N}{\partial x^N} \end{array} \right]$$ (2.24)

and the Jacobian $J'$

$$J' = \det \left[ \frac{\partial x^a}{\partial x^b} \right] = \left| \begin{array}{cccc} \frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \cdots & \frac{\partial x^1}{\partial x^N} \\ \frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \cdots & \frac{\partial x^2}{\partial x^N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x^N}{\partial x^1} & \frac{\partial x^N}{\partial x^2} & \cdots & \frac{\partial x^N}{\partial x^N} \end{array} \right|$$ (2.25)

We note that

$$\frac{\partial x^a}{\partial x^i} = \frac{\partial x^a}{\partial x^1} \frac{\partial x^1}{\partial x^i} + \frac{\partial x^a}{\partial x^2} \frac{\partial x^2}{\partial x^i} + \cdots + \frac{\partial x^a}{\partial x^N} \frac{\partial x^N}{\partial x^i} = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} \frac{\partial x^b}{\partial x^i}$$ (2.26)

Noting that for independent coordinates

$$\frac{\partial x^a}{\partial x^c} = \begin{cases} 0, & a \neq c \\ 1, & a = c \end{cases}$$ (2.27)

we can generally write

$$\sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} \frac{\partial x^b}{\partial x^c} = \delta^a_c.$$ (2.28)
Consider two points \( P \) and \( Q \) on the Manifold with dimension \( N \). Suppose these points are separated by infinitesimal interval so that if the coordinates of \( P \) is \( x^a \) and that of \( Q \) is \( x^a + dx^a \), then we can write
\[
dx^a = \frac{\partial x^a}{\partial x^1} dx^1 + \frac{\partial x^a}{\partial x^2} dx^2 \ldots \frac{\partial x^a}{\partial x^N} dx^N = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} dx^b, \tag{2.29}\]
where the summation is evaluated at \( P \). Similarly for the interval between \( x^a \) and \( x^a + dx^a \), in the none-primed coordinate system, we can write
\[
dx^a = \frac{\partial x^a}{\partial x^0} dx^0 + \frac{\partial x^a}{\partial x^0} dx^2 \ldots \frac{\partial x^a}{\partial x^0} dx^N = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^0} dx^b, \tag{2.30}\]
here also the summation is evaluated at \( P \).

*Einstein’s summation convention*: whenever an index occurs twice in an expression, once as subscript and once as a superscript, imply a summation over the index. An index should not occur more than twice. For example, according to Einstein’s summation convention, the summations in Eq. (2.29) and (2.30) can be expressed
\[
dx^a = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} dx^b = \frac{\partial x^a}{\partial x^0} dx^b \tag{2.31}\]
and
\[
dx^a = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^0} dx^b = \frac{\partial x^a}{\partial x^0} dx^b. \tag{2.32}\]
The index \( a \), which is known as the *free index*, can take any value from 1 to \( N \). The index \( b \), which is known as the *dummy index* and it must be summed up from 1 to \( N \).

### 2.4 The Riemannian geometry

*The local geometry of a Manifold*: The local geometry of a manifold is determined by defining the invariant ‘distance’ or (as we say in the Minkowski spacetime Manifold) the interval \( ds \) between points \( P \) with coordinate \( x^a \) and \( Q \) with coordinates \( x^a + dx^a \). This distance can be assigned in general to be a well-behaved function \( f(x^a, dx^a) \) of the coordinates \( x^a \) and \( dx^a \)
\[
ds^2 = f(x^a, dx^a). \tag{2.33}\]
In the general theory of relativity we are confined to a Manifold where the interval \( ds \) can be described by the equation of the form
\[
ds^2 = \sum_{a=1}^{N} \sum_{b=1}^{N} g_{ab}(x) dx^a dx^b, \tag{2.34}\]
or simply using the *Einstein’s summation convention*

\[ ds^2 = g_{ab}(x) \, dx^a dx^b. \]  

(2.35)

A geometry of a Manifold defined by Eq. (2.35) is known as the Riemannian geometry if \( ds^2 > 0 \). As we have seen in the case of Minkowski spacetime manifold the interval \( ds^2 \) can also be negative (spacelike) or zero (lightlike). In such cases the geometry is referred as *pseudo Riemannian geometry* and the manifold can be referred as *pseudo Riemannian* or simply *Riemannian manifold*. The function \( g(x) \) is know as the metric function where \( g_{ab}(x) \) represent the element of a metric tensor.

**Transformation of the interval:** applying the relation in Eq. (2.32), we can express

\[ dx^a = \frac{\partial x^a}{\partial x'^c} \, dx'^c, \quad dx^b = \frac{\partial x^b}{\partial x'^d} \, dx'^d \]  

(2.36)

so that the interval can be transformed as

\[ ds^2 = g_{ab}(x) \, \frac{\partial x^a}{\partial x'^c} \, \frac{\partial x^b}{\partial x'^d} \, dx'^c dx'^d, \]  

(2.37)

or

\[ ds^2 = g'_{cd}(x') \, dx'^c dx'^d, \]  

(2.38)

where

\[ g'_{cd}(x') = g_{ab}(x) \, \frac{\partial x^a}{\partial x'^c} \, \frac{\partial x^b}{\partial x'^d}. \]  

(2.39)

Note that \( x = x(x') \).

### 2.5 Intrinsic and extrinsic geometry and the metric

A given geometry of dimension \( M \) defined by the metric equation

\[ ds^2 = g_{ab}(x) \, dx^a dx^b. \]  

(2.40)

and embedded in a higher dimension manifold of dimension \( N, (N > M) \) is said to be:

(a) **Intrinsic:** when the geometry remains unchanged as viewed in the higher dimensional manifold.

(b) **Extrinsic:** when the geometry is different as viewed in the higher dimensional manifold.

Before we see two examples that have intrinsic and extrinsic geometries let’s consider the metric equation for a curved surface in three dimensional Euclidean space.
We know that this surface can be defined by a function \( g \), that depends on \((x, y, z)\) in Cartesian, \((r, \theta, \varphi)\) in spherical, or \((r, \varphi, z)\) in cylindrical coordinates. Suppose we represent these coordinates by \((x^1, x^2, x^3)\), then we may write the function that defines the surface as

\[
s(\mathbf{x}) = g(\mathbf{x}) = g(x^1, x^2, x^3) \tag{2.41}
\]

Now let’s consider a point \( P \) on this surface that has coordinates \((x^1, x^2, x^3)\). Suppose we consider another point \( Q \) with coordinates \((x^1 + dx^1, x^2 + dx^2, x^3 + dx^3)\) we may define the surface between these two points as. This displacement is just the differential of Eq. (2.41)

\[
ds^2 = \left( \frac{\partial g}{\partial x^1} dx^1 + \frac{\partial g}{\partial x^2} dx^2 + \frac{\partial g}{\partial x^3} dx^3 \right) \cdot \left( \frac{\partial g}{\partial x^1} dx^1 + \frac{\partial g}{\partial x^2} dx^2 + \frac{\partial g}{\partial x^3} dx^3 \right) \tag{2.42}
\]

or using Einstein’s summation convention the geometry of the surface between the two points

\[
ds^2 = (ds) \cdot (ds) = \left( \sum_{a=1}^{3} \frac{\partial g}{\partial x^a} dx^a \right) \cdot \left( \sum_{b=1}^{3} \frac{\partial g}{\partial x^b} dx^b \right) = \sum_{a=1}^{3} \sum_{b=1}^{3} \frac{\partial g}{\partial x^a} \frac{\partial g}{\partial x^b} dx^a dx^b. \tag{2.43}
\]

Again using Einstein’s summation convention and the notation

\[
g_{ab}(\mathbf{x}) = \frac{\partial g}{\partial x^a} \frac{\partial g}{\partial x^b} \tag{2.44}
\]

we may write

\[
ds^2 = g_{ab}(\mathbf{x}) dx^a dx^b. \tag{2.45}
\]

which is the metric equation that we defined earlier. Note that the metric tensor in this case is a \(3 \times 3\) matrix given by

\[
G = \begin{bmatrix}
\frac{\partial g}{\partial x^1} & \frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^3} \\
\frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^1} & \frac{\partial g}{\partial x^3} \\
\frac{\partial g}{\partial x^3} & \frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^1}
\end{bmatrix} \tag{2.46}
\]

and we can easily see that this matrix is symmetric as

\[
\frac{\partial g}{\partial x^a} \frac{\partial g}{\partial x^b} = \frac{\partial g}{\partial x^b} \frac{\partial g}{\partial x^a} \tag{2.47}
\]

Suppose the coordinates for an orthonormal set like the Cartesian, spherical, or cylindrical, we note that

\[
G = \begin{bmatrix}
\left( \frac{\partial g}{\partial x^1} \right)^2 & 0 & 0 \\
0 & \left( \frac{\partial g}{\partial x^2} \right)^2 & 0 \\
0 & 0 & \left( \frac{\partial g}{\partial x^3} \right)^2
\end{bmatrix} \tag{2.47}
\]
Extrinsic geometry: One simple example of extrinsic geometry is 2-D cylindrical geometry. In order to see that let’s consider a plane geometry in a 3-D Euclidean Manifold. Let’s this plane depends on \((x^1, x^2)\). We recall from theoretical physics I generally a plane:

**Equation of a plane**: If \(\vec{N} = ai + bj + ck\) is normal (perpendicular) to a plane, then the dot product of the vector \(\vec{N}\) and the vector \(\vec{r} - \vec{r}_0\)

\[
\vec{r} - \vec{r}_0 = (x - x_0) \hat{x} + (y - y_0) \hat{y} + (z - z_0) \hat{z}
\]

is zero,

\[
a(x - x_0) + b(y - y_0) + c(z - z_0) = 0
\]

This defines the equation of the plane. It can be rewritten as

\[
a x + b y + c z = d
\]

where

\[
d = ax_0 + by_0 + cz_0
\]

For our notation we may write

\[
a x^1 + b x^2 + c x^3 = d
\]
where
\[ d = ax_0^1 + bx_0^2 + cx_0^3. \] (2.53)

For a plane that depends on only \((x^1, x^2)\), we may write the interval as
\[ ds^2 = (dx^1)^2 + (dx^2)^2. \] (2.54)

Now let’s consider the 2-D cylindrical surface which we can construct using our 2-D plane. Suppose the cylinder has radius \(a\) with its axis along the \(z\)-axis (which I call it \(x^3\)-axis). Using cylindrical coordinates this surface can be defined by the function
\[ g(a, \varphi, z) = g(a, x^2, x^3) \] (2.55)
we may write the interval, just using geometrical visualization, as
\[ ds^2 = a^2 (dx^2)^2 + (dx^3)^2. \] (2.56)

This is a 2-D surface embedded in a 3-D manifold. It has cylindrically curved geometry when it is viewed in this 3-D Euclidean manifold. But you can actually obtain this geometry from the plane geometry from Eq. (2.41) by simply using the transformation
\[ x^1 = ax^2, x^2 = x^3. \]

Such kind of geometry is not intrinsic and it is called extrinsic. Its curvature is extrinsic and is a result of the way it is embedded in the three dimensional space..

**Example 2.3**  Intrinsic geometry: One simple example of intrinsic geometry is 2-D spherical geometry embedded in a 3-D Euclidean Manifold. In Fig. 2.3 we see a 3-D infinitesimal volume. Assume the sphere a radius \(a\). Then the surface defined by a pair of points on this sphere separated by a distance \(ds\) can be expressed as
\[ ds^2 = (a d\theta)^2 + (a \sin(\theta) d\varphi)^2 = a^2 (d\theta)^2 + a^2 \sin^2(\theta) (d\varphi)^2. \] (2.57)

or if we use \((x^2, x^3)\) for \((\theta, \varphi)\) we can write
\[ ds^2 = a^2 (dx^2)^2 + a^2 \sin^2(\theta) (dx^3)^2. \] (2.58)

The is a 2-D surface embedded in a 3-D manifold. You can not obtain this geometry from the plane geometry like the 2-D cylindrical geometry. Such kind of geometry is intrinsic. This means the geometry of a sphere is intrinsic curved because we can not transform Eq.(2.58) to the Euclidean form
\[ ds^2 = (dx^1)^2 + (dx^2)^2 \] (2.59)
over the whole surface by any coordinate transformation. Note that this can be done locally but not for the whole spherical surface.
Example 2.4  Non-Euclidean geometry: