Indian Institute of Technology
Bombay

Special Theory of Relativity & Lagrangian Formalism

Author: Neha Pedgaonkar

Roll no.: 22B1805

Department: Engineering Physics

Contents

Chapter 1

Introduction: What is Relativity?

What the principle of relativity essentially states is the following:

The laws of physics take the same form in all frames of reference, moving with constant velocity concerning one another.

First consider an example from 'everyday experience' – a train carriage moving smoothly at a constant speed on a straight and level track – this is a 'frame of reference'. Suppose that in this carriage is a pool table. If you were a passenger on this carriage and you decided to play a game of pool, one of the first things that you would notice is that in playing any shot, you would have to make no allowance whatsoever for the motion of the train. Any judgement of how to play a shot, as learned by playing the game back home or in the local pool hall, would apply equally well on the train, irrespective of how fast the train was moving. If we consider that what is taking place here is the innate application of Newton's Laws to describe the motion and collision of the pool balls, we see that no adjustment has to be made to these laws when playing the game on the moving train.

This argument can be turned around. Suppose the train windows are covered, and the carriage is well insulated so that there is no immediate evidence to the senses as to whether or not the train is in motion. It might nevertheless still be possible to determine if the train is in motion by carrying out an experiment, such as playing a game of pool. But, as described above, a game of pool proceeds in exactly the same way as if it were being played back home – no change in shot-making is required. There is no indication from this experiment as to whether or not the train is in motion. There is no way of knowing whether, on pulling back the curtains, you are likely to see the countryside hurtling by or to find the train sitting at a station. In other words, what the principle of relativity means is that it is not possible to determine whether or not the train

carriage is moving.

This idea can be extended to encompass other laws of physics. To this end, imagine a collection of spaceships with engines shut off, all drifting through space. Each spaceship constitutes a 'frame of reference', an idea that will be better defined later. On each of these ships, a series of experiments is performed: a measurement of the half-life of uranium, a measurement of the outcome of the collision of two billiard balls, an experiment in thermodynamics, e.g. a measurement of the specific heat of a substance, a measurement of the speed of light radiated from a nearby star: any conceivable experiment. If the results of these experiments are later compared, what is found is that in all cases (within experimental error), the results are identical. In other words, the various laws of physics being tested here yield exactly the same results for all the spaceships, in accordance with the principle of relativity.

Thus, quite generally, the principle of relativity means that it is not possible, by considering any physical process whatsoever, to determine whether or not one or the other of the spaceships is 'in motion'. The results of all the experiments are the same on all the spaceships, so there is nothing that definitely singles out one spaceship over any other as being the one that is stationary. It is true that from the point of view of an observer on any one of the spaceships, it is the others that are in motion. But the same statement can be made by an observer in any spaceship. All that we can say for certain is that the spaceships are in relative motion, and we cannot claim that one of them is 'truly' stationary while the others are all 'truly' moving

Chapter 2

Lorentz Transformation

In deriving this transformation, we will eventually use the constancy of the speed of light. Still, first, we will derive the general form that the transformation law must take purely from kinematic/symmetry considerations. The starting point is to consider two inertial frames, S and S', where S' is moving with a velocity v_x relative to S.

Let us suppose that when the two origins coincide, the times on the clocks in each frame of reference are set to read zero, that is, t = $t' = 0$. Now consider an event that occurs at the point (x, y, z, t) as measured in S. The same event occurs at (x', y', z', t') in S'. What we are after is a set of equations that relate these two sets of coordinates.

We are going to assume a number of things about the form of these equations, all of which can be fully justified but which we will introduce more or less on the basis that they seem intuitively reasonable.

First, because the relative motion of the two reference frames is in the X direction, it is reasonable to expect that all distances measured at right angles to the X direction will be the same in both S and S', i.e.[5]

$$y = y' \text{ and } z = z'$$

We now assume that (x, t) and (x', t') are related by the linear transformations

$$x' = Ax + Bt$$

$$t' = Cx + Dt$$

Why linear? Assuming that space and time are homogeneous tells us that a linear relation is the only *possibility*[6]. What it amounts to saying is that it should not matter where in space we choose our origin of the spatial coordinates to be, nor should it matter when we choose the origin of time, i.e. the time that we choose to set as t = 0.

Now consider the origin $O^{'}$ of $S^{'}$. This point is at $x^{'} = 0$, which, if substituted into Eq. (22), gives

$$Ax + Bt = 0$$

where x and t are the coordinates of $O^{'}$ as measured in S, i.e. at time t, the origin $O^{'}$ has the X coordinate x, where x and t are related by $Ax + Bt = 0$. This can be written

$$\frac{x}{t} = -\frac{B}{A}$$

but x/t is just the velocity of the origin $O^{'}$ as measured in S. This origin will be moving at the same speed as the whole reference frame, so then we have

$$-\frac{B}{A} = v_x$$

which gives $B = v_x A$, which can be substituted into Eq. (22) to give

$$x^{'} = A(x \ v_x t)$$

---

[5]If we assumed, for instance, that $z = kz^{'}$, then it would also have to be true that $z^{'} = kz$ if we reverse the roles of S and $S^{'}$, which tells us that $k^2 = 1$ and hence that $k = \pm1$. We cannot have $z = z^{'}$ as the coordinate axes are clearly not 'inverted', so we must have $z = z^{'}$.

[6]In general, $x^{'}$ will be a function of x and t, i.e. $x^{'} = f(x, t)$ so that we would have $dx^{'} = fxdx + ftdt$ where fx is the partial derivative of f with respect to x, and similarly for ft. Homogeneity, then means that these partial derivatives are constants. In other words, a small change in x and t produces the same change in $x^{'}$ no matter where in space or time the change takes place.

If we now solve Eq. (22) and Eq. (23) for x and t, we get

$$x = \frac{Dx' + v_x At'}{AD - BC}$$

$$t = \frac{At' - Cx'}{AD - BC}$$

If we now consider the origin O of the reference frame S, that is, the point x = 0, and apply the same argument as just used above, and note that O will be moving with a velocity v_x with respect to S', we get

$$-\frac{v_x A}{D} = -v_x$$

Comparing this with the above equation, we see that: A = D
and hence the above transformations from S' to S will be, after substituting for D and B:

$$x = \frac{A(x' + v_x t')}{A^2 + v_x AC}$$

$$t = \frac{A(t' - (C/A)x')}{A^2 + v_x AC}$$

which we can compare with the original transformation from S to S'

$$x' = A(x - v_x t)$$

$$t' = A(t + (C/A)x).$$

Any difference between the two transformation laws can only be due to the fact that the velocity of S' with respect to S is v_x and the velocity of S with respect to S' is v_x. So, given the transformation laws that give the S coordinates in terms of the S' coordinates, the corresponding equations going the other way can be obtained simply by swapping the primed and unprimed variables and changing the sign of v_x. If that is to be the case, then the factor $A^2 + v_x AC$ must be unity, i.e.

$$A^2 + v_x AC = 1$$

and it also suggests that C/A is proportional to v_x to guarantee the change in sign that occurs in passing from the expression for t to the one for $t^{'}$. Thus, we have

$$A^2(1 + v_x(\text{C/A})) = 1$$

from which we get

$$A = \frac{1}{\sqrt{1 + v_x C/A}}$$

If we now use the clue that C/A is proportional to v_x to try a substitution $\text{C/A} = v_x/V^2$ where V is a quantity with the units of velocity yet to be determined, we have

$$A = \frac{1}{\sqrt{1 - (v_x/V)^2}}$$

so that finally the transformation laws become

$$x^{'} = (\text{x-}v_x\text{t})/\sqrt{1 - (v_x/V)^2}$$

$$y^{'} = \text{y}$$

$$z^{'} = \text{z}$$

$$t^{'} = (\text{t-}v_x/\text{V}^2)x/\sqrt{1 - (v_x/V)^2}$$

This is a remarkable and very general result that depends purely on the assumed homogeneity and isotropy of space. At no stage have we mentioned light or any other physical quantity for that matter, and yet we have been able to pin down the transformation laws relating coordinate systems for two different inertial frames of reference, at least as far as there is only one undetermined quantity left, namely V. More information is needed to determine its value, but if we were to choose $V = \infty$, then we find that these transformation equations reduce to the Galilean transformation! However, we have yet to make use of Einstein's second proposal. In doing so, we are able to determine V and find that V has an experimentally determinable, finite value.

To this end, let us suppose that when the two origins coincide, the clocks at O and $O^{'}$ both read zero and also suppose that, at that instant, a flash of light is emitted from the coincident points O and $O^{'}$. In the frame of reference S,

this flash of light will be measured as lying on a spherical shell centred on O whose radius is growing at the speed c. However, by the second postulate, in the frame of reference S', the flash of light will also be measured as lying on a spherical shell centred on O' whose radius is also growing at the speed c. Thus, in S, if the spherical shell passes a point P with spatial coordinates (x, y, z) at time t, then by our definition of synchronization, we must have:

$$x^2 + y^2 + z^2 - c^2t^2 = 0$$

The flash of light passing the point P in space at time t defines an event with spacetime coordinates (x, y, z, t). This event will have a different set of coordinates (x', y', z', t') relative to the frame of reference S', but by our definition of synchronization, these coordinates must also satisfy the following:

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0$$

We want to find how the two sets of coordinates (x, y, z, t) and (x', y', z', t') are related in order for both the above equations to hold true. But we generally know that these coordinates must be related by the transformation laws obtained above. If we substitute these expressions, we get:

$[1 - (cv_x/V^2)^2]x^2 + [1 - (v_x/V)^2]y^2 + [1 - (v_x/V)^2]z$

$-[1 - (v_x/c)^2](ct)^2 - 2v_x[1 - (c/V)^2]xt = 0$

it is straightforward to confirm that this requires V = c, i.e. the general transformation with V = c, guarantees that the two spheres of light are expanding at the same rate, that is at the speed c, in both inertial frames of reference. Introducing a quantity γ defined by

$$\gamma = \frac{1}{\sqrt{1 - (v_x/c)^2}}$$

We are left with the final form of the transformation law consistent with the light always being observed to be travelling at the speed c in all reference frames:

$$x' = \gamma(\text{x} - v_x\text{t})$$

$$y' = \text{y}$$

$$z' = \text{z}$$

$$t' = \gamma(\text{t} - (v_x/c^2)x).$$

These are the equations of the Lorentz transformation. We can find the inverse transformation either by solving the above equation for x, y, z, and t in terms of $x^{'}$, $y^{'}$, $z^{'}$ and $t^{'}$, or else by simply recognizing, as was mentioned above in the derivation of this transformation, that if $S^{'}$ is moving with velocity v_x relative to S, then S is moving with velocity v_x relative to $S^{'}$. Consequently, all that is required is to exchange the primed and unprimed variables and change the sign of v_x in the above equation. The result of either method is

$$x = \gamma(x^{'} + v_x t^{'})$$

$$y^{'} = y$$

$$z^{'} = z$$

$$t = \gamma(t^{'} + (v_x/c^2)x^{'})$$

These equations were first obtained by Lorentz looking for a mathematical transformation that left Maxwell's equations unchanged in form. However, he did not assign any physical significance to his results. It was Einstein who first realized the true meaning of these equations and, consequently, with this greater insight, was able to derive them without reference at all to Maxwell's equations. The importance of his insight goes to the heart of relativity. Although the use of a flash of light played a crucial role in deriving the transformation equations, the final result simply establishes a connection between the two sets of space-time coordinates associated with a given event, this event being the passage of a flash of light past the point (x, y, z) at time t, as measured in S, or $(x^{'}, y^{'}, z^{'})$ at time $t^{'}$, as measured in $S^{'}$. The transformation equations, therefore, represent a property that space and time must have in order to guarantee that light will always be observed to have the same speed c in all inertial frames of reference. But given that these transformation equations represent an intrinsic property of space and time, it can only be expected that the behaviour of other material objects, which may have nothing whatsoever to do with light, will also be influenced by this fundamental property of space and time. This is the insight that Einstein had that the Lorentz transformation was saying something about the properties of space and time and the consequent behaviour that matter and forces must have in order to be consistent with these properties.

Later, we will see that the speed of light acts as an upper limit to how fast any material object can travel, be it light electrons or rocket ships. In addition, we shall see that anything that travels at this speed c will always be observed to do so from all frames of reference. Light just happens to be one of the things in the universe that travels at this particular speed. Subatomic particles called neutrinos also apparently travel at the speed of light, so we could have formulated our arguments above on the basis of an expanding sphere of neutrinos! The constant c, therefore represents a characteristic property of space and time, and only less significantly is it the speed at which light travels.

Two immediate conclusions can be drawn from the Lorentz Transformation. Firstly, suppose that $v_x > c$, i.e., that $S^{'}$ is moving relative to S at a speed greater than the speed of light. In that case, we find that $\gamma^2 < 0$, i.e., γ, is imaginary, so both position and time in Eq. (43) become imaginary. However, position and time are both physical quantities that must be measured as real numbers. In other words, the Lorentz transformation becomes physically meaningless if $v_x > c$. This immediately suggests that it is a physical impossibility for a material object to attain a speed greater than c relative to any reference frame S. The frame of reference in which such an object would be stationary will then also be moving at the speed v_x, but as we have just seen, in this situation, the transformation law breaks down. We shall see later how the laws of dynamics are modified in special relativity, one of the consequences of this modification being that no material object can be accelerated to a speed greater than c.

Secondly, we can consider the form of the Lorentz Transformation in the mathematical limit $v_x << c$. We find that $\gamma = 1$ so that Eq. (43) becomes the equations of the Galilean Transformation, Eq. (1). (Though this also requires that the x dependent term in the time transformation equation to be negligible, which it will be over small enough distances). Thus, at low enough speeds, any unusual results due to the Lorentz transformation would be unobservable.

Perhaps the most startling aspect of the Lorentz Transformation is the appearance of a transformation for time. The result obtained earlier for the Galilean Transformation agrees with, indeed, it was based on our 'common sense' notion that time is absolute, i.e. that time passes in a manner completely independent of the state of motion of any observer. This is certainly not the case with the Lorentz Transformation, which leads, as we shall see, to the conclusion that moving clocks run slow. This effect, called time dilation, and its companion effect, length contraction, will now be discussed.

2.1 Length Contraction

The first of the interesting consequences of the Lorentz Transformation is that length no longer has an absolute meaning: the length of an object depends on its motion relative to the frame of reference in which its length is being measured. Let us consider a rod moving with a velocity v_x relative to a frame of reference S and lying along the X axis. This rod is then stationary relative to a frame of reference $S^{'}$ which is also moving with a velocity v_x relative to S.

Figure 6: A rod of length at rest in reference frame $S^{'}$ which is moving with a velocity v_x with respect to another frame S.

As the rod is stationary in $S^{'}$, the ends of the rod will have coordinates $x^{'}_1 and x^{'}_2$ which remain fixed as functions of the time in $S^{'}$. The length of the rod, as
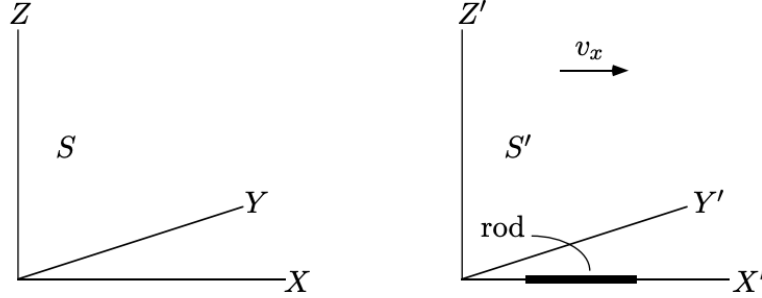
measured in $S^{'}$ is then

$$l_0 = \text{x}^{'}\,_2 - x^{'}\,_1$$

where l_0 is known as the proper length of the rod, i.e. l_0 is its length as measured in a frame of reference in which the rod is stationary. Now, suppose that we want to measure the length of the rod as measured with respect to S. In order to do this, we measure the X coordinates of the two ends of the rod at the same time, t, as measured by the clocks in S. Let x2 and x1 be the X coordinates of the two ends of the rod as measured in S at this time t. It is probably useful to be aware that we could rephrase the preceding statement in terms of the imaginary synchronized clocks introduced in Section 2.1 and Section 8 by saying that 'the two clocks positioned at x2 and x1 both read t when the two ends of the rod coincided with the points x2 and x1.' Turning now to the Lorentz Transformation equations, we see that we must have

$$\text{x}^{'}\,_1 = \gamma(x1 - \text{v}_x\text{t}))$$

$$\text{x}^{'}\,_2 = \gamma(x2 - \text{v}_x\text{t}))$$

We then define the length of the rod as measured in the frame of reference S to be

$$l = \text{x2 - x1}$$

where the important point to be re-emphasized is that this length is defined in terms of the positions of the ends of the rods as measured at the same time t in S. Using the above sets of equations, we find

$$l_0 = \text{x}^{'}\,_2 - x^{'}\,_1 = \gamma(x2 - x1) = \gamma\text{l}$$

we eventually get: l is less than l_0

Thus, the length of the rod as measured in the frame of reference S with respect

to which the rod is moving is shorter than the length as measured from a frame of reference S' relative to which the rod is stationary. A rod will be observed to have its maximum length when it is stationary in a frame of reference. The length so-measured, l_0, is known as its proper length.

This phenomenon is known as the Lorentz-Fitzgerald contraction. It is not the consequence of some force 'squeezing' the rod, but it is a real physical phenomenon with observable physical effects. Note however that someone who actually looks at this rod as it passes by will not see a shorter rod. If the time that is required for the light from each point on the rod to reach the observer's eye is taken into account, the overall effect is that of making the rod appears as if it is rotated in space.

2.2 Time Dilation

Perhaps the most unexpected consequence of the Lorentz transformation is the way in which our 'commonsense' concept of time has to be drastically modified. Consider a clock C' placed at rest in a frame of reference S' at some point x' on the X axis. Suppose once again that this frame is moving with a velocity v_x relative to some other frame of reference. At a time $t'_1 registered by clock C'$ there will be a clock C1 in the S frame of reference passing the position of C':
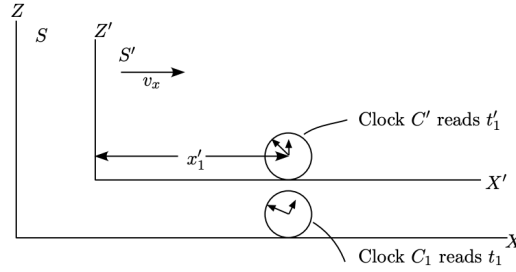

Figure 8: Clock C' stationary in S' reads t'_2 when it passes clock C2 stationary in S, at which instant C2 reads t2.

This clock C2 will read: $t2 = \gamma(t'_2 + v_x x'/c^2)$.

Thus, from the above equations, we have:

$$\Delta t = t2 - t1 = \gamma(t'_2 - t'_1) = \gamma\Delta t'$$

Once again, since

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v_x}{C}\right)^2}} > 1$$

if $v_x < $ C we have

$$\Delta \text{t} > \Delta t^{'}$$

In order to interpret this result, suppose that $\Delta t^{'}$ is the time interval between two 'ticks' of the clock $C^{'}$. Then, according to the clocks in S, these two 'ticks' are separated by a time interval t which, is $> \Delta t^{'}$. Thus, the time interval between 'ticks' is longer, as measured by the clocks in S, than what it is measured to be in $S^{'}$. In other words, from the point of view of the frame of reference S, the clock (and all the clocks in $S^{'}$) are running slow. It appears from S that time is passing more slowly in $S^{'}$ than it is in S. This is the phenomenon of time dilation. A clock will be observed to run at its fastest when it is stationary in a frame of reference. The clock is then said to be measuring proper time.

This phenomenon is just as real as length contraction. One of its best-known consequences is that of the increase in the lifetime of a radioactive particle moving at a speed close to that of light. For example, it has been shown that if the lifetime of a species of radioactive particle is measured while stationary in a laboratory to be T', then the lifetime of an identical particle moving relative to the laboratory is found to be given by T = γT', in agreement with Eq. (54) above.

2.3   Simultaneity

Another consequence of the transformation law for time is that events which occur simultaneously in one frame of reference will not in general occur simultaneously in any other frame of reference. Thus, consider two events, 1 and 2 which are simultaneous in S, i.e. t1 = t2, but which occur at two different places, x1 and x2. Then, in $S^{'}$, the time interval between these two events is

$$t_2' \text{ - } t_1' = \gamma(\text{t2}\ \ v_x\text{x2}/c^2)\ \ \gamma(\text{t1}\ \ v_x\text{x1}/c^2)$$
$$= \gamma(\text{x1 - x2})v_x/c^2$$

$$\neq 0 \, as \, x1 \neq x2$$

Here, t_1' is the time registered on the clock in $S^{'}$, which coincides with the position x1 in S at the instant t1 that the event 1 occurs and similarly for t_2'. Thus,

events which appear simultaneous in S are not simultaneous in $S^{'}$. In fact, the order in which the two events, 1 and 2, are found to occur will depend on the sign of x1 x2 or v_x. It is only when the two events occur at the same point (i.e. x1 = x2) that the events will occur simultaneously in all frames of reference.

Chapter 3

Lagrangian Mechanics

In the early centuries, Fermat posed his famous Fermat's Principle of Least Time, which states the principle governing the motion of light in a medium. This principle uses the fundamental "extremum principle" of physics. The extremum principle appears in almost all branches of physics. One such principle, "Hamilton's Principle", is one the most powerful tools we use in nearly all sectors of physics.

In Newtonian Mechanics, we have been solving equations of motion involving vectors. I will discuss a much more robust alternative to Newton's law, which is used in Special Relativity, General Relativity, etc., with some apparent modifications. Before going to the actual statement of the theorem, let's look at some terminologies used.

We define a strange combination of kinetic energy(T) and potential energy(V); T-V is Lagrangian(L). L = T-V ; This combination of kinetic and potential energy finds use in many advanced areas of physics. It often simplifies our problem. Instead of vectors, it uses scalar, i.e., kinetic and potential energy. Let's define another quantity Action ;

$$S[x(t)] = \int_a^b L(x, v, t)dt$$

where v = $\dot{x}$.

3.1   The Principle of Stationary Action

The principle of stationary action claims that among all the possible trajectories between the endpoints, the particle assumes the trajectory with stationary action (local maxima, local minima or a saddle point). Using this principle, we arrive at Euler's Lagrange equation,

$$\frac{d}{dt}(\frac{\partial L}{\partial v}) = \frac{\partial L}{\partial x}$$

, where

$$v_i = \dot{x}_i$$

Applying Euler Lagrange equation on a general kinetic energy and time-independent potential energy yields Newton's equation of motion, i.e.

$$m\ddot{x} = -\frac{\partial V}{\partial x}$$

;

Another unique property of Euler's Lagrange equation that makes it special is that if it holds in a coordinate $(x_1, x_2, ..., x_i)$, then it holds in any other ccoordinate

$$q_i = q(x_1, x_2, ..., x_i, t)$$

While using Newton's equation, we needed pseudo forces while working with generalised coordinates such as in rotating or accelerating frames; we are allowed to use Euler's Lagrange equation for any general coordinate without worrying about its explicit time or position dependence.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) = \frac{\partial L}{\partial q}$$

, where

$$v_i = \dot{q}_i$$

3.2   Symmetry and conservation laws

I am being symmetric concerning a specific coordinate implies that the quantity is invariant concerning a change in that coordinate, i.e.

$$\frac{dQ}{dq} = 0$$

. Suppose the Lagrangian of a system is invariant concerning a coordinate q_i, then for ith coordinate;

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) = \frac{\partial L}{\partial q} = 0$$

, this implies

$$\frac{\partial L}{\partial v} = c$$

, which is independent of time . In other words, a quantity $\frac{\partial L}{\partial v}$ is a conserved quantity. Suppose Lagrangian is symmetric about x_i co-ordinate , then $\frac{\partial L}{\partial v} = m\dot{x}$ is conserved ($v_i = \dot{x}$) . Similarly, corresponding to a quantity q_i, we call $\frac{\partial L}{\partial v}$ as the generalised momentum corresponding to coordinate q_i, and its dimension needs not be MLT^{-1}.

3.3 Time Translation Symmetry

When Lagrangian is symmetric concerning time translation, then the Energy of the system is conserved. We can't use the same analysis that we did for any other spatial coordinate because time is something that flows without any interrupt; we can't access control on the flow of time, i.e. we aren't allowed to assume any random evolution of time coordinate as we did for any other coordinate. So, before analysis, we define a quantity Hamiltonian 'H' as

$$H = \sum_{i=1}^{n} \dot{q} \frac{\partial L}{\partial v} - L$$

. Generally, Hamiltonian represents the total energy of a particle when the coordinates q_i are independent of time and $\dot{x}$. Therefore, we will use E instead of H for the rest of our purposes. After calculating, we get,

$$\frac{\partial E}{\partial t} = -\frac{\partial L}{\partial t}$$

when Lagrangian is symmetric about time coordinates. Therefore, Energy conservation is a consequence of time translation symmetry.

3.4 Noether's Theorem

This leads to one of the most beautiful results in physics. For every continuous symmetry in the system, there exists a conservation law.