Infinitesimal Lorentz Transformations & A Relativistic Hamiltonian

Lecture 10

Physics 411
Classical Mechanics II

September 19th, 2007

We have the relativistic Lagrangian, or at least one form of the relativistic Lagrangian. The question now becomes, what are the conserved quantities? Or, which is equivalent, what are the transformations that leave the associated Hamiltonian unchanged? We already know the answer to this, as the Lorentz transformations were developed to leave the metric unchanged, we know that the Lagrangian is insensitive to frame. But, it is of theoretical utility, at the very least, to carefully develop the connection between a general transformation like Lorentz boosts or spatial rotations, and their infinitesimal counterparts – these linearized transformations are the ones that are relevant to generators, and hence to constants of the motion. So we start by establishing, for rotations and Lorentz boosts, that it is possible to build up a general rotation (boost) out of infinitesimal ones.

We can then sensibly discuss the generators of infinitesimal transformations as a stand-in for the full transformation. But in order to check that the Poisson bracket of a generator with the Hamiltonian vanishes, we must also have the Hamiltonian – this is not so easy to develop, so we spend some time discussing the role of the Hamiltonian in free particle relativistic mechanics.

10.1 Infinitesimal Transformations

We will look at how infinitesimal transformations can be used to build up larger ones – this is an important idea in physics, since it allows us to focus on simple, linearized forms of more general transformations to understand the action of a transformation on a physical system.

10.1.1 Rotations

Going back to three dimensional space with Cartesian coordinates, we saw that an infinitesimal rotation through an angle $d\Omega$ about an axis $\hat{\boldsymbol{\Omega}}$ as shown in Figure 10.1 could be written as (rotating the coordinate system, now, rather than a vector – hence the change in sign)

$$\bar{\mathbf{x}} = \mathbf{x} - d\Omega\,\hat{\boldsymbol{\Omega}} \times \mathbf{x}. \tag{10.1}$$

Suppose we take $\hat{\boldsymbol{\Omega}} = \hat{z}$, so the rotation occurs in the $x - y$ plane, and we

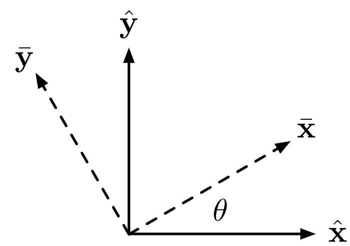

Figure 10.1: Two coordinate systems related by a rotation.

let $d\Omega \equiv \theta$. In order to keep track of the number of successive applications of this transformation, we can use a superscript on the coordinate axes, so one rotation is:

$$\mathbf{x}^1 = \mathbf{x} - \theta\,\hat{z} \times \mathbf{x} \tag{10.2}$$

with $\mathbf{x} \equiv x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}}$. This operation can be written as matrix-vector multiplication:

$$\begin{pmatrix} x^1 \\ y^1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{10.3}$$

Suppose we act on the $\mathbf{x}^1$ coordinates with another infinitesimal rotation:

$$\begin{aligned} \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} &= \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} + \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} \\ &= \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} \right]^2 \begin{pmatrix} x \\ y \end{pmatrix}, \end{aligned} \tag{10.4}$$

and the pattern is clear – for the n^{th} iterate, we have:

$$\begin{pmatrix} x^n \\ y^n \end{pmatrix} = \left[\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\equiv \mathbb{I}} + \theta \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\equiv \mathbb{A}} \right]^n \begin{pmatrix} x \\ y \end{pmatrix}. \tag{10.5}$$

Suppose we want to rotate through a net angle Θ using n iterations, then $\theta = \frac{\Theta}{n}$, and we focus on the matrix term – it is easy enough to determine the powers of this simple matrix:

$$\left(\mathbb{I} + \frac{\Theta}{n}\,\mathbb{A}\right)^n = \left[\frac{1}{2}\begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^T \begin{pmatrix} \frac{n-i\Theta}{n} & 0 \\ 0 & \frac{n+i\,\Theta}{n} \end{pmatrix} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^*\right]^n, \tag{10.6}$$

and noting the identity:

$$\mathbb{I} = \frac{1}{2}\begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^* \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^T, \tag{10.7}$$

we can write:

$$\left(\mathbb{I} + \frac{\Theta}{n}\,\mathbb{A}\right)^n = \frac{1}{2}\begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^T \begin{pmatrix} \left(\frac{n-i\Theta}{n}\right)^n & 0 \\ 0 & \left(\frac{n+i\Theta}{n}\right)^n \end{pmatrix} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^*. \tag{10.8}$$

Finally, we want to take the limit of this expression as $n \longrightarrow \infty$. To do this, consider the binomial expansion of the terms in the diagonal matrix:

$$\left(1 - \frac{i\,\Theta}{n}\right)^n \sim 1 - i\Theta - \frac{1}{2}\,n\,(n-1)\left(\frac{\Theta}{n}\right)^2 + \frac{1}{6}\,i\,n\,(n-1)\,(n-2)\left(\frac{\Theta}{n}\right)^3 + \ldots, \tag{10.9}$$

as $n \longrightarrow \infty$, we have:

$$\begin{aligned}\left(1 - \frac{i\,\Theta}{n}\right)^n &\longrightarrow 1 - i\,\Theta - \frac{1}{2}\,\Theta^2 + \frac{1}{6}\,i\,\Theta^3 + \ldots \\ &= e^{-i\,\Theta}.\end{aligned} \tag{10.10}$$

Finally, then, the limit of the full matrix is

$$\begin{aligned}\left(\mathbb{I} + \frac{\Theta}{n}\,\mathbb{A}\right)^n &\longrightarrow \frac{1}{2}\begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^T \begin{pmatrix} e^{-i\,\Theta} & 0 \\ 0 & e^{i\,\Theta} \end{pmatrix} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}^* \\ &= \begin{pmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{pmatrix}.\end{aligned} \tag{10.11}$$

The point, in the end, is that to build up a net rotation through an angle Θ from the infinite application of infinitesimal θ, we get

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{10.12}$$

as always.

That is heartening, but what is so great about this procedure? Going back to our discussion of infinitesimal generators and Hamiltonians, we know that the infinitesimal coordinate transformation:

$$\mathbf{X} = \mathbf{x} + \boldsymbol{\Omega} \times \mathbf{x} \tag{10.13}$$

leaves the Hamiltonian unchanged, and hence is associated with a conserved quantity. Using indexed notation,

$$X^\alpha = x^\alpha + \theta\, \epsilon_{\alpha\beta\gamma}\, \Omega^\beta\, x^\gamma, \tag{10.14}$$

and in the language of generators, we have coordinate transformation generated by[1] $J = p_\alpha\, f^\alpha$:

$$\begin{aligned} J &= p^\alpha\, \epsilon_{\alpha\beta\gamma}\, \Omega^\beta\, x^\gamma = \Omega^\beta\, \epsilon_{\beta\gamma\alpha}\, x^\gamma\, p^\alpha \\ &= \frac{1}{2}\, \Omega^\beta\, \epsilon_{\beta\gamma\alpha}\, (x^\gamma\, p^\alpha - p^\gamma\, x^\alpha)\,. \end{aligned} \tag{10.15}$$

For this reason, we often say that $M^{\alpha\beta} = x^\alpha\, p^\beta - p^\alpha\, x^\beta$ is the generator of infinitesimal rotations (there are three of them, selected by the axis represented in $\boldsymbol{\Omega}$). As a doubly-indexed object, it appears there are 9 components to the matrix $\mathbb{M}$, but it is antisymmetric, so only has three independent components.

Note that we could have simplified life by starting from the full transformation, which is known to be invariant, and linearizing about the parameter Θ:

$$\begin{pmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & \Theta \\ -\Theta & 0 \end{pmatrix} + O(\Theta^2) \tag{10.16}$$

10.1.2 Infinitesimal Lorentz Transformations

If we consider a $D = 1 + 1$ dimensional Lorentz "boost" along a shared $\hat{\mathbf{x}}$ axis, then the matrix representing the transformation is:

$$\begin{pmatrix} c\bar{t} \\ \bar{x} \end{pmatrix} = \begin{pmatrix} \cosh\eta & \sinh\eta \\ \sinh\eta & \cosh\eta \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{10.17}$$

[1] We will be a little sloppy with indices in the following expression, so that the Levi-Civita symbol's role is clear. Remember that we are in flat Euclidean space with Cartesian coordinates, so there is no numerical distinction between contravariant and covariant tensors.

where η is the rapidity, and $\cosh\eta = \gamma$, $\sinh\eta = -\beta\gamma$ for $\beta \equiv v/c$. If we take the infinitesimal parameter to be precisely η (a boost to a frame moving "slowly"), then we can develop the expansion to linear order in η (compare with (10.16)):

$$\begin{pmatrix} c\bar{t} \\ \bar{x} \end{pmatrix} \sim \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & \eta \\ \eta & 0 \end{pmatrix} \right] \begin{pmatrix} ct \\ x \end{pmatrix}. \tag{10.18}$$

This gives us the infinitesimal transformation associated with Lorentz boosts – but what is its form analagous to (10.14)? Suppose we try to represent the linearized Lorentz boost as a generalized cross product. In $D = 3+1$, we of course have Levi-Civita with four indices: $\epsilon^{\alpha\beta\gamma\delta}$ defined as usual. The linear transformation should be . . . linear in x^μ, of course, but then we need a doubly-indexed object to soak up the additional indices – our "reasoning" gives, thus far:

$$X^\alpha = x^\alpha + \eta\, \epsilon^{\alpha\beta\gamma\delta}\, Q_{\beta\gamma}\, x_\delta. \tag{10.19}$$

Now suppose we exploit the analogy with rotation further – in order to rotate the $x-y$ plane, we needed a vector $\hat{\mathbf{\Omega}} \sim \hat{\mathbf{z}}$ for (10.1). In the full four-dimensional space-time, there are two directions orthogonal to the $t-x$ plane, both y and z, so why not set $Q_{\beta\gamma} = \delta_{\beta y}\delta_{\gamma z}$? In this specific case, we recover (referring to the transformed coordinates using capital letters, here, and putting the c directly into the temporal coordinate so that T is shorthand for $c\,T$):

$$T = t + \eta x \qquad X = x + \eta t \qquad Y = y \qquad Z = z, \tag{10.20}$$

as desired. In general, $Q_{\beta\gamma}$ plays the same role as Ω^α did for rotations. Only the antisymmetric part of $Q_{\beta\gamma}$ contributes, so we expect to have six independent quantities here. As icing on the cake, notice that if we set $Q_{\beta\gamma} = \delta_{\beta 0}\,\delta_{\gamma 3}$, we get:

$$T = t \qquad X = x + \eta\, y \qquad Y = y - \eta\, x \qquad Z = z, \tag{10.21}$$

precisely the infinitesimal rotation we studied before. So both boosts and rotations are included here (as they should be).

Turning to the generator of the above transformation, we have

$$\begin{aligned} J &= \epsilon^{\alpha\beta\gamma\delta}\, Q_{\beta\gamma}\, x_\delta\, p_\alpha \\ &= Q_{\beta\gamma}\, \epsilon^{\beta\gamma\delta\alpha}\, x_\delta\, p_\alpha. \end{aligned} \tag{10.22}$$

We see that there are six independent generators, depending on our choice of $Q_{\beta\gamma}$. Finally, we associate the tensor:

$$\boxed{M^{\alpha\beta} = x^\alpha\, p^\beta - p^\alpha\, x^\beta} \tag{10.23}$$

with the six conserved quantities.

We've strayed relatively far from our starting point, and the preceding argument is meant to motivate only – we will show that this set of quantities is indeed conserved, but first we have to develop the Hamiltonian associated with the relativistic Lagrangian.

10.2 Relativistic Hamiltonian

In order to show that the generator of Lorentz transformations is conserved, we need the Hamiltonian. This has its own physics as we shall see by moving back and forth between proper time and coordinate time parametrization.

Remember, we are only dealing with the free particle Lagrangian (and Hamiltonian) – we have yet to introduce any potentials. So, start with

$$L = -m\,c\,\sqrt{-\dot{x}^\mu(\tau)\,\eta_{\mu\nu}\,\dot{x}^\nu(\tau)}, \tag{10.24}$$

and we want the Legendre transform of this. In preparation, define the canonical momenta as usual:

$$p_\alpha = \frac{\partial L}{\partial \dot{x}^\alpha} = \frac{m\,c\,\dot{x}^\mu\,\eta_{\mu\alpha}}{\sqrt{-\dot{x}^\mu\,\eta_{\mu\nu}\,\dot{x}^\nu}}. \tag{10.25}$$

Now, in proper time parametrization, we have, by definition:

$$-c^2\,dt^2 + dx^2 + dy^2 + dz^2 = -c^2\,d\tau^2 \longrightarrow c^2\,\dot{t}^2 - \dot{x}^2 - \dot{y}^2 - \dot{z}^2 = c^2, \tag{10.26}$$

so that $\sqrt{-\dot{x}^\mu\,\eta_{\mu\nu}\,\dot{x}^\nu} = c$, and we can use this to simplify the momenta:

$$p_\alpha = m\,\dot{x}^\mu\,\eta_{\mu\alpha} = m\,\dot{x}_\alpha. \tag{10.27}$$

In order to understand the physics of this new four-vector, we can go back to temporal parametrization, then p_0 and p_1 give a sense for the whole vector

– remember that $\dot{t} = \sqrt{\frac{1}{1-v^2/c^2}}$:

$$\begin{aligned} p_0 &= -c\,\frac{dt}{d\tau} = -\frac{m\,c}{\sqrt{1-\frac{v^2}{c^2}}} \\ p_1 &= m\,\frac{dx}{d\tau} = m\,\frac{dx}{dt}\dot{t} = \frac{m\,v^x}{\sqrt{1-\frac{v^2}{c^2}}}, \end{aligned} \tag{10.28}$$

where v^x is the velocity in coordinate time $\frac{dx(t)}{dt}$ and v^2 is the coordinate-time velocity magnitude. So we get the usual sort of contravariant form:

$$\boxed{p^\alpha \doteq \begin{pmatrix} \frac{m\,c}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{m\,\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix}.} \tag{10.29}$$

10.2.1 The Physics of p^0

Before moving on to the full relativistic Hamiltonian, we will make contact with the "relativistic classical" result by considering the Lagrangian back in temporal parametrization, and forming the Hamiltonian as we normally would.

$$L^* = -m\,c^2\sqrt{1-\frac{v^2}{c^2}} \tag{10.30}$$

has conjugate momenta

$$\mathbf{p} = \frac{\partial L^*}{\partial \mathbf{v}} = \frac{m\,\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{10.31}$$

so that the Hamiltonian, representing the total energy is:

$$H^* = \mathbf{p}\cdot\mathbf{v} - L = \frac{m\,v^2}{\sqrt{1-\frac{v^2}{c^2}}} + m\,c^2\sqrt{1-\frac{v^2}{c^2}} \quad = \frac{m\,c^2}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{10.32}$$

This total energy can be broken down into the rest energy of a particle with $v = 0$, namely $m\,c^2$ and "the kinetic" portion, $H^* - m\,c^2$. Regardless, we see that p^0 is precisely the total energy (divided by c):

$$p^\alpha \doteq \begin{pmatrix} \frac{E}{c} \\ \frac{m\,\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix}. \tag{10.33}$$

You have guessed, I hope, the reason for the delay – if we take the canonical momentum defined by (10.25), and attempt to form the relativistic Hamiltonian, we get:

$$p_\alpha \, \dot{x}^\alpha = -m \, c \, \sqrt{-\dot{x}^\mu \, \eta_{\mu\nu} \, \dot{x}^\nu}, \tag{10.34}$$

and this is precisely L, so subtracting L from it will give $H = p_\alpha \, \dot{x}^\alpha - L = 0$, not a very useful result for dynamics. The problem, if one can call it that, is in the form of the Lagrangian – basically its reparametrization invariance – so useful for the interpretation of equations of motion (via proper time), is in a way, too large to define H. To put it another way, there are a wide variety of Lagrangians that describe free particle special relativity and reduce to the classical one in some limit.

For now, we can focus on generating a functionally efficient Hamiltonian. Consider, for example, the equations of motion that come from our $L = -m \, c \, \sqrt{-\dot{x}^\mu \, \eta_{\mu\nu} \, \dot{x}^\nu}$ – in proper time parametrization, they reduce to:

$$m \, \ddot{x}^\mu = 0. \tag{10.35}$$

So we are, once again, describing a "straight line", the general solution is $x^\mu = A^\mu \, \tau + B^\mu$ with constants used to set initial conditions and impose the proper time constraint.

From this, we see that there are other Lagrangians that would lead to the same equation of motion – for example $L^+ = \frac{1}{2} \, m \, \dot{x}^\mu \, \eta_{\mu\nu} \, \dot{x}^\nu$. And this Lagrangian has the nice property that the Hamiltonian is nontrivial. We now have canonical momenta: $p_\alpha = \frac{\partial L^+}{\partial \dot{x}^\alpha} = m \, \dot{x}^\mu \, \eta_{\mu\alpha}$. The Legendre transform of the Lagrangian is:

$$\boxed{H^+ = \frac{1}{2 \, m} p_\alpha \, \eta^{\alpha\beta} \, p_\beta} \tag{10.36}$$

and the equations of motion are precisely $m \, \ddot{x}^\mu = 0$. Operationally, then, this Hamiltonian does what we want. In particular, it provides a way to check the constancy of the infinitesimal Lorentz generators, our original point.

We will check them directly from (10.23), without the baggage of additional constants:

$$
\begin{aligned}
[H^+, M^{\alpha\beta}] &= [H^+, x^\alpha\, p^\beta] - [H^+, x^\beta\, p^\alpha] \\
&= \frac{\partial H^+}{\partial x^\gamma}\,\frac{\partial x^\alpha\, p^\beta}{\partial p_\gamma} - \frac{\partial H^+}{\partial p_\gamma}\,\frac{\partial x^\alpha p^\beta}{\partial x^\gamma} - (\alpha \leftrightarrow \beta) \\
&= -\frac{1}{m}\, p^\gamma\, \delta^\alpha_\gamma\, p^\beta + \frac{1}{m}\, p^\gamma\, \delta^\beta_\gamma\, p^\alpha \\
&= 0,
\end{aligned}
\tag{10.37}
$$

so indeed, we have $M^{\alpha\beta}$ constant (all six of them).