Contents

Part I
Relativity

1 Special Relativity (SR)

1.1 Introduction

Early experiments and observations:
• *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.
• *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).
• *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.
• *Maxwell's equations are not invariant under Galilean transformations, but are invariant under Lorentz transformtions.* It was believed that 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, 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. 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) **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) **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, force, etc. such that they are Lorentz-invariant, and reduce to their pre-SR forms in the low velocity limit.

1.2 Lorentz transformations

The first thing to note about relativity is that 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, and when, and according to whose clock. Because of this, every coordinate system we consider has its own set of observers sprinkled throughout it, each observer knows his/her location ($\vec{x}$) in that coordinate system, and each carries his/her own clock (t). In any given coordinate system all the clocks are synchronized. 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 at a given place at a given time, and event 2 is the absorption of that light signal at a different place at a later time. 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: $dx^2 + dy^2 + dz^2 = c^2dt^2$. In another coordinate system, S' that same interval between emission and absorption can be described as $dx'^2 + dy'^2 + dz'^2 = c^2dt'^2$. Placing all the terms in each of the above equations on the same side we can write,

$$-c^2dt^2 + dx^2 + dy^2 + dz^2 = 0 = -c^2dt'^2 + dx'^2 + dy'^2 + dz'^2, \tag{1}$$

Where $c^2dt^2 - dx^2 - dy^2 - dz^2$, or its equivalent in any other frame, is an interval between two null-separated events in space-time. 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. This interval is 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, $d\tau^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$, i.e. keeps it *invariant*, and also preserves the *form* of the interval.[1] This interval looks just like the distance between two points in the usual Cartesian system, 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. This means that the projections of the segment onto all 4 coordinate axes are $\sqrt{dx^2} = dx$, etc. (as usual), and $\sqrt{-c^2dt^2} = icdt$. So the axes of our space-time are: x, y, z, and ict, and all are mutually orthogonal.

The interval $d\tau^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$ 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 θ in the anti-clockwise direction) we have:

$$\begin{pmatrix} ict' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ict \\ x \\ y \\ z \end{pmatrix} \tag{2}$$

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

$$x' = x\cos\theta - y\sin\theta \quad \text{and} \quad y' = x\sin\theta + y\cos\theta \tag{3}$$

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

$$\cos i\theta = \cosh\theta \quad \text{and} \quad \sin i\theta = i\sinh\theta \tag{4}$$

[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.

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

$$\begin{pmatrix} ict' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh\theta & -i\sinh\theta & 0 & 0 \\ i\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ict \\ x \\ y \\ z \end{pmatrix} \tag{5}$$

And the new, i.e. primed time and x coordinates become

$$ict' = ict\cosh\theta - ix\sinh\theta \quad \text{and} \quad x' = -ct\sinh\theta + x\cosh\theta \tag{6}$$

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 θ 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:

$$\cosh\theta = (1 - v^2)^{-1/2} = \gamma \quad \text{and} \quad \sinh\theta = \gamma v, \tag{7}$$

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

$$t' = \gamma(t - xv/c^2) \quad \text{and} \quad x' = \gamma(x - tv) \tag{8}$$

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:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{9}$$

The entire set of Lorentz transformations includes rotations in all the possible planes of space-time, plus linear translations. However, it is only the rotations involving the ict 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

$$v = \tanh\theta = \tanh(\theta_1 + \theta_2) = \frac{\tanh\theta_1 + \tanh\theta_2}{(1 + \tanh\theta_1 \tanh\theta_2)} = \frac{v_1 + v_2}{1 + v_1 v_2} \tag{10}$$

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

1.2.2 Simultaneity/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 $u = (v + u')/(1 + vu')$. Transforming between the frames gives us (each entry is a pair of t and x coordinates):

$$\begin{array}{lll} & \textit{event \#1} & \textit{event \#2} \\ S & (0,0) & (t, x = ut) \\ S' & (0,0) & (t' = \gamma[t - xv], \quad x' = u't' = \gamma[x - tv]) \end{array} \tag{11}$$

$t' = \gamma t(1 - vu)$. t' and γ 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 the S frame 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.

1.2.3 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). 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". 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 he gets back home, i.e. after decelerating, turning around, and making the return journey) would have aged less by a factor of γ 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. 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 γ factor. As emphasised at the beginning of this section, it takes more than one observer to measure time dilation and length contraction! Finally, note that time dilation and length contraction are real, and not optical illusions.

1.2.4 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 γ 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 θ 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 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. (These two types of tensors are called contravariant and covariant.)

1.3.2 Four-velocity and four-momentum

We define the components of the velocity 4-vector of frame S' as seen from S as follows:

$$U = d\tau/d\tau' = \frac{(dt, dx, dy, dz)}{(dt', 0, 0, 0)} = \frac{(dt, dx, dy, dz)}{dt/\gamma} = \gamma(1, v_x, v_y, v_z). \tag{12}$$

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

$$|U|^2 = |\frac{d\tau}{d\tau'}|^2 = \frac{(dt^2 - dx^2 - dy^2 - dz^2)}{(dt/\gamma)^2} = \gamma^2(1 - |\vec{v}|^2) = 1. \tag{13}$$

Since we defined the 4-velocity as differentiation $d\tau/d\tau'$, where $d\tau$ and $d\tau'$ are the invariant proper time, it is not a surprise that the magnitude of the velocity 4-vector is invariant under coordinate transformations. The 4-velocity is a tensor because $d\tau = (dt, dx, dy, dz)$ is (transforms as) a tensor.

Next, we define momentum 4-vector, $P = m_0 U = m_0\gamma(1, v_x, v_y, v_z)$, where m_0 is the mass as measured in the object's rest frame, and is called the rest mass. Observer in some other frame, say S will measure the mass to be γm_0. (Massless particles have $|P| = 0$.) The time component of the 4-momentum is $P^0 = m_0(1 - v^2)^{-1/2} \approx m_0 + \frac{1}{2}m_0 v^2 + ...$, suggesting that "total mass" should include, in addition to the rest mass, the object's 'classical' kinetic energy. This also suggests that the maximum energy contained in (or, can be extracted from) an object is $m_0 c^2$. Because $P = m_0 U = m_0\gamma(1, v_x, v_y, v_z)$, the squared magnitude of the 4-momentum is $|P|^2 = -m_0^2$, or, writing it in terms of its components in an arbitrary interial frame, and putting back the factors of c where needed, -$m_0^2 c^2 = -E^2/c^2 + |\vec{p}|^2$, where $\vec{p}$ is the 3-momentum.

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 γ, and the length of the volume element (along the direction of motion) is contacted by γ, so the volume density is now $\rho\gamma^2$. The two factors of γ 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 α, β component is the flow of the α component of momentum through the x_β=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.

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 is it characterized by energy density, ρ 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: $T^{\alpha\alpha} = (\rho, p, p, p)$.

2 General Relativity (GR)

2.1 Motivation for GR

We have 'updated' concepts like velocity, acceleration, and force to also 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 with Laplacian operator to include time derivative 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.

[2]Note that 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.

It is important to remember that the EP only provides a path to GR. Tt 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.*

Time dilation can be calculated approximately, 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. Next we place observer A next to a source of EM radiation, which is affixed to the top of the elevator, and radiates at frequency ν_0. Observer B is at the bottom of the elevator. It takes light from EM source $t = h/c$ to get to B. In that time the velocity of B changes by $v = gt$, so B sees radiation blueshifted, with $\nu \to \nu + \Delta\nu$. Since v is non-relativistic, we can use the regular Doppler formula, with

$$\Delta\nu/\nu = v/c = hg/c^2 = (\Phi_A - \Phi_B)/c^2 = -\Delta\Phi/c^2, \tag{14}$$

which is solved by

$$\nu/\nu_0 = e^{-\Delta\Phi/c^2} \approx (1 - \Delta\Phi/c^2) \tag{15}$$

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 Δt_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. The result we just derived 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 Δt_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; in other words, the time dilation effect we have derived is not Doppler. And yet they do not agree on the rates of their clocks: A's as seen by B, or B's as seen by A. The interpretation that we will adopt of this is that A and B must be living in a non-flat space-time geometry, and so the factor of $(1 + \Delta\Phi/c^2)$ 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 is not yet the full GR.

One could argue that the main usefulness of the EP is conceptual. In SR's inertial frames a free object traces out a straight path in space-time. Deflection of light, throwing of a ball, and other such experiments suggest that if free objects in the presence of gravity (i.e. in GR) are to follow straight lines, then the underlying geometry must be curved, and so the underlying geometry of a general space-time in non-Minkowskian.

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 between two neighboring points in space-time (events) is given by

$$d\tau^2 = g_{00}dt^2 + g_{11}dx^2 + g_{22}dy^2 + g_{33}dz^2. \tag{16}$$

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 Φ 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 pontential.) We see that the GR analogs of the Newtonian potential Φ are the $g_{\mu\nu}$'s. The spatial components of the metric are $g_{ii} = (1 - 2\Phi)$, again, in the weak field approximation. Notice that the $g_{\mu\nu}$'s are closely related to Newtonian potential. The full weak field metric is

$$d\tau^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)dr^2 \tag{17}$$

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) is extremized by 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. 17 we get $dt = (1 - 2\Phi)dr$, where we have used approximations of the sort $(1 - 2\Phi)^{-1} \approx (1 + 2\Phi)$ and $(1 - 2\Phi)^{1/2} \approx (1 - \Phi)$ because Φ is small. Integrating dt along two lines of sight gives different results for the time it takes light to get to us, because both the geometrical path length ($\int dr$) and the time delay due to gravitational potential ($-2 \int \Phi dr$) 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 particles

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.

$$\frac{dU'^{\alpha}}{d\tau} = \frac{d^2 x'^{\alpha}}{d\tau^2} = 0 \tag{18}$$

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 $\eta_{\alpha\beta}$. 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. 18 in the accelerated (unprimed) coordinate frame. The result is the geodesic equation of motion,

$$\frac{d^2 x^{\alpha}}{d\tau^2} = \Gamma^{\alpha}_{\beta\gamma} \frac{dx^{\beta}}{d\tau} \frac{dx^{\gamma}}{d\tau} \tag{19}$$

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 fictions force, one should be able to have Γ's disappear by an appropriate selection of a coordinate frame, just what we had in eq. 18. $\Gamma^{\alpha}_{\beta\gamma}$'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. $\Gamma^{\alpha}_{\beta\gamma}$'s are coodinate dependent to the degree that they can be made to vanish in some coordinate systems, so these quantities, at least by themselves, will not be able to characterize the geometry. We need something else...

As an aside, it is interesting to note that eq. 19 can be arrived at via a more elegant route. We note that proper time $\int (d\tau^2)^{1/2}$ is maximized for particles that remain inertial (recall the twin paradox). Generalizing, we speculate that proper time is *extremized* for these particles, and that the statement is also true in GR. This suggests that proper time might be the Lagrangian of a particle (equivalent to $L = K.E. - P.E.$ in classical dynamics). Applying Euler-Lagrange equation using proper time as the Lagrangian results in eq. 19, the GR geodesic equation of motion. It is reassuring to recall that, in classical dynamics, applying Euler-Lagrange equation to Lagrangian of the form $L = K.E. - P.E.$ gives Newton's 2nd law of motion.

2.5 Deviation equations: characterizing curvature

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. This tensor is the Riemann curvature tensor, $R_{\mu\nu\sigma\delta}$, and is a function of the metric tensor and its first and second derivatives only. 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 Φ, i.e. the gravitational force. Second derivatives of the potential, on the other hand, are related to the matter density in the Newtonian analoque, $\Sigma_i \partial^2\Phi_i/\partial x_i^2 = \nabla^2\Phi \propto \rho$, and do not go away as long as there are sources of gravity, i.e. $\rho \neq 0$. In fact, the GR equivalent of Newtonian tidal forces

$$\frac{d^2\Delta x_i}{dt^2} = -\Sigma_j \frac{\partial^2\Phi}{\partial x_i \partial x_j} d\Delta x_j \tag{20}$$

(where Δx_i is the separation between two neighboring test particles) is

$$\frac{d^2\Delta x^\alpha}{d\tau^2} = -R^\alpha_{\beta\gamma\beta}\Delta x^\gamma \tag{21}$$

(In eq. 20 j indices are being summed over, in eq. 21 γ's are being summed over.) Since tidal forces can be detected from any reference frame, and cannot be made to vanish they are *the true measure of gravity*.

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, $\nabla^2\Phi = 4\pi G\rho$. The latter connects the potential which characterizes gravity (LHS) and density, the source term that generates gravity (RHS). Relating these is the big problem that took Einstein a long time to solve. In GR the LHS would contain some function of the Riemann curvature tensor, while the RHS should contain the source term, i.e. 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, and only tensors obey that. (2) Einstein limited his search to tensors that contain only up to second order derivate of the metric tensor. Riemann curvature tensor is an excellent candidate. (3) Since stress-energy is a second rank tensor, i.e. has two indices, the LHS should contain only second rank tensor(s) as well. Because of the various symmetries of $R_{\mu\nu\sigma\delta}$ (only 20 of its $4^4 = 256$ components are independent) only two unique rank-2 tensors can be constructed out of it, Ricci tensor and Ricci scalar, $\mathtt{R}_{\mu\nu}$ and $\mathtt{R}$, so the LHS will be some linear combination of $\mathtt{R}_{\mu\nu}$ and $\mathtt{R}g_{mu\nu}$. (4) The divergence of the LHS must be equal to 0, because as we saw earlier, in the absence of external forces or sources the divergence of the stress-energy is 0. This 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:

$$\mathtt{R}_{\mu\nu} - \frac{1}{2}\mathtt{R}g_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{22}$$

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 Λ term is written on the RHS of eq. 22, 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.