# Contents

I Relativity

1 Special Relativity (SR)
   1.1 Introduction ............................................. 2
   1.2 Lorentz transformations ...................................... 3
      1.2.1 Addition of velocities .................................. 4
      1.2.2 The role of observers in SR ............................ 4
      1.2.3 Simultaneity and causality .............................. 5
      1.2.4 Time dilation and length contraction .................... 5
      1.2.5 Relativistic Doppler .................................... 5
   1.3 Particle mechanics .......................................... 6
      1.3.1 Tensors ................................................. 6
      1.3.2 Displacement, velocity, momentum and acceleration four-vectors .................. 6
   1.4 Fluid mechanics ............................................ 7
      1.4.1 Stress-Energy or Energy-momentum Tensor ................. 7
   1.5 Maxwell’s Equations are Lorentz invariant .................. 8

2 General Relativity (GR)
   2.1 Motivation for GR ........................................... 9
   2.2 The Equivalence Principle (EP) .............................. 9
   2.3 The Weak Field limit of GR .................................. 10
   2.4 Motion of a single particle .................................. 11
   2.5 Motion of a collection of particles: characterizing curvature ....................... 11
   2.6 Field equations ............................................. 12
   2.7 Non-linearity of GR .......................................... 13
Part I
Relativity

1 Special Relativity (SR)

1.1 Introduction

Early experiments and observations:

• \textit{c is finite}. Romer, 1675: Observed Io’s orbital period around Jupiter as Jupiter approached and receded from Earth; concluded that velocity of light is finite.

• \textit{c is independent of the velocity of the emitter}. In an eclipsing binary with equal mass stars maximum redshift of star 1 and maximum blueshift of star 2 are observed to take place at the same time. This challenges the Newtonian prescription for addition of velocities (Galilean transformations).

• \textit{c is independent of the frame from which it is measured}. Michelson-Morley, 1887: looked for ether, but measured same value for \(c\) regardless of the lab frame’s orientation and motion with respect to ether.

• \textit{Maxwell’s equations are not invariant under Galilean transformations, but are invariant under Lorentz transformations}. If Galilean transformations (GT) are the physically correct ones, then Maxwell’s equations are valid only in one special frame, defined by ‘ether’, where \(c\) would take on its speed of light value. In all other frames, GT would make Maxwell’s equations take on a different form in different frames, and \(c\) would have different values, determined by the Galilean law of addition of velocities, \(c' = c + v\). It was generally known that Maxwell’s equations were invariant under Lorentz transformations, but that was considered a mathematical curiosity, rather than physical reality.

Einstein chose to start with the premise that Lorentz transformations did represent physical reality, and that Maxwell’s equations retain the same form irrespective of the reference frame. Conclusion: there is no ether; therefore Galilean transformation, and the corresponding law of addition of velocities \((v = v_1 + v_2)\) does not hold in general. Need to formulate a new theory, where \(c\) is the same in all inertial frames, as it already is in Maxwell’s equations.

Einstein’s 2 postulates:

(1) \textbf{Principle of Relativity}. Physics experiments must be reproducible, therefore physical laws must have the same form in all inertial frames, i.e. physical laws must be form invariant, or covariant. Inertial frames are defined by Newton’s 1st law. Any reference frame moving at constant velocity w.r.t. an inertial frame is itself an inertial frame. In inertial frames free particles travel along straight lines with constant velocity. PR restated: all inertial frames are completely equivalent for the performance of physical experiments.

(2) \textbf{Constancy of} \(c\). Speed of light is constant in vacuum, and is independent of the velocity of the emitter and the absorber. It is like a property of space rather than property of the emitter/absorber.

We will use (2) to construct the basic machinery of SR, in particular, derive transformations, LT, that will allow us to go between frames which are moving with \(v \leq c\) with respect to each other. Once that’s done, (1) and Lorentz transformations can be used to reformulate/redefine concepts like velocity, momentum, acceleration, force, etc. such that they are Lorentz-invariant, and reduce to their pre-SR forms in the low velocity limit.
1.2 Lorentz transformations

We start with the 2nd postulate (speed of light is the same in all frames) and consider propagation of light. Consider 2 events: event 1 is the emission of light signal, and event 2 is the absorption of that light signal. The spatial and temporal locations of these 2 events are diligently recorded by observers in two inertial frames: \( S' \) is moving at \( v \) as seen from \( S \). In frame \( S \) light travels a distance \( \sqrt{dx^2 + dy^2 + dz^2} \) in an interval of time equal to \( dt: \, dx^2 + dy^2 + dz^2 = c^2 dt^2 \). In another coordinate system, \( S' \) that same interval between emission and absorption can be described as \( dx'^2 + dy'^2 + dz'^2 = c^2 dt'^2 \). (One can then ask what linear transformation will satisfy these, and thereby deduce Lorentz transformations, but we will take a simpler route.) Placing all the terms in each of the above equations on the same side we write,

\[
-c^2 dt'^2 + dx'^2 + dy'^2 + dz'^2 = 0 = -c^2 dt^2 + dx^2 + dy^2 + dz^2
\]

where \(-c^2 dt'^2 + dx'^2 + dy'^2 + dz'^2\), or its equivalent in any other frame, is an interval between two null-separated events in space-time. Note that this keeps \( c \) constant in all frames. This interval in space-time has the same length, equal to zero for light, in both coordinate systems, and thus in all inertial coordinate frames. For particles and objects other than photons this interval is not null, but it is still invariant in all frames. It is called proper time, \( d\tau \), and is the time measured in the rest frame of the particle.

Since \( d\tau \) is the same in all frames, we should look for a transformation between coordinate systems that preserves the magnitude of the interval, \( d\tau^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \), i.e. keeps it invariant, and also preserves the form of the interval. This interval looks just like the distance between two points in the usual Cartesian system of Euclidean space, where \( dr = \sqrt{dx^2 + dy^2 + dz^2} \), except that here we have an additional dimension of time, and the \( dt^2 \) segment has a sign opposite to that of the spatial-dimension segments.

The interval \( d\tau^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \) between any two given events has a fixed length in all frames. A transformation that has the property of preserving length of segments is a rotation or translation of coordinate frames. Familiar rotation involves the spatial axes only, for example, for rotation in the \( x-y \) plane by angle \( \theta \) in the anti-clockwise direction we have:

\[
\begin{pmatrix}
ict' \\
x' \\
y' \\
z'
\end{pmatrix} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \theta & -\sin \theta & 0 \\
0 & \sin \theta & \cos \theta & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
ict \\
x \\
y \\
z
\end{pmatrix}
\]

(2)

Here \( t \) and \( z \) coordinates remain the same in both frames, while

\[
x' = x \cos \theta - y \sin \theta \quad \text{and} \quad y' = x \sin \theta + y \cos \theta
\]

(3)

If a rotation involves the time axis, say, is in the \( ict-x \) plane\(^1\) then the angle of rotation has to be not \( \theta \), but \( i\theta \). Using hyperbolic trigonometry we have

\[
\cos i\theta = \cosh \theta \quad \text{and} \quad \sin i\theta = i \sinh \theta
\]

(4)

With these, the rotation matrix in the \( ict-x \) plane becomes,

\[
\begin{pmatrix}
ict' \\
x' \\
y' \\
z'
\end{pmatrix} = \begin{pmatrix}
\cosh \theta & -i\sinh \theta & 0 & 0 \\
i\sinh \theta & \cosh \theta & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
ict \\
x \\
y \\
z
\end{pmatrix}
\]

(5)

\(^1\)Note that because the squares of \( dt \) and \( dx, dy, dz \) intervals do not all have the same sign we are dealing with non-Euclidean geometry. The space-time defined by these is called Minkowski space-time. However, for now we will pretend that we are working in Euclidean space, where one of the axes is imaginary, \( ict \).
And the new, i.e. primed time and $x$ coordinates become

$$ict' = ic \cosh \theta - ix \sinh \theta \quad \text{and} \quad x' = -ct \sinh \theta + x \cosh \theta$$

(6)

We will assume from now on that frame $S'$ is moving with respect to $S$ in the +ve $x$ direction, with $v$; and $y$ and $z$ velocities of $S'$ w.r.t. $S$ are 0. How does the rotation angle $\theta$ relate to $v$? A person sitting at the origin of $S'$ will record $x' = 0$, and from the second of eqns. 6 we get $x/t = v = \tanh \theta$. Using the identity $\cosh^2 \theta - \sinh^2 \theta = 1$, and those derived from it we get:

$$\cosh \theta = (1 - v^2)^{-1/2} = \gamma \quad \text{and} \quad \sinh \theta = \gamma v,$$

(7)

where we have introduced the Lorentz-$\gamma$ factor, which is never less than 1 and approaches infinity for $v$ close to $c$. Lorentz transformation now becomes (from eqs. 6)

$$t' = \gamma (t - xv/c^2) \quad \text{and} \quad x' = \gamma (x - tv)$$

(8)

and you can verify directly that the length of space-time interval is preserved. Equivalently, the Lorentz transformations are sometime written as follows, avoiding imaginary numbers:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

(9)

Rotations involving time axis are called boosts. The entire set of Lorentz transformations includes boosts, rotations, plus linear translations. However, it is only the rotations involving the time axis which can cause confusion because they are so different from our everyday experience.

1.2.1 Addition of velocities

Frame $S''$ moves with respect to frame $S'$ at velocity $v_1$, while frame $S'$ moves with respect to frame $S$ at $v_2$. What is the speed of $S''$ as seen from $S$? The rotation angles add up linearly (and their subscripts indicate which velocities they correspond to), so that

$$v = \tanh \theta = \tanh (\theta_1 + \theta_2) = \frac{\tanh \theta_1 + \tanh \theta_2}{1 + \tanh \theta_1 \tanh \theta_2} = \frac{v_1 + v_2}{1 + v_1 v_2}$$

(10)

this is always less than 1, i.e. less than $c$, in accordance with the 1st postulate. (Recall the shape of $\tanh \theta$ vs. $\theta$.)

1.2.2 The role of observers in SR

A major source of confusion in studying relativity arises from the choice of observers who record time and location of events. Because light has a finite speed of propagation (signals travel at finite velocities), it is very important to keep track of who is observing what, when, and according to whose clock. One can divide most of SR experiments (thought experiments or real experiments) into two classes:

(i) Those where each $S$ and $S'$ are populated with numerous observers stationed at every coordinate location, each carrying their own clock that tells the local time. Clock of all these observers in a given frame are synchronized. Coordinates $t, x, \text{etc.}$ are those measured by these observers. Lorentz transformations help you go between frames. The standard time dilation and length contraction experiments (§ 1.2.4 ) are in this category.

(ii) Those where only a single observer is involved. Here your calculations have to take into account the finite speed of propagation of signals; Lorentz transformations do not give the full answer. Examples are relativistic Doppler effect and rotation and shape change of moving boxes (§ 1.2.5).
1.2.3 Simultaneity and causality

Frame $S'$ is moving with respect to frame $S$ with velocity $v$ along the positive $x$-axis. A person standing still in $S'$ throws a ball into a basket (also stationary in $S'$) with velocity $u'$, along the positive $x'$-axis. Event #1 corresponds to the ball being released from the person’s hands; event #2 is the ball falling into the basket. By design, let the two frames’ origins coincide at the moment and location of event #1 (hence all the entries in the left hand column of the table below are zeros).

Person stationary at the origin in $S$ seen the ball’s velocity as $u = \frac{v + u'}{1 + vu'}$, same as eq. 10.

Transforming between the frames gives us (each entry is a pair of $t$ and $x$ coordinates):

<table>
<thead>
<tr>
<th>event #1</th>
<th>event #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>$(0,0)$</td>
</tr>
<tr>
<td>$S'$</td>
<td>$(0,0)$</td>
</tr>
<tr>
<td></td>
<td>$(t,x) = ut$</td>
</tr>
<tr>
<td></td>
<td>$(t' = \gamma[t - xv], \ x' = u't' = \gamma[x - tv])$</td>
</tr>
</tbody>
</table>

$t' = \gamma t(1 - vu)$, $t'$ and $\gamma$ are always positive, by definition, and $u$ cannot exceed 1 (in units of $c$). $v$ also cannot exceed 1, therefore $(1 - vu)$ is always positive, and hence $t$ is always positive. In other words, event #2 in $S$ always takes place after event #1 (at $t = 0$), or, if both $v$ and $u$ are 1, then the two events happen simultaneously. This establishes causality in SR, i.e. event #2 can never precede event #1. So causality is universally agreed upon, while simultaneity is not, since $t' \neq t$, i.e. observers in the two frames do not agree on the time when the ball hits the basket.

1.2.4 Time dilation and length contraction

These effects can be shown to result from the same transformations as we’ve just used.

The fact that moving clocks tick slower (time dilation) can be demonstrated as follows. Suppose a clock is sitting at the origin of frame $S'$. It makes the first tick at $(t', x') = (0,0)$ and the second tick at $(t', x') = (dt', 0)$ as seen by observers in its own frame. In frame $S$ the first tick is detected at $(t, x) = (0,0)$, while the second tick is detected at $(dt, dx)$, where $dx = vdt$ (second of eqns. 8).

Note that $dx$ is not zero, i.e. the second tick in frame $S$ is detected by a different observer than the one who detected the first tick. Since the spacetime interval between the two ticks is invariant, $dt'^2 = dt^2 - dx^2 = dt^2(1 - v^2)$. Therefore, $\gamma dt' = dt$, and so interval between ticks always takes longer if the clock is seen to be moving (here, by observers in $S$). Hence “moving clocks runs slow”.

The situation is completely symmetric: a clock in $S$ will be seen to run slow by those in $S'$. Substituting clocks with people and ticks by heartbeats, and remembering that person’s age is proportional to the duration of a heartbeat, we see that the traveling person (once they gets back home, i.e. after decelerating, turning around, and making the return journey) would have aged less by a factor of $\gamma$ compared to a person who stayed at home.

Length contraction means that moving objects are always shorter along their direction of motion, as seen from the lab frame (by two observers). Objects are always longest in their own rest frames. The effect is symmetric: a spaceship moving through the Galaxy looks shorter to the inhabitants of Galaxy’s planets; and the Galaxy looks shorter to the space-travelers, by the same $\gamma$ factor. As emphasised at the beginning of this section, it takes more than one observer to measure the standardly quoted time dilation and length contraction! Finally, note that time dilation and length contraction are real, and not optical illusions.

1.2.5 Relativistic Doppler

In contrast to the time dilation measurement here we have only one observer, standing still at the origin of frame $S$. The light source moves at velocity $v$ along with frame $S'$. In this case the
observer needs to wait for the light signal from the moving source to get back to him, which adds an extra \( vdt \), in the observer’s \( S \) frame, or \( vdt'\gamma \), where \( dt' \) is the interval between emitted signals in clock’s rest frame \( S' \). The interval between ticks seen by observer in \( S \) is, \( dt = \gamma(1 + v)dt' \). This reduces to the usual Doppler when \( \gamma \) approaches 1. Time dilation implies that sources moving perpendicular to the line of sight will appear redshifted to the observer. When the source is moving at an angle \( \theta \) with respect to the \(+x\) axis, the signal return time will be \( vdt\cos\theta \), and the observer at \( S \) will measure the interval between the two ticks as \( dt = \gamma(1 + v\cos\theta)dt' \). Note that the time dilation portion of \( dt \), namely \( \gamma dt' \) is unaffected.

1.3 Particle mechanics

1.3.1 Tensors

Instead of 3D vector quantities that we are familiar with from Newtonian mechanics, SR uses 4-vectors. The 4-vectors are a subclass of tensors, objects that transform in a certain way between frames, and in doing so preserve their form. Vectors are tensors of rank 1, scalars are tensors of rank 0, while tensors that are written as ‘matrices’, for example the metric tensor, are of rank 2. A good starting example of a tensor is the 4-vector \( d\tau \), with components \((dt, dx, dy, dz)\). We already know that it preserves its form under LT. Like other rank 1 tensors, it preserves its magnitude under Lorentz transformations. Other tensors can be defined based on \( d\tau = (dt, dx, dy, dz) \), and derivatives w.r.t. coordinates.

There are two types of tensors defined by how they transform between frames. Contravariant tensors transforms as

\[
dx{\nu} = \sum_{j=0}^{3} \frac{\partial x'^{\nu}}{\partial x^{j}} dx^{j}
\]

(12)

Covariant tensors transforms between frames as (where \( f \) is the function being differentiated)

\[
\frac{\partial f}{\partial x^{\nu}} = \sum_{j=0}^{3} \frac{\partial x^{j}}{\partial x'^{\nu}} \frac{\partial f}{\partial x^{j}}
\]

(13)

The two transformation are different because \( \frac{\partial x'^{\nu}}{\partial x^{j}} \) are not the same as \( \frac{\partial x^{j}}{\partial x'^{\nu}} \). One can show that by using the Lorentz transformations, and recalling that the partial derivative with respect to \( x \), keeping \( t \) constant, is

\[
\frac{\partial}{\partial x} = \frac{\partial x'}{\partial x'} \bigg|_{t} \cdot \frac{\partial}{\partial x'} + \frac{\partial t'}{\partial x'} \bigg|_{t} \cdot \frac{\partial}{\partial t'}
\]

(14)

And similarly for \( t \).

One can also show that the distinction between covariant and contravariant transformations vanishes in Euclidean geometry; they become the same. Verify this by using the usual \( x-y \) rotation matrix as an example of transformation between two Euclidean frames, and calculate the transformation coefficients for \( dx \) and \( d\tau \).

1.3.2 Displacement, velocity, momentum and acceleration four-vectors

The length squared of the displacement 4-vector, which we already encountered (eq. 1), can be written as a dot product with itself, or as a matrix multiplication

\[
d\tau^{2} = d\tau \cdot d\tau = \eta_{\alpha\beta} dx^{\alpha} dx^{\beta}
\]

(15)
where $\eta_{\alpha\beta}$ is the metric tensor and encodes the geometry of space. In Minkowski space $\eta$ has no off diagonal components (it diagonal components are -1, 1, 1, 1) so there are no $\alpha \neq \beta$ terms.

We define the components of the velocity 4-vector as follows:

$$U^\alpha = dx^\alpha/d\tau = (dt, dx, dy, dz)/(d\tau) = (1, v_x, v_y, v_z).$$

Here, $d\tau$ is the proper time, which is the time measured by the moving observer, in his own rest frame. The magnitude of the 4-velocity is

$$|U|^2 = U \cdot U = \eta_{\alpha\beta} U^\alpha U^\beta = (dt^2 + dx^2 + dy^2 + dz^2)/(d\tau)^2 = -\gamma^2 (1 - |\vec{v}|^2) = -1. \quad (17)$$

The magnitude of the velocity 4-vector, being a scalar, is invariant under coordinate transformations. The 4-velocity is a tensor because its constructed by dividing the displacement 4-vector, $d\tau = (dt, dx, dy, dz)$ (a tensor), by an invariant, $d\tau$.

Next, we define momentum 4-vector by multiplying the velocity 4-vector by multiplying it by $m_0$, $P = m_0 U = m_0 \gamma (1, v_x, v_y, v_z)$, where $m_0$ is the mass as measured in the object’s rest frame, and is called the rest mass; it is an invariant. Thus by construction $P$ is also a tensor of rank 1. Just constructing a tensor is nice, but what is more important is to see if it is useful. The time component of the 4-momentum is

$$P^0 = m_0 (1 - v^2)^{-1/2} \approx m_0 + \frac{1}{2} m_0 v^2 + ..., \text{ suggesting that "total mass" should include, in addition to the rest mass, the object’s 'classical' kinetic energy. So it appears that this terms represented the energy of the particle, } E. \text{ The other three components correspond to classical momenta, } \vec{p}. \text{ The squared magnitude of the 4-momentum is: }$$

$$|P|^2 = E^2/c^2 + |\vec{p}|^2 = -m_0^2 c^2. \quad (18)$$

and is an invariant because we constructed it from $U$, which obeys eq. 17, and $m_0$ which is an invariant. The last expression is the 4-momentum magnitude expressed in the rest frame of the particle, and also says that the maximum energy content of a particle is $m_0 c^2$. So the newly defined 4-momentum turns out to be a useful construct; it encapsulated useful physics.

Four-acceleration is defined as $dU/d\tau$. One can show that the dot product $\vec{A} \cdot U = 0$, an invariant, and represents a statement that does not have a direct counterpart in classical 3-mechanics (where acceleration can be parallel to velocity).

1.4 Fluid mechanics

1.4.1 Stress-Energy or Energy-momentum Tensor

Four momentum describes individual particles. How do we describe fluid? Consider how its density is different when fluid moves with respect to the lab frame. The mass of each particle increases by $\gamma$, and the length of the volume element (along the direction of motion) is contacted by $\gamma$, so the volume density is now $\rho \gamma^2$. The two factors of $\gamma$ suggest that the appropriate ‘momentum’ must have two factors of $U$, so we define a 4x4 entity as $T^\alpha{}_{\beta} = \rho U^\alpha U^\beta$. This is a tensor, i.e. it tranforms as a tensor between frames. In general, the $\alpha, \beta$ component is the flow of the $\alpha$ component of momentum through the $x_\beta=$constant surface. The $T^{0i}$ (where $i = 1, 3$) and the $T^{i0}$ components are then the energy flux and the momentum density, respectively, and are the same thing. Energy flux is conduction of energy in some spatial direction, i.e. ‘heat’ conductivity. The diagonal elements $T^{ii}$ are the 3 pressure components, and the off-diagonal $T^{ij}$ elements are the stresses, or viscosities.
One can show that the divergence of stress-energy tensor is zero in the absence of external forces: \( \frac{\partial}{\partial x^\beta} T^{\alpha \beta} = 0 \). For \( \alpha = 0 \), i.e. the time component, this becomes the equation of conservation of mass, or continuity equation, while for \( \alpha = 1, 2, 3 \), it gives the three (\( x, y, z \)) components of the equation of motion, or the equations for the conservation of momentum in three separate spatial directions. This shows that the stress-energy tensor we have constructed, guided by the concept of momentum of an individual particle, is a valuable quantity for describing relativistic fluids.

In cosmology we will be interested in a special type of fluid: a perfect fluid is one that has no viscosity, stresses, or conductivity. In its rest frame it is characterized by energy density, \( \rho \), and internal pressure, \( p \). Because of assumed isotropy, \( p \) is the same in all three directions. Off diagonal components are zero (no stresses), so the stress-energy tensor of a perfect fluid has only the diagonal components: \( T^{\alpha \alpha} = (\rho, p, p, p) \).

### 1.5 Maxwell’s Equations are Lorentz invariant

The homogeneous ME’s are

\[
\vec{\nabla} \cdot \vec{B} = 0 \quad \text{absence of magnetic monopoles} \tag{19}
\]

\[
\vec{\nabla} \times \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = 0 \quad \text{Faraday’s Law} \tag{20}
\]

The inhomogeneous ME’s are

\[
\vec{\nabla} \cdot \vec{E} = \rho \quad \text{Gauss’ Law for electric charges} \tag{21}
\]

\[
\vec{\nabla} \times \vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} = \frac{\vec{J}}{c} \quad \text{Ampere’s Law} \tag{22}
\]

\( \vec{E} \) and \( \vec{B} \) are related to forces that act on charged particles. By analogy with Newtonian gravity we can rewrite ME’s using potentials, recalling that one can often define force as a gradient of a potential.

Define vector and scalar E&M potentials, \( \vec{A} \) and \( \phi \) as follows. Using eq. 19, \( \vec{B} = \vec{\nabla} \times \vec{A} \), because \( \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0 \). Using eq. 20, define \( \vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \vec{\nabla} \phi \), because \( \vec{\nabla} \times (\vec{\nabla} \cdot \phi) = 0 \).

What is the relation between \( \vec{A} \) and \( \phi \)? The Lorenz condition is that the three components of \( \vec{A} \) and one component of \( \phi \) form a 4-vector, \( \Phi = (\phi, \vec{A}) \), and the divergence of \( \Phi \) is zero, establishing the connection between \( \vec{A} \) and \( \phi \). Divergence operator is Lorentz invariant, and so the following statement is true in all Lorentz frames:

\[
\vec{\nabla} \cdot \vec{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} = \frac{\partial \Phi^\alpha}{\partial x^\alpha} = 0 \tag{23}
\]

Using the two inhomogeneous equations, 21 and 22, one can show that

\[
\vec{\nabla} \cdot \vec{J} + \frac{\partial \rho}{\partial t} = \frac{\partial J^\alpha}{\partial x^\alpha} = 0, \tag{24}
\]

so the charge and current density form a 4-vector, and its divergence is Lorentz invariant.

Now we have two tensors, \( \Phi \) and \( \vec{J} \), and one can show that they are connected by

\[
(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2) \Phi = \Box \Phi = \vec{J} \tag{25}
\]

Hence all 4 Maxwell’s equations can be written as one Lorentz invariant equation, which preserves its form under coordinate transformations, even though the specific values of the components
of $\Phi$ and $J$ will change. Also note that the components of $\vec{E}$ and $\vec{B}$ are not Lorentz invariant; they are part of E&M field-strength rank 2 tensor, $F^{\alpha\beta}$ which transforms under LT.

Equation 25 is the E&M equivalent of Newtonian gravity’s Poisson equation, $\nabla^2 \Phi_{\text{Newt}} = 4\pi G \rho$; both are field equations for the corresponding theories. Our next goal is to find the field equations for relativistic gravity.

2 General Relativity (GR)

2.1 Motivation for GR

We have ‘updated’ concepts like velocity, acceleration, and force to be Lorentz invariant. The Maxwell equations of electromagnetism are already Lorentz invariant. What about gravity? A simple-minded inclusion of gravity into SR does not work: Poisson equation where the Laplacian is augmented to include $\partial^2 / \partial t^2$ predicts wrong amplitude and sense for the precession of Mercury, and does not predict bending of light. So obviously including gravity is more difficult than that. The search for the Poisson equation equivalent for GR leads to the Einstein’s field equations.

2.2 The Equivalence Principle (EP)

This is equivalence between gravitating and inertial masses. The proportionality of the two can be experimentally verified to a great precision, but it took a leap of faith for Einstein to assume that the two are not just proportional, but are exactly the same. Making the assumption that the two types of masses are the same, and so equating $m \ddot{a}$ and $GmM/r^2$, we see that acceleration due to gravity will be the same for objects of all masses, $\ddot{a} = GM/r^2$. The observational result of this—that a feather and a bowling ball when released simultaneously from a tower through a vacuum tube will land on the ground at the same time—was known since Galileo. Einstein fully exploited the far reaching consequences of this observation. If two particles of very different masses (both much smaller than the mass of Earth in this particular example) follow same paths, they must be following ‘tracks’ determined by curved geometry of space-time. The fact that space-time in the presence of gravity must be curved is also apparent if you watch the path of a ball which has been thrown parallel to the surface of the Earth. Once released, the ball is a free particle, not acted upon by forces, and yet it traces a curved path\(^2\). Replacing the concept of gravity with the concept of curved geometry allows us to maintain that the particle is following a straight path. This substitution of geometry in place of gravity also gets rid of the spooky action-at-a-distance concept: for example, the falling motion of a ball is merely a response to the local shape of space-time, not Earth’s gravitational attraction.

It is important to remember that the EP only provides a path to GR. It does not give the full description of the relativistic space-time because it ignores the key feature of gravity, tidal forces, which make the distances between neighboring falling particles change.

Equivalence principle predicts two effects: gravitational bending of light and gravitational time dilation.

In its approximate form time dilation can be calculated entirely from the Newtonian point of view, because Newton and Einstein must agree in the limit of weak gravitational fields and nonrelativistic velocities. First, we replace the Earth with an elevator, height $h$, which is being accelerated upwards at $g$. Then we place observer $A$ next to a source of EM radiation, which is

\(^2\)In GR it is the space-time that is curved, not just the 3D space. When we watch a thrown ball we are seeing just the space portion of the ball’s trajectory; its space-time trajectory would appear a lot less curved because the ball covers more time than space.
affixed to the top of the elevator, and radiates at frequency \( \nu_0 \). Observer \( B \) is at the bottom of the elevator. In that time the velocity of \( B \) changes by \( v = gt \), so \( B \) sees radiation blueshifted, with \( \nu \rightarrow \nu + \Delta \nu \). Since \( v \) is non-relativistic, we can use the regular Doppler formula, with

\[
\Delta \nu / \nu = v / c = h g / c^2 = (\Phi_A - \Phi_B) / c^2 = -\Delta \Phi / c^2, \tag{26}
\]

which is solved by

\[
\nu / \nu_0 = e^{-\Delta \Phi / c^2} \approx (1 - \Delta \Phi / c^2) \tag{27}
\]

So the intervals between two successive wave crests (or clock ticks) as viewed by \( A \) and \( B \) are related by \( \Delta t_B = \Delta t_A (1 + \Delta \Phi / c^2) \), where \( \Delta t_B \) is the interval that \( B \) ascribed to \( A \)’s clock. Since \( \Delta \Phi / c^2 < 0 \), \( B \) concludes that his clock is running slower than \( A \)’s. Now, replacing back the elevator with Earth, we note that \( A \) is located at a higher gravitational potential (higher above the surface of Earth), and \( B \) is located deeper within the potential well. This result can be extended to a case where observer \( A \) is at zero potential (i.e. very far from any source of gravity), and measures clock tick interval as \( \Delta t_0 \), and \( B \) is at some non-zero potential, and measures \( \Delta t = \Delta t_0 (1 + \Phi / c^2) \). \( A \) and \( B \) are not moving with respect to one another so this time dilation is not Doppler. There are two possible interpretations (1) clocks have different rates, but space-time is Minkowskian, (2) clock rates are the same, but the time component in the metric of the geometry has value different from that in \( \eta_{\alpha \beta} \), i.e. space-time is not Minkowskian. Einstein adopted (2), and the factor \( (1 + \Delta \Phi / c^2) \) is related to the time coefficient of the metric tensor.

The amount of deflection of light calculated based on the EP is not the full correct value, as determined by experiment, for example, the bending of light from distant stars that just grazes the edge of the Sun’s surface as seen from Earth. This is because the EP does not give the full GR.

### 2.3 The Weak Field limit of GR

In the Minkowski case the metric tensor \( \eta_{\mu \nu} \) is diagonal with the elements \((-1, 1, 1, 1)\). In a general geometry, the elements take on other values, and are represented by \( g_{\mu \nu} \). In other words, the interval squaread between two neighboring events in space-time is given by

\[
d\tau^2 = g_{00} dt^2 + g_{11} dx^2 + g_{22} dy^2 + g_{33} dz^2. \tag{28}
\]

Using the results of the previous section we can get the time-time component of the metric: since \( \Delta t \approx \Delta t_0 (1 + \Phi) \), and \( \Phi \) is small, \( \Delta t^2 \approx \Delta t_0^2 (1 + 2\Phi) \), hence \( g_{00} \approx -(1 + 2\Phi) \). (The minus sign appears because when \( \Phi = 0 \), \( g_{00} \) must become -1. I’ve also dropped factors of \( c^2 \) that divide the potential.) We see that the GR analogs of the Newtonian potential \( \Phi \) are the \( g_{\mu \nu} \)’s.

The spatial components of the metric cannot be obtained from a simple thought experiment. They can be deduced from the full theory, in the weak field regime. They are \( g_{ii} = (1 - 2\Phi) \). The full weak field metric is

\[
d\tau^2 = -(1 + 2\Phi) dt^2 + (1 - 2\Phi) dr^2 \tag{29}
\]

where \( d\tau \) is the invariant interval. This metric is used in cosmological applications of propagation of gravitational waves, and gravitational lensing. For example, imagine an intervening galaxy causing light from a background quasar to take more than one path to us because the corresponding action (integral of the Lagrangian along the path, \( S = \int d\tau \)) is extremized along more than one path from quasar to us. What we see as a result are multiple images of the same background quasar. Light travels along null geodesics, so \( d\tau^2 = 0 \) for light along each path/image. So from eq. 29 we get \( dt = (1 - 2\Phi) dr \), where we have used approximations of the sort \((1 - 2\Phi)^{-1} \approx (1 + 2\Phi)\) and
(1 − 2Φ)^{1/2} ≈ (1 − Φ) because Φ is small. Once we find all the paths from the source to us that the light can take, and hence locate all the images, we can calculate how long (in our frame) it took light to travel along these paths. Integrating dt along these paths gives different results for different paths/images, because both the geometrical path length (∫ dr) and the time delay due to gravitational potential (∫ 2Φ dr) are different for the two paths. These time delays have been measured for a number of distant quasars, providing yet another observational verification of GR’s predictions.

2.4 Motion of a single particle

The EP suggests that gravity can be expressed as geometry with non-Minkowskian metric, so space will have some curvature. We need to figure out how to describe that curvature.

For a given particle moving through space-time we can always find an instantaneous frame in which the particle is not being accelerated, i.e.

\[
\frac{dU^{\alpha}}{d\tau} = \frac{d^2x^{\alpha}}{d\tau^2} = 0
\]  

(30)

This is the freely falling frame, where one would experience weightlessness. (In terms of geometry going into a freely-falling frame corresponds to finding a ‘tangent space’ to a curved space at a given location, similar to a tangent plane to a sphere at some point; in a small enough patch around that point a plane approximates a sphere pretty well.) This instantaneous frame is described by the metric η_{αβ}. Suppose we now look at the same particle from another frame, in which the particle is seen to be accelerating. We can change coordinate systems, and rewrite eq. 30 in the accelerated (unprimed) coordinate frame. The result is the geodesic equation of motion,

\[
\frac{d^2x^{\alpha}}{d\tau^2} = \Gamma^{\alpha}_{\beta\gamma} \frac{dx^{\beta}}{d\tau} \frac{dx^{\gamma}}{d\tau}
\]  

(31)

where Γ’s are called Christoffel symbols (a.k.a. the affine connection). These are analogous to forces (first gradients of the potential) in Newtonian language. Since in GR gravity is a fictitious force, one should be able to have Γ’s disappear by an appropriate selection of a coordinate frame, just what we had in eq. 30. Γ^{α}_{βγ}’s are 0 in some frames and non-0 in others, hence they are not coordinate invariant, and not tensors.

The curvature of a given geometry is an intrinsic property of that geometry, and as such, its characterization should not depend on the coordinate system that we paint on that geometry. Γ^{α}_{βγ}’s are coordinate dependent to the degree that they can be made to vanish in some coordinate systems. To sum up, quantities arising from considering motion of a single particle will not be able to characterize the geometry, we need to consider motion of a collection of particles.

2.5 Motion of a collection of particles: characterizing curvature

The basic idea can be summarized by a simple example: two particles freely falling in a gravitational field, like falling towards Earth will (in addition to approaching Earth) also come closer to each other. This is the way to tell that gravity is present; just acceleration towards Earth is not enough. This coming together of particles is described by tidal forces, which are 2nd derivatives of the potential,

\[
\frac{d^2\Delta x^i}{dt^2} = -\Sigma_{j} \frac{\partial^2 \Phi}{\partial x^i \partial x^j} d\Delta x^j
\]  

(32)
where $\Delta x_i$ is the separation between two neighboring test particles. In GR the equivalent of Newtonian tidal forces is

$$\frac{d^2 \Delta x^\alpha}{d\tau^2} = -R^\alpha_{\beta\gamma\delta} \Delta x^\gamma \tag{33}$$

(In eq. 32 $j$ indices are being summed over; ignore the fact that the $i$’s don’t follow the Einstein convention. In eq. 33 $\gamma$’s are being summed over; $\beta$’s are the time components.) Since tidal forces can be detected from any reference frame, and cannot be made to vanish they are the true measure of gravity. $R^\alpha_{\beta\gamma\delta}$ is the Riemann curvature tensor, and is a function of the metric tensor and its first and second derivatives only.

Here’s another way of explaining why we need $R^\alpha_{\beta\gamma\delta}$. The geometry is specified by the metric tensor. However, $g_{\mu\nu}$’s are coordinate dependent whereas the curvature of a given space is an intrinsic property of that space and is not coordinate dependent. In fact, many different sets of $g_{\mu\nu}$’s can describe the same space. So we need to come up with a tensor that will describe the intrinsic curvature of space regardless of how we draw the coordinate axes in that space. It has to have the second derivatives of $g_{\mu\nu}$’s since the first derivatives can be gotten rid of—in the Newtonian analog these correspond to the spatial gradients of $\Phi$, i.e. the gravitational force. Second derivatives of the potential, on the other hand, are related to the matter density in the Newtonian analogue, $\Sigma_i \frac{\partial^2 \Phi_i}{\partial x^2_i} = \nabla^2 \Phi \propto \rho$, and do not go away as long as there are sources of gravity, i.e. $\rho \neq 0$.

2.6 Field equations

Our goal is to incorporate gravity into relativity; in particular we are looking for equation(s) that would replace Poisson equation,

$$\nabla^2 \Phi = 4\pi G \rho, \tag{34}$$

that connects the potential which characterizes gravity (LHS) with density, the source term that generates gravity (RHS). Relating these is the big problem that took Einstein a long time to solve. In GR good candidates for the LHS would contain some function of the Riemann curvature tensor, while the RHS should contain the source term, i.e. some function of the stress-energy tensor.

Einstein arrived at the right equations (as far as we know by testing these against experiments) by simple arguments plus some guesswork. Guesswork was involved because it is a new piece of physics that was being sought; nothing that can be derived from existing physical laws would be new physics. Here are some helpful hints, i.e. conditions that have to be satisfied by the field equations:

(1) Only tensors should go into the equation because the whole equation should remain form invariant under generalized coordinate transformations;

(2) Einstein limited his search to tensors that contain only up to second order derivatives of the metric tensor. Riemann curvature tensor is an excellent candidate;

(3) Because stress-energy tensor which will go in to the RHS of the field equation is a rank 2 tensor, the LHS should also contain only 2nd rank tensor(s). Because of the various symmetries of $R_{\mu\nu\rho\sigma}$ (only 20 of its $4^4 = 256$ components are independent) only two unique rank 2 tensors can be constructed out of it, the Ricci tensor $R_{\mu\nu}$ and the Ricci scalar $R$, which when multiplied by the metric tensor gives rank 2 tensor, so the LHS will be some linear combination of $R_{\mu\nu}$ and $R g_{\mu\nu}$;

(4) We saw earlier that in the absence of external forces or sources the divergence of the stress-energy, which is on the RHS of the field equation, is 0. Therefore the divergence of the LHS must also be 0. This condition fixes the constants multiplying the two terms on the LHS;

(5) When all the components of $T_{\mu\nu}$ are zeros then the geometry of space-time must be flat, i.e. zero space-time curvature, and zero Ricci tensor.
All these conditions lead to a unique solution:

\[ R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu} \] (35)

The constant $8\pi G$ comes from comparison with the result in the Newtonian limit, and $G$ is determined from experiments.

If condition (5) above is not applied then we are free to add a constant to the equation, $-\Lambda g_{\mu\nu}$, and (1)-(4) conditions will still be satisfied. Comparing the form of $g_{\mu\nu}$ to the form of the stress energy tensor of a perfect fluid in its rest frame, we conclude that this extra term must satisfy $\rho = -p$, i.e. must have negative pressure, because there is no such thing as negative mass. These days the $\Lambda$ term is written on the RHS of eq. 35, and is considered to be a contribution to the stress-energy. What is its physical origin? It is generally thought to be due to the vacuum energy density, though other interpretations exist as well.

### 2.7 Non-linearity of GR

A major difference between GR and Newtonian gravity is that the latter is a linear theory, while the former is not: if there are two sources of gravity, like two stars, than in Newtonian gravity one would linearly superimpose the two gravitational potentials and there is no additional ‘interaction’ term. In GR there is: the energy in the gravitational field acts like additional source of gravitation (and changes with the separation of the two stars). Einstein’s field equations are the answer to the complex problem of how to include this effect. Because of their non-linearity, ‘solving’ the field equations proceeds by a different route than usual equation-solving: one has to guess at the right form of $T_{\mu\nu}$ and the corresponding $R_{\mu\nu\lambda\kappa}$ and see if they satisfy the field equations. Exact solutions are rare: Schwarzschild, Kerr, Friedmann, and Weak Field equations. All these are highly symmetric. Friedmann equation governs the evolution of a perfectly smooth Universe.