Chapter 8

Field Theory

8.1 Lagrangian Mechanics for Fields

In sections 5.3 and 5.4.1 we considered the continuum limit of a chain of point masses on stretched string. We had a situation in which the potential energy had interaction terms for particle A which depended only on the relative displacements of particles in the neighborhood of A. If we take our coordinates to be displacements from equilibrium, and consider only motions for which the displacement $\eta = \eta(x, y, z, t)$ becomes differentiable in the continuum limit, then the leading term in the potential energy is proportional to the square of derivatives in the spatial coordinates. For our points on a string at tension τ, with mass density ρ, we found

$$\begin{aligned} T &= \frac{1}{2}\rho \int_0^L \dot{y}^2(x)dx, \\ U &= \frac{\tau}{2}\int_0^L \left(\frac{\partial y}{\partial x}\right)^2 dx, \end{aligned}$$

and we can write the Lagrangian as an integral of a **Lagrangian density** $\mathcal{L}(y, \dot{y}, y', x, t)$ over x. Actually for our string we had no y or x or t dependence, because we ignored gravity $U_g = \int \rho g y(x, t)dx$, and had a homogeneous string whose properties were also time independent. In general, however, such dependence is quite possible. In section 5.4.2, we considered a three dimensional object, and discussed the equations for the displacement of the atoms in a crystal. Then the fields $\vec{\eta}$ were the three components of the displacement of a particle, as a function of the three coordinates (x, y, z)

determining the particle, as well as time. Thus the generalized coordinates are the functions $\eta_i(x, y, z, t)$, and the Lagrangian density will depend on these, their gradients, their time derivatives, as well as possibly on x, y, z, t. Thus

$$\mathcal{L} = \mathcal{L}(\eta_i, \frac{\partial \eta_i}{\partial x}, \frac{\partial \eta_i}{\partial y}, \frac{\partial \eta_i}{\partial z}, \frac{\partial \eta_i}{\partial t}, x, y, z, t)$$

and

$$\begin{aligned} L &= \int dx\, dy\, dz\, \mathcal{L}, \\ I &= \int dx\, dy\, dz\, dt\, \mathcal{L}. \end{aligned}$$

The actual motion of the system will be given by a particular set of functions $\eta_i(x, y, z, t)$, which are functions over the volume in question and of $t \in [t_I, t_f]$. The function will be determined by the laws of dynamics of the system, together with boundary conditions which depend on the initial configuration $\eta_i(x, y, z, t_I)$ and perhaps a final configuration. Generally there are some boundary conditions on the spatial boundaries as well. For example, our stretched string required $y = 0$ at $x = 0$ and $x = L$, for all values of t.

Before taking the continuum limit we say that the configuration of the system at a given t was a point in a large N dimensional configuration space, and the motion of the system is a path $\Gamma(t)$ in this space. In the continuum limit $N \to \infty$, so we might think of the path as a path in an infinite dimensional space. But we can also think of this path as a mapping $t \to \eta(\cdot, \cdot, \cdot, t)$ of time into the (infinite dimensional) space of functions on ordinary space.

Hamilton's principal says that the actual path is an extremum of the action. If we consider small variations $\delta\eta_i(x, y, z, t)$ which vanish on the boundaries, then

$$\delta I = \int dx\, dy\, dz\, dt\, \delta\mathcal{L} = 0$$

determines the equations of motion.

Note that what is varied here are the functions η_i, not the coordinates (x, y, z, t). x, y, z do not represent the position of some atom — they represent a label which tells us which atom it is that we are talking about. Often they are chosen to be the equilibrium position of that atom, but they are fixed labels independent of the motion. It is the $\eta_i(\vec{x})$, for each $\vec{x}$, which are the dynamical degrees of freedom, specifying the configuration of the system. In our discussion of section 5.4 η_i specified the displacement from equilibrium, but here we generalize to an arbitrary set of dynamical fields[1].

The variation of the Lagrangian density is

$$\begin{aligned}
&\delta\mathcal{L}(\eta_i, \frac{\partial\eta_i}{\partial x}, \frac{\partial\eta_i}{\partial y}, \frac{\partial\eta_i}{\partial z}, \frac{\partial\eta_i}{\partial t}, x, y, z, t) \\
&\quad = \sum_i \frac{\partial\mathcal{L}}{\partial\eta_\rangle}\delta\eta_i + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_\rangle/\partial\S)}\delta\frac{\partial\eta_i}{\partial x} + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_\rangle/\partial\dagger)}\delta\frac{\partial\eta_i}{\partial y} \\
&\qquad + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_\rangle/\partial\ddagger)}\delta\frac{\partial\eta_i}{\partial z} + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial\eta_\rangle/\partial\sqcup)}\delta\frac{\partial\eta_i}{\partial t}.
\end{aligned}$$

Notice there is no variation of x, y, z, and t, as we discussed.

The notation is getting awkward, so we need to reintroduce the notation $A_{,j} = \partial A/\partial r_j$, for $r_j = (x, y, z)$. In fact, we see that $\partial/\partial t$ enters in the same way as $\partial/\partial x$, so it is time to introduce notation which will become crucial when we consider relativistic dynamics, even though we are not doing so here. So we will consider time to be an additional component of the position, called the zeroth rather than the fourth component. We will also change our notation for coordinates to anticipate needs from relativity, by writing the indices of coordinates as superscripts rather than subscripts. Thus we write $x^0 = ct$, where c will eventually be taken as the speed of light, but for the moment is an arbitrary scaling factor. Until we get to special relativity, one should consider whether an index is raised or lowered as irrelevant, but they are written here in the place which will be correct once we make the distinction between them. In particular the Kronecker delta is now written ${\delta_\mu}^\nu$. For the partial derivatives we now have

$$\partial_\mu := \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{c\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right),$$

for $\mu = 0, 1, 2, 3$, and write $\eta_{,\mu} := \partial_\mu\eta$. If there are several fields η_i, then $\partial_\mu\eta_i = \eta_{i,\mu}$. The comma represents the beginning of differentiation, so we must not use one to separate different ordinary indices.

In this notation, we have

$$\delta\mathcal{L} = \sum_i \frac{\partial\mathcal{L}}{\partial\eta_i}\delta\eta_i + \sum_i \sum_{\mu=0}^{3} \frac{\partial\mathcal{L}}{\partial\eta_{i,\mu}}\delta\eta_{i,\mu},$$

[1]Note in particular that $\{\eta_i\}$ is not the set of coordinates of phase space as it was in the last chapter.

and

$$\delta I = \int \left(\sum_i \frac{\partial \mathcal{L}}{\partial \eta_i} \delta \eta_i + \sum_i \sum_{\mu=0}^{3} \frac{\partial \mathcal{L}}{\partial \eta_{i,\mu}} \delta \eta_{i,\mu} \right) d^4x,$$

where[2] $d^4x = c dx\, dy\, dz\, dt$. Except for the first term, we integrate by parts,

$$\delta I = \int \left[\sum_i \frac{\partial \mathcal{L}}{\partial \eta_i} - \sum_i \sum_{\mu=0}^{3} \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial \eta_{i,\mu}} \right) \right] \delta \eta_i d^4x,$$

where we have thrown away the boundary terms which involve $\delta\eta_i$ evaluated on the boundary, which we assume to be zero. Inside the region of integration, the $\delta\eta_i$ are independent, so requiring $\delta I = 0$ for all functions $\delta\eta_i(x^\mu)$ implies

$$\sum_\mu \frac{d}{dx^\mu} \frac{\partial \mathcal{L}}{\partial \eta_{i,\mu}} - \frac{\partial \mathcal{L}}{\partial \eta_i} = 0. \tag{8.1}$$

We have written the equations of motion (which are now partial differential equations rather than coupled ordinary differential equations), in a form which looks like we are dealing with a relativistic problem, because t and spatial coordinates are entering in the same way. We have not made any assumption of relativity, however, and our problem will not be relativistically invariant unless the Lagrangian density is invariant under Lorentz transformations (as well as translations).

Now consider how the Lagrangian changes from one point in space-time to another, including the variation of the fields, assuming the fields obey the equations of motion. Then the total derivative for a variation of x^μ is given by the chain rule

$$\frac{d\mathcal{L}}{dx^\mu} = \left.\frac{\partial \mathcal{L}}{\partial x^\mu}\right|_\eta + \sum_i \frac{\partial \mathcal{L}}{\partial \eta_i} \eta_{i,\mu} + \sum_{\nu i} \frac{\partial \mathcal{L}}{\partial \eta_{i,\nu}} \eta_{i,\nu,\mu}.$$

As we did previously with d/dt, we are using "total" derivative notation d/dx^μ to represent the variation from a change in one x^μ, including the changes induced in the fields which are the arguments of $\mathcal{L}$, though it is still a partial derivative in the sense that the other three x^ν need to be held fixed while varying x^μ.

---

[2] We have also multiplied I by c, which does no harm in finding the extrema.

Plugging the equations of motion into the second term,

$$\begin{aligned}
\frac{d\mathcal{L}}{dx^\mu} &= \frac{\partial\mathcal{L}}{\partial x^\mu} + \sum_i \left[\sum_\nu \frac{d}{dx^\nu}\left(\frac{\partial\mathcal{L}}{\partial\eta_{i,\nu}}\right)\right]\eta_{i,\mu} + \sum_{i\nu}\frac{\partial\mathcal{L}}{\partial\eta_{i,\nu}}\eta_{i,\mu,\nu} \\
&= \frac{\partial\mathcal{L}}{\partial x^\mu} + \sum_\nu \frac{d}{dx^\nu}\left(\sum_i \frac{\partial\mathcal{L}}{\partial\eta_{i,\nu}}\eta_{i,\mu}\right).
\end{aligned}$$

Thus

$$\sum_\nu \frac{d}{dx^\nu} T_\mu{}^\nu = -\frac{\partial\mathcal{L}}{\partial x^\mu}, \tag{8.2}$$

where the **stress-energy** tensor $T_\mu{}^\nu$ is defined by

$$T_\mu{}^\nu(x) = \sum_i \frac{\partial\mathcal{L}}{\partial\eta_{i,\nu}}\eta_{i,\mu} - \mathcal{L}\delta_\mu{}^\nu. \tag{8.3}$$

We will often talk about $T_\mu{}^\nu$ as a function of x^ρ, understanding that x dependence to include the implicit dependence through the fields, for T is a function of $x^\mu, \eta_i(x)$ and $\eta_{i,\mu}(x)$. It is that total derivative that we are evaluating on the left of equation (8.2), but it is often written as a partial derivative, $\sum \partial_\nu T_\mu{}^\nu = -\partial_\mu\mathcal{L}$, understanding that $T_\mu{}^\nu(x^\rho)$ depends on x^ρ through its field dependence as well as any explicit dependence. But the partial derivatives on the right of that equation do not include the variations through the fields. Sorry about that, it is just the way it is always written.

Note that if the Lagrangian density has no *explicit* dependence on the coordinates x^μ, equation (8.2) tells us the stress-energy tensor satisfies an equation $\sum_\nu \partial_\nu T_\mu{}^\nu = 0$ which is a continuity equation.

What does that mean? In fluid mechanics, we have the equation of continuity

$$\partial\rho/\partial t + \vec{\nabla}\cdot(\rho\vec{v}) = 0,$$

which expresses the conservation of mass. That equation has the interpretation that the rate of change in the mass contained in some volume is equal to the flux into the volume, because $\rho\vec{v}$ is the flow of mass outward past a unit surface area. In general, if we have a scalar field $\rho(\vec{x}, t)$ which, together with a vector field $\vec{j}(\vec{x}, t)$, satisfies the equation

$$\frac{\partial\rho}{\partial t}(\vec{x}, t) + \nabla\cdot\vec{j}(\vec{x}, t) = 0, \tag{8.4}$$

we can interpret ρ as the density of, and $\vec{j}$ as the flow of, a material property which is conserved. Given any volume V with a boundary surface S, the rate at which this property is flowing out of the volume, $\int_S \vec{j}\cdot d\vec{S} = \int_V \nabla\cdot\vec{j}\,dV$, is the rate at which the total amount of the substance in the volume is decreasing, $\int_V -(d\rho/dt)dV$. If we define $j^0 = c\rho$, we can rewrite this equation of continuity (8.4), as $\sum_\nu \partial_\nu j^\nu = 0$, and we say that j^ν is a conserved current[3]. If we integrate over the whole volume of our field, we can define a total "charge" $Q(t) = \int_V j^0(\vec{x},t)/c\, d^3x$, and its time derivative is

$$\frac{d}{dt}Q(t) = \int_V \frac{d\rho}{dt}(\vec{x},t)\,d^3x = -\int_V \nabla\cdot\vec{j}(\vec{x},t)\,d^3x = -\int_S \vec{j}\cdot d\vec{S}.$$

We see that this is the integral of the divergence of a vector current $\vec{j}$, which by Gauss' law becomes a surface integral of the flux of j out of the volume of our system. We have been sloppy about our boundary conditions, but in many cases it is reasonable to assume there is no flux out of the entire volume, either because of boundary conditions, as in a stretched string, or because we are working in an infinite space and expect any flux to vanish at infinity. Then the surface integral vanishes, and we find that the charge Q is conserved.

We have seen that when the lagrangian density has no explicit x^μ dependence, for each value of μ, ${T_\mu}^\nu$ represents such a conserved current. Thus we should have four conserved currents $(J_\mu)^\nu := {T_\mu}^\nu$, each of which gives a conserved "charge"

$$Q_\mu(t) = \int_V {T_\mu}^0(\vec{x},t)\,d^3x = \text{constant}.$$

We will return to what these conserved quantities are in a moment.

In dynamics of discrete systems we defined the momenta $p_i = \partial L/\partial\dot{q}_i$, and defined the Hamiltonian as $H = \sum_i p_i\dot{q}_i - L(q,p,t)$. In considering the continuum limit of the loaded string, we noted that the momentum corresponding to each point particle (of vanishing mass) disappears in the limit, but the appropriate thing to do is define a momentum density

$$P(x) = \frac{\delta}{\delta\dot{y}(x)}L = \frac{\delta}{\delta\dot{y}(x)}\int \mathcal{L}(y(x'),\dot{y}(x'),x',t)dx' = \left.\frac{\partial\mathcal{L}}{\partial\dot{y}}\right|_x,$$

having defined both the "variation at a point" $\delta/\delta\dot{y}(x)$ and the lagrangian density $\mathcal{L}$. In considering the three dimensional continuum as a limit, say

[3]More accurately, the set of four fields $j^\nu(\vec{x},t)$ is a conserved current.

on a cubic lattice, $L = \int d^3x\mathcal{L}$ is the limit of $\sum_{ijk} \Delta x\Delta y\Delta z L_{ijk}$, where L_{ijk} depends on $\vec{\eta}_{ijk}$ and a few of its neighbors, and also on $\dot{\vec{\eta}}_{ijk}$. The conjugate momentum to $\vec{\eta}(i,j,k)$ is $\vec{p}_{ijk} = \partial L/\partial\dot{\vec{\eta}}_{ijk} = \Delta x\Delta y\Delta z\partial L_{ijk}/\partial\dot{\vec{\eta}}_{ijk}$, which would vanish in the continuum limit. So we define instead the momentum density

$$\pi_\ell(x,y,z) = (\vec{p}_{ijk})_\ell/\Delta x\Delta y\Delta z = \partial L_{ijk}/\partial(\dot{\vec{\eta}}_{ijk})_\ell = \partial\mathcal{L}/\partial\dot{\eta}_\ell(x,y,z).$$

The Hamiltonian

$$\begin{aligned} H &= \sum \vec{p}_{ijk}\cdot\dot{\vec{\eta}}_{ijk} - L = \sum \Delta x\Delta y\Delta z\vec{\pi}(x,y,z)\cdot\dot{\vec{\eta}}(xyz) - L \\ &= \int d^3x \left(\vec{\pi}(\vec{r})\cdot\dot{\vec{\eta}}(\vec{r}) - \mathcal{L}\right) = \int d^3x\,\mathcal{H}, \end{aligned}$$

where the **Hamiltonian density** is defined by $\mathcal{H}(\vec{r}) = \vec{\pi}(\vec{r})\cdot\dot{\vec{\eta}}(\vec{r}) - \mathcal{L}(\vec{r})$. This assumed the dynamical fields were the vector displacements $\vec{\eta}(\vec{r},t)$, but the same discussion applies to any set of dynamical fields $\eta_\ell(\vec{r},t)$, even if η refers to some property other than a displacement. Then

$$\mathcal{H}(\vec{r}) = \sum_\ell \pi_\ell(\vec{r})\dot{\eta}_\ell(\vec{r}) - \mathcal{L}(\vec{r}).$$

where

$$\pi_\ell(\vec{r}) = \frac{\partial\mathcal{L}}{\partial\dot{\eta}_\ell(\vec{r})} = \frac{1}{c}\frac{\partial\mathcal{L}}{\partial\eta_{\ell,0}(\vec{r})}.$$

Notice from (8.3) that ${T_\mu}^0 = c\sum_\ell \pi_\ell\eta_{\ell,\mu} - {\delta_\mu}^0\mathcal{L}$, and in particular ${T_0}^0 = \sum_\ell \pi_\ell\dot{\eta}_\ell - \mathcal{L} = \mathcal{H}$ is the Hamiltonian density, which we see is one component of the stress-energy tensor.

Consider again the case where $\mathcal{L}$ does not depend explicitly on $(\vec{x},t)$, so $\sum_{\nu=0}^3 \partial_\nu {T_\mu}^\nu = 0$, which, as we have seen, tells us that the four currents $(J_\mu)^\nu := {T_\mu}^\nu$ are conserved currents, leading to conserved "charges" $Q_\mu = \int_V {T_\mu}^0\, d^3x$. For $\mu = 0$, ${T_0}^0$ is the hamiltonian density, so under appropriate conditions Q_0 is the conserved total energy. Then ${T_0}^j$ should be the j component of the flow of energy. As an example, let's return to thinking of η_i as the displacement, and make the small deviation approximation of section 5.4.2. If we consider a small piece $d\vec{S}$ of the surface of a volume V, then the inside is exerting a force $dF_i = \sum_j \mathbf{P}_{ij}dS_j$ on the outside, and if the surface is moving with velocity $\vec{v}$, the inside is doing work $\sum_i v_i dF_i = \vec{v}\cdot\mathbf{P}\cdot d\vec{S}$. But $\vec{v} = d\vec{\eta}/dt$ or $v_i = c\eta_{i,0}$, so energy is flowing out of the volume at a rate

$$\begin{aligned}
-\frac{dE}{dt} &= c\int_S \vec{\eta}_{,0}\cdot\mathbf{P}\cdot d\vec{S} = c\int_V \sum_{ij}\partial_j\left(\eta_{i,0}\mathbf{P}_{ij}\right)) \\
&= c\int_V \sum_j \partial_j {T_0}^j = c\int_V \sum_{ij}\partial_j\left(\frac{\partial\mathcal{L}}{\partial\eta_{i,j}}\eta_{i,0}\right)
\end{aligned}$$

which encourages us to conclude

$$\mathbf{P}_{ij} = \frac{\partial\mathcal{L}}{\partial\eta_{i,j}}.$$

A force on the surface of our volume transfers not only energy but also momentum. In fact, the force A exerts on B represents the rate of momentum transfer from A to B, and the force per unit area across a surface gives the flux of momentum across that surface. As the outside is exerting a force $-dF_i = -\sum_j \mathbf{P}_{ij}dS_j$ on the inside, this force will cause the momentum P_i of the inside of the volume to be changing at a rate

$$\begin{aligned}
\frac{d}{dt}P_i &= \int_S -\sum_j \mathbf{P}_{ij}dS_j = -\int_V \sum_j \partial_j\mathbf{P}_{ij} = -\int_V \sum_j \partial_j\frac{\partial\mathcal{L}}{\partial\eta_{i,j}} \\
&= \int_V \left(\frac{d}{cdt}\frac{\partial\mathcal{L}}{\partial\eta_{i,0}}\right) - \frac{\partial\mathcal{L}}{\partial\eta_i},
\end{aligned}$$

where in the last step we used the equations of motion. If it were not for the last term, we would take this as expected, because we would expect, if the Lagrangian is of the usual form, that the momentum density would be $\frac{\partial\mathcal{L}}{\partial\dot{\eta}_i} = \frac{\partial\mathcal{L}}{\partial c\eta_{i,0}}$. We will return to the interpretation of this last term after we discuss what happens in its absence.

Cyclic coordinates

In discrete mechanics, when L was independent of a coordinate q_i, even though it depended on $\dot{q}_i$, we called the coordinate cyclic or ignorable, and found a conserved momentum conjugate to it. In particular, if we use the center-of-mass coordinates in an isolated system those will be ignorable coordinates and the conserved momentum of the system will be their conjugate variables. In field theory, however, the center of mass is not a suitable dynamical variable. The variables are not $\vec{x}$ but $\eta_i(\vec{x},t)$. For fields in general, $\mathcal{L}(\eta,\dot\eta,\nabla\eta)$ depends on spatial derivatives of η as well, and we may ask whether we need to require absence of dependence on $\nabla\eta$ for a coordinate to be cyclic. Independence of both η and $\nabla\eta$ implies independence on an infinite number of discrete coordinates, the values of $\eta(\vec{r})$ at every point $\vec{r}$, which is too restrictive a condition for our discussion. We will call a coordinate field η_i cyclic if $\mathcal{L}$ does not depend directly on η_i, although it may depend on its derivatives $\dot\eta_i$ and $\nabla\eta_i$.

The Lagrange equation then states

$$\sum_\mu \partial_\mu \frac{\partial\mathcal{L}}{\partial\eta_{i,\mu}} = 0, \quad \text{or } \frac{d}{dt}\pi_i + \sum_j \partial_j \frac{\partial\mathcal{L}}{\partial\eta_{i,j}} = 0,$$

which constitutes continuity equations for the densities $\pi_i(\vec{r},t)$ and currents $(\vec{j}_i)_\ell = \partial\mathcal{L}/\partial\eta_{i,j}$. If we integrate this equation over all space, and define

$$\Pi_i(t) = \int \pi_i(\vec{r})d^3r,$$

then the derivative $d\Pi/dt$ involves the integral of a divergence, which by Gauss' law is a surface term

$$\frac{d\Pi(t)}{dt} = -\int \frac{\partial\mathcal{L}}{\partial\eta_{i,j}}(dS)_j.$$

If we assume the spatial boundary conditions are such that we may ignore this boundary term, we see that the $\Pi_i(t)$ will be constants of the motion. These are the total canonical momentum conjugate to η, and not, except when η represents a displacement, the components of the total ordinary momentum of the system.

If we considered our continuum with η_i representing the displacement, and placed it in a gravitational field, we would have an additional potential energy $\int_V \rho g\eta_3$, and our equation for $d\pi_i/dt$ would have an extra term corresponding to the volume force:

$$F_i^{\rm vol} + F_i^{\rm surf} = \Delta V \frac{d\pi_i}{dt} = \Delta V\left(-\sum_j \partial_j \frac{\partial\mathcal{L}}{\partial\eta_{i,j}} + \frac{\partial\mathcal{L}}{\partial\eta}_i\right),$$

so

$$F_i^{\rm vol} = \Delta V \frac{\partial\mathcal{L}}{\partial\eta}_i = -\rho g\hat{e}_z\Delta V,$$

as expected, and the total momentum is not conserved.

From equation (8.3) we found that if $\mathcal{L}$ is independent of $\vec{x}$, the stress-energy tensor gives conserved currents. Linear momentum conservation in field dynamics is connected not to ignorable coordinates but to a lack of dependence on the labels. This is best viewed as an invariance under a transformation of all the fields, $\eta_i(\vec{x}) \to \eta_i(\vec{x}+\vec{a})$, for a constant vector $\vec{a}$. This is a change in the integrand which can be undone by a change in the variable of integration, $\vec{x} \to \vec{x}\,' = \vec{x}+\vec{a}$, under which the Lagrangian is unchanged if the integration is over all space and the Lagrangian density does not depend explicitly on $\vec{x}$. This is a special case of conserved quantities arising because of symmetries, a topic we will pursue in the next section.

Before we do so, let us return to our treatment of elasticity in the linear continuum approximation of a solid, with the dynamical fields being the displacements $\eta_i(\vec{x},t)$. We saw that the stress tensor $\mathbf{P}_{ij} = \partial\mathcal{L}/\partial\eta_{i,j}$, and if we intend to describe a material obeying the generalized Hooke's law,

$$\frac{\partial\mathcal{L}}{\partial\eta_{i,j}} = \mathbf{P}_{ij} = -\frac{\alpha-\beta}{3}\delta_{ij}\,\mathrm{Tr}\,\mathbf{S} - \beta\mathbf{S}_{ij} = -\frac{\alpha-\beta}{3}\delta_{ij}\sum_k \eta_{k,k} - \frac{\beta}{2}\left(\eta_{i,j}+\eta_{j,i}\right).$$

This suggests a term in the Lagrangian

$$\mathcal{L}_1 = \frac{\beta-\alpha}{6}\left(\sum_k \eta_{k,k}\right)^2 - \frac{\beta}{8}\left(\eta_{i,j}+\eta_{j,i}\right)^2.$$

We will also need a kinetic energy term to give a momentum density, which we would expect to be just $\vec{\pi} = \rho\dot{\vec{\eta}}$, so we take that term to be

$$\mathcal{L}_2 = \frac{c^2}{2}\rho\sum_i \eta_{i,0}^2.$$

Finally, if we have a volume force $\vec{E}(\vec{r})$ due to some external potential $-\vec{\eta}\cdot\vec{E}$, this will be from $\mathcal{L}_3 = \vec{\eta}\cdot\vec{E}$. Thus our total lagrangian density is

$$\mathcal{L} = \frac{c^2}{2}\rho\sum_i \eta_{i,0}^2 + \frac{\beta-\alpha}{6}\left(\sum_k \eta_{k,k}\right)^2 - \frac{\beta}{8}\left(\eta_{i,j}+\eta_{j,i}\right)^2 + \vec{\eta}\cdot\vec{E}.$$

Now

$$\frac{\partial\mathcal{L}}{\partial\eta_{i,j}} \quad = \quad \frac{\beta-\alpha}{3}\delta_{ij}\sum_k \eta_{k,k} - \frac{\beta}{2}\left(\eta_{i,j}+\eta_{j,i}\right)$$

$$\frac{\partial \mathcal{L}}{\partial \eta_{i,0}} = c^2 \rho \eta_{i,0}$$
$$\frac{\partial \mathcal{L}}{\partial \eta_i} = E_i$$

so the equations of motion are

$$0 = \partial_\mu \frac{\partial \mathcal{L}}{\partial \eta_{i,\mu}} - \frac{\partial \mathcal{L}}{\partial \eta_i} = \rho \ddot{\eta}_i + \sum_j \left[\frac{\beta - \alpha}{3} \delta_{ij} \sum_k \eta_{k,k,j} - \frac{\beta}{2} \left(\eta_{i,j,j} + \eta_{j,i,j} \right) \right] - E_i,$$

or

$$\rho \ddot{\vec{\eta}} = \left(\frac{\alpha}{3} + \frac{\beta}{6} \right) \nabla (\nabla \cdot \vec{\eta}) + \frac{\beta}{2} \nabla^2 \vec{\eta} + \vec{E},$$

in agreement with (5.9).

8.2 Special Relativity

We have commented several times that a continuous symmetry of the dynamics, such as invariance under translation or rotation, is reflected in conservation laws. We will give a formal development of Noether's theorem, which makes this connection generally, in the next section. When we do that, we will certainly want to consider relativistic invariance, so first we will revise and clarify our notation appropriately.

So we now consider the symmetry known as special relativity, the postulate that the laws of physics are equally valid in all inertial reference frames. We will assume familiarity with the basic ideas[4], so we will only deal with notational issues here. The relation of coordinates in different inertial reference frames is determined by the invariance of

$$(ds)^2 = -c^2 (dt)^2 + (dx)^2 + (dy)^2 + (dz)^2,$$

where c is the speed of light in vacuum. This looks something like the Pythagorian length, except that the time component is scaled and has the wrong sign. The scaling is not a problem, we could just choose to define $x^0 = ct$ and measure time with x^0 in meters. Then we can treat the space-time coordinates as a four-vector[5] $x^\mu = (ct, x, y, z)$. The minus sign is more

[4]The student who has not learned about Einstein's theory is referred to Smith ([15]) or French ([5]) for elementary introductions.

[5]Actually x^μ is a position in space-time and not truly a vector, a distinction discussed in section (1.2.1) but not important here.

significant, so that $(ds)^2$ is not a true length. We introduce the Minkowski metric tensor

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

so we can write[6]

$$(ds)^2 = \sum_{\mu\nu} \eta_{\mu\nu} dx^\mu dx^\nu.$$

Notice we have defined x^μ with superscripts rather than subscripts, and any vector (or tensor) with such indices is said to be **contravariant**. From any such vector V^μ we can also define a **covariant** vector

$$V_\mu = \sum_\nu \eta_{\mu\nu} V^\nu.$$

This is a somewhat trivial distinction in special relativity, only changing the sign of the zeroth component[7]. But it is useful, because it enables us to define an invariant inner product $\sum_\mu V_\mu W^\mu$. One can also make a contravariant vector from a covariant one, $W^\mu = \sum_\nu \eta^{\mu\nu} W_\nu$, where $\eta^{\mu\nu}$ is the inverse[8], as a matrix, of $\eta_{\mu\nu}$. We will also redefine the Einstein summation convention: an index which occurs twice is summed over only if it appears once upper and once lower. (Otherwise it is probably a mistake!) We also redefine what we mean by the square of a vector V^μ: $V^2 := \eta_{\mu\nu} V^\mu V^\nu = V_\mu V^\mu$ and not $\sum_\mu (V^\mu)^2$.

The relationship between coordinates in different inertial frames,

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

is given by the Lorentz transformation matrix $\Lambda^\mu{}_\nu$. The invariance of $(ds)^2$ tells us

$$\eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}, \tag{8.5}$$

---

[6] Note that this is not a two-form, as η is symmetric.

[7] In general relativity $\eta_{\mu\nu}$ is replaced by the metric tensor $g_{\mu\nu}$ which is a dynamical degree of freedom of space-time rather than a fixed matrix, and this distinction becomes less trivial.

[8] $\sum_\rho \eta^{\mu\rho}\eta_{\rho\nu} = \delta^\mu{}_\nu = 1$ if $\mu = \nu$ and 0 otherwise. Note the Kronecker delta function needs one upper and one lower index in order to be properly covariant, and in fact it and η are different forms of the same tensor, using the usual lowering or raising procedures with η. Don't be misled by the fact that for each μ and ν, $\eta^{\mu\nu}$ is the same as $\eta_{\mu\nu}$.

which says that Λ is pseudo-orthogonal.

We have defined position to be naturally described by a contravariant vector, but some objects are naturally defined as covariant. In particular, the partial derivative operator

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \text{ is, for } \partial_\mu x^\nu = \delta^\nu_\mu.$$

With this four dimensional notation we see that time translation and spatial translations are unified in $x^\mu \to x^\mu + c^\mu$, and rotations are just special cases of Lorentz transformations, with

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R & \\ 0 & & & \end{pmatrix}.$$

As for rotations, we may ask how objects transform under Lorentz transformations. For rotations, we saw that in addition to scalars and vectors, we may have tensors with multiple indices. The same is true in relativity — a large class of covariant objects may be written in terms of multiple indices, and the transformation properties are simply multiplicative. First of all, how does a covariant vector transform? From $V'^\mu = \Lambda^\mu{}_\nu V^\nu$ and the lowered forms $V'_\rho = \eta_{\rho\mu} V'^\mu = \eta_{\rho\mu}\Lambda^\mu{}_\nu V^\nu = \eta_{\rho\mu}\Lambda^\mu{}_\nu \eta^{\nu\sigma} V_\sigma$, we see that $V'_\rho = \Lambda_\rho{}^\sigma V_\sigma$, where we have used η's to lower and raise the indices on the Lorentz matrix, $\Lambda_\rho{}^\sigma = \eta_{\rho\mu}\Lambda^\mu{}_\nu \eta^{\nu\sigma}$. So we see that covariant indices transform with $\Lambda_\rho{}^\sigma$. Note that $\Lambda_\rho{}^\sigma \Lambda^\rho{}_\tau = \eta_{\rho\mu}\Lambda^\mu{}_\nu \eta^{\nu\sigma}\Lambda^\rho{}_\tau = \eta_{\tau\nu}\eta^{\nu\sigma} = \delta^\sigma_\tau$, where the second equality follows from (8.5), so $\Lambda_\rho{}^\sigma = (\Lambda^{-1})^\sigma{}_\rho$. Note also that the order of indices matters, $\Lambda_\mu{}^\nu \neq \Lambda^\nu{}_\mu$.

Now more generally we may define a multiply-indexed tensor $T^{\mu_1\ldots\mu_j}{}_{\nu_1\ldots\nu_k}{}^{\mu_{j+1}\ldots\mu_\ell}$ and it will transform with each index suitably transformed:

$$T'^{\mu'_1\ldots\mu'_j}{}_{\nu'_1\ldots\nu'_k}{}^{\mu'_{j+1}\ldots\mu'_\ell} = \prod_{i=1}^{\ell} \Lambda^{\mu'_i}{}_{\mu_i} \prod_{n=1}^{k} \Lambda_{\nu'_n}{}^{\nu_n} T^{\mu_1\ldots\mu_j}{}_{\nu_1\ldots\nu_k}{}^{\mu_{j+1}\ldots\mu_\ell}. \tag{8.6}$$

If we contract two indices, they don't contribute to the transformation:

$$T'_\mu{}^\mu = \Lambda_\mu{}^\nu \Lambda^\mu{}_\rho T_\nu{}^\rho = (\Lambda^{-1})^\nu{}_\mu \Lambda^\mu{}_\rho T_\nu{}^\rho = \delta^\nu_\rho T_\nu{}^\rho = T_\nu{}^\nu.$$

So we see that we can make an invariant object (a scalar) by contracting all indices. We should mention that in addition to tensors, another possible transformation possibility is that of a spinor, but we will not explore that here.

For a point particle, the momentum three-vector is coupled by Lorentz transformation to the energy[9], $P^\mu = (E/c, \vec{p})$. Then we see that to make an invariant,

$$P^\mu P^\nu \eta_{\mu\nu} = \vec{p}^2 - E^2/c^2 = -m^2c^2.$$

We are going to be interested in infinitesimal Lorentz transformations, with $\Lambda^\mu{}_\nu = \delta^\mu_\nu + \epsilon L^\mu{}_\nu$. From the condition (8.5) for Λ to be a Lorentz transformation, we have

$$\eta_{\mu\nu}\left(\delta^\mu_\rho + \epsilon L^\mu{}_\rho\right)\left(\delta^\nu_\sigma + \epsilon L^\nu{}_\sigma\right) = \eta_{\rho\sigma} + \epsilon\left(\eta_{\mu\sigma}L^\mu{}_\rho + \eta_{\rho\nu}L^\nu{}_\sigma\right) + \mathcal{O}(\epsilon^2) = \eta_{\rho\sigma},$$

so

$$\eta_{\mu\sigma}L^\mu{}_\rho + \eta_{\rho\nu}L^\nu{}_\sigma = L_{\sigma\rho} + L_{\rho\sigma} = 0,$$

so the condition is that L is antisymmetric when its indices are both lowered. Thus $L_{\cdot\cdot}$ is a 4×4 antisymmetric real matrix, and has 6 independent parameters, and the infinitesimal Lorentz transformations form a 6 dimensional Lie algebra.

Now we are ready to discuss symmetries more generally.

8.3 Noether's Theorem

We have seen in several ways that there is a connection between conserved quantities and an invariance of the dynamics under some continuous transformations. First we saw, in discrete dynamics, that ignorable coordinates have conserved conjugate momenta. A coordinate is ignorable if the Lagrangian is unchanged under its translation, $\phi \to \phi + c$, for arbitrary c. In particular invariance under translation of all coordinates $\vec{r}_j \to \vec{r}_j + \vec{c}$ leads to conservation of the total momentum. In field theory momentum conservation is not associated with ignorable field coordinates, but rather to invariance under translations of the labels, that is, under $\eta_\ell(\vec{r}) \to \eta_\ell(\vec{r} + \vec{c})$, which is a consequence of $\vec{r}$ being an integration variable, so changing it makes no difference as long as $\mathcal{L}$ has no explicit dependence on it, and as long as we

[9]Why P^0 rather than P_0 for the energy? In quantum mechanics we have $\vec{p}$ associated with the gradient operator, $\vec{p} = -i\hbar\vec{\nabla}$, and a partial derivative is naturally covariant. But the energy is $H = i\hbar\partial/\partial t$, because Schrödinger arbitrarily chose that sign when he wrote down his equation. So if we write $P_\mu = -i\hbar\partial/\partial x^\mu$, we have $P_\mu = (-E/c, \vec{p})$.

are integrating over all $\vec{r}$. For rotations in discrete mechanics we saw that one component of $\vec{L}$ could be considered conserved because ϕ is ignorable, but the other two components, which are also conserved, must be attributed to a less obvious symmetry, that of rotations about directions other than z.

Now we will discuss more generally the relationship between symmetries and conserved quantities, a general connection given in a famous theorem by **Emmy Noether**[10]. Symmetry means the dynamics is unchanged under a change in the values of the degrees of freedom $\eta \to \eta'$ which will in general depend on those degrees of freedom. In the discrete case the dependence is commonly on a related set, such as the new x component of the electric field experienced by a point charge being dependent on all three old components under a general rotation. In the case of fields, it would in principle be possible for the new field $\eta'_\ell(\vec{x})$ to depend on all the values of all fields at all points in space, but this is not useful to consider. We might consider only local symmetries, for which it depends only on the old fields at the same point, $\eta'_k(\vec{x})$, which might for example be the case for considering the spins of atoms under rotation of all the spins. But if we want to consider the more fundamental symmetry under a true rotation, for which the atoms are also rotating, we need to consider a symmetry which relates new fields at x' to old fields at x, where the symmetry maps $x \to x'$ as well as transforming the fields. Then we find that the new field $\eta'_\ell(\vec{x}\,')$ depends on the old fields at a different point $\vec{x}$. This is what we have in the case of translation we just discussed, as well as for rotations and other possible symmetries. These symmetries may be thought of in a passive sense as having the physics unchanged when we translate, rotate, or boost (in a relativistic theory) the coordinate system describing the physics. Then the new coordinates x'_μ describe the same physical point as the old x_μ, with a definite map $\Phi : x \mapsto x'$ which describes the change of coordinates of space(time). While the physics at that point is unchanged, its description in terms of fields may be, so we need to consider a rule for transforming the fields, which gives $\eta_\ell(x'_\mu)$ as a function of fields at x_ν.

We will only be concerned with continuous symmetries, which can be generated by infinitesimal transformations, so we can consider an infinitesimal transformation with $x'_\mu = x_\mu + \delta x_\mu$, along with a rule that gives the change of $\eta'_\ell(x'_\mu)$ from the set of $\eta_k(x_\nu)$. For a scalar field, like temperature, under a

[10]This section relies heavily on Goldstein, "Classical Mechanics", 2nd Ed., section 12-7.

rotation, we would define the new field

$$\eta'(x') = \eta(x),$$

but more generally the field may also change, in a way that may depend on other fields,

$$\eta'_i(x') = \eta_i(x) + \delta\eta_i(x; \eta_k(x)).$$

This is what you would expect for a vector field $\vec{E}$ under rotations, because the new E'_x gets a component from the old E_y.

To say that

$$x_\mu \to x'_\mu, \qquad \eta_i \to \eta'_i$$

is a symmetry means, at the least, that if $\eta_i(x)$ is a specific solution of the equations of motion, the set of transformed fields $\eta'_i(x')$ is also a solution. The equations of motion are determined by varying the action, so if the corresponding actions are equal for each pair of configurations $(\eta(x), \eta'(x'))$, so are the equations of motion. Notice here that what we are saying is that the same Lagrangian function applied to the fields η'_i and integrated over $x' \in \mathcal{R}'$ should give the same action as $S = \int_{\mathcal{R}} \mathcal{L}(\eta_i(x)...)d^4x$, where $\mathcal{R}'$ is the range of x' corresponding to the domain $\mathcal{R}$ of x. [Of course our argument applies also if $\delta x_\mu = 0$, when the transformation does not involve a change in coordinates. Such symmetries are called *internal symmetries*, with isospin an example.]

Actually, the above condition that the actions be unchanged is far more demanding than is needed to insure that the same equations of motion arise. The variations required to derive the equations of motion only compare actions for field configurations unchanged at the boundaries, so if the actions

$$S' = \int_{\mathcal{R}'} \mathcal{L}(\eta'_i(x'), \partial'_\mu\eta'_i(x'), x')d^4x' \text{ and } S = \int_{\mathcal{R}} \mathcal{L}(\eta_i(x), \partial_\mu\eta_i(x), x)d^4x \quad (8.7)$$

differ by a function only of the values of η_i on the boundary $\partial\mathcal{R}$, they will give the same equations of motion. Even in quantum mechanics, where the transition amplitude is given by integrating $e^{iS/\hbar}$ over all configurations, a change in the action which depends only on surface values is only a phase change in the amplitude. In classical mechanics we could also have an overall change multiplying the Lagrangian and the action by a constant $c \neq 0$, which would still have extrema for the same values of the fields, but we will not consider such changes because quantum mechanically they correspond to changing Planck's constant.

The Lagrangian density is a given function of the old fields $\mathcal{L}(\eta_i, \partial_\mu \eta_i, x_\mu)$. If we substitute in the values of $\eta(x)$ in terms of $\eta'(x')$ we get a new density $\mathcal{L}'$, defined by

$$\mathcal{L}'(\eta'_i, \partial'_\mu \eta'_i, x'_\mu) = \mathcal{L}(\eta_i, \partial_\mu \eta_i, x_\mu) \left| \frac{\partial x^\nu}{\partial x'^\mu} \right|,$$

where the last factor is the Jacobian of the transformation $x \to x'$, required because these are densities, intended to be integrated. This change in functional form for the Lagrangian is not the symmetry transformation, for as long as $x \leftrightarrow x'$ is one-to-one, the integral is unchanged

$$\begin{aligned} \int_{\mathcal{R}'} \mathcal{L}'(\eta'_i(x'), \partial'_\mu \eta'_i(x'), x') d^4x' &= \int_{\mathcal{R}'} \mathcal{L}(\eta_i(x), \partial_\mu \eta_i(x), x) \left| \frac{\partial x^\nu}{\partial x'^\mu} \right| d^4x' \\ &= \int_{\mathcal{R}} \mathcal{L}(\eta_i(x), \partial_\mu \eta_i(x), x) d^4x = S \qquad (8.8) \end{aligned}$$

regardless of whether this transformation is a symmetry.

We see that the change in the action, $\delta S = S' - S$, which must vanish up to surface terms for a symmetry, may be written as an integral over $\mathcal{R}'$ of the variation of the Lagrangian density, $\delta S = \int_{\mathcal{R}'} \delta \mathcal{L}$, with

$$\begin{aligned} \delta\mathcal{L}(\eta'_i(x'), \partial'_\mu \eta'_i(x'), x') &:= \mathcal{L}(\eta'_i(x'), \partial'_\mu \eta'_i(x'), x') - \mathcal{L}'(\eta'_i(x'), \partial'_\mu \eta'_i(x'), x') \\ &= \mathcal{L}(\eta'_i(x'), \partial'_\mu \eta'_i(x'), x') - \mathcal{L}(\eta_i(x), \partial_\mu \eta_i(x), x) \left| \frac{\partial x^\nu}{\partial x'^\mu} \right|. \qquad (8.9) \end{aligned}$$

Here we have used the first of Eq. (8.7) for S' and Eq. (8.8) for S.

Expanding to first order, the Jacobian is

$$\det \left| \frac{\partial x'^\mu}{\partial x^\nu} \right|^{-1} = \det \left(\delta^\mu_\nu + \partial_\nu \delta x^\mu \right)^{-1} = \left(1 + \operatorname{Tr} \frac{\partial \delta x^\mu}{\partial x^\nu} \right)^{-1} = 1 - \partial_\mu \delta x^\mu, \quad (8.10)$$

while

$$\begin{aligned} \mathcal{L}(\eta'_i(x'), \partial'_\mu \eta'_i(x'), x') &= \mathcal{L}(\eta_i(x), \partial_\mu \eta_i(x), x) \\ &\quad + \delta\eta_i \frac{\partial \mathcal{L}}{\partial \eta_i} + \delta(\partial_\mu \eta_i) \frac{\partial \mathcal{L}}{\partial \partial_\mu \eta_i} + \delta x^\mu \frac{\delta \mathcal{L}}{\delta x^\mu}, \quad (8.11) \end{aligned}$$

Thus[11]

$$\delta\mathcal{L} = \mathcal{L}\partial_\mu \delta x^\mu + \delta\eta_i \frac{\partial \mathcal{L}}{\partial \eta_i} + \delta(\partial_\mu \eta_i)\frac{\partial \mathcal{L}}{\partial\partial_\mu \eta_i} + \delta x^\mu \frac{\delta \mathcal{L}}{\delta x^\mu}, \tag{8.12}$$

and if this is a divergence, $\delta\mathcal{L} = \partial_\mu \Lambda^\mu$ for some Λ^μ, we will have a symmetry.

There are subtleties in this expression[12]. The last term involves a derivative of $\mathcal{L}$ with its first two arguments fixed, and as such is not the derivative with respect to x^μ with the *functions* η_i fixed. For this reason we used a different symbol, because it is customary to use ∂_μ to mean only that x^ν is fixed for $\nu \neq \mu$, and not to indicate that the other arguments of $\mathcal{L}$ are held fixed. That form of derivative is the stream derivative,

$$\frac{\partial \mathcal{L}\Big(\eta_i(x), \partial_\mu \eta_i(x), x\Big)}{\partial x^\nu} = \frac{\delta \mathcal{L}\Big(\eta_i(x), \partial_\mu \eta_i(x), x\Big)}{\delta x^\nu} + (\partial_\nu \eta_i)\frac{\partial \mathcal{L}}{\partial \eta_i} + (\partial_\nu \partial_\mu \eta_i)\frac{\partial \mathcal{L}}{\partial\partial_\mu \eta_i}.$$

Note also that $\delta\eta_i(x) = \eta_i'(x') - \eta_i(x)$ is not simply the variation of the field at a point, $\mathfrak{d}\eta_i(x) = \eta_i'(x) - \eta_i(x)$, but includes in addition the change $(\delta x^\mu)\partial_\mu \eta_i$ due to the displacement of the argument. Thus

$$\delta\eta_i(x) = \mathfrak{d}\eta_i(x) + (\delta x^\nu)\partial_\nu \eta_i. \tag{8.13}$$

The variation with respect to $\partial_\mu' \eta_i'$ needs to be examined carefully, because the δ variation effects the coordinates, and therefore in general $\partial_\mu \delta\eta_i \neq \delta\partial_\mu \eta_i$. By definition,

$$\begin{aligned}
\delta\partial_\mu \eta_i &= \partial\eta_i'/\partial x'^\mu\big|_{x'} - \partial\eta_i/\partial x^\mu\big|_x \\
&= \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu}\left[\eta_i + (\delta x^\rho)\partial_\rho \eta_i + \mathfrak{d}\eta_i\right]\Big|_x - \partial\eta_i/\partial x^\mu\big|_x \\
&= -\left(\partial_\mu \delta x^\nu\right)\partial_\nu \eta_i + \frac{\partial}{\partial x^\mu}\left[(\delta x^\rho)\partial_\rho \eta_i + \mathfrak{d}\eta_i\right] \\
&= (\delta x^\nu)\partial_\mu \partial_\nu \eta_i + \mathfrak{d}\partial_\mu \eta_i
\end{aligned} \tag{8.14}$$

where in the last line we used $\partial_\mu \mathfrak{d}\eta_i = \mathfrak{d}\partial_\mu \eta_i$, because the $\mathfrak{d}$ variation *is* defined at a given point and *does* commute with ∂_μ.

---

[11] This is the equation to use on homework.

[12] There is also a summation understood on the repeated i index as well as on the repeated μ index.

Notice that the δx^ν terms in (8.13) and (8.14) are precisely what is required in (8.11) to change the last term to a full stream derivative. Thus

$$\begin{aligned}\mathcal{L}(\eta_i'(x'), \partial_\mu'\eta_i'(x'), x') &= \mathcal{L}(\eta_i(x), \partial_\mu\eta_i(x), x) \\ &\quad + \mathfrak{d}\eta_i \frac{\partial \mathcal{L}}{\partial \eta_i} + \mathfrak{d}\partial_\mu\eta_i \frac{\partial \mathcal{L}}{\partial \partial_\mu \eta_i} + \delta x^\mu \frac{\partial \mathcal{L}}{\partial x^\mu}, \qquad (8.15)\end{aligned}$$

where now $\partial\mathcal{L}/\partial x^\mu$ means the stream derivative, including the variations of $\eta_i(x)$ and its derivative due to the variation δx^μ in their arguments.

Inserting this and (8.10) into the expression (8.9) for $\delta\mathcal{L}$, we see that the change of action is given by the integral of

$$\begin{aligned}\delta\mathcal{L} &= (\partial_\mu \delta x^\mu)\,\mathcal{L} + \delta x^\mu \frac{\partial \mathcal{L}}{\partial x^\mu} + \mathfrak{d}\eta_i \frac{\partial \mathcal{L}}{\partial \eta_i} + \mathfrak{d}\partial_\mu\eta_i \frac{\partial \mathcal{L}}{\partial \partial_\mu \eta_i} \\ &= \frac{\partial}{\partial x^\mu}\left(\delta x^\mu \mathcal{L} + \mathfrak{d}\eta_i \frac{\partial \mathcal{L}}{\partial \partial_\mu \eta_i}\right) + \mathfrak{d}\eta_i \left(\frac{\partial \mathcal{L}}{\partial \eta_i} - \frac{\partial}{\partial x^\mu}\frac{\partial \mathcal{L}}{\partial \partial_\mu \eta_i}\right) \qquad (8.16)\end{aligned}$$

We will discuss the significance of this in a minute, but first, I want to present an alternate derivation.

Observe that in the expression (8.7) for S', x' is a dummy variable and can be replaced by x, and the difference can be taken at the same x values, except that the ranges of integration differ. That is,

$$S' = \int_{\mathcal{R}'} \mathcal{L}\left(\eta'(x), \partial_\mu \eta'(x), x\right) d^4x,$$

and this differs from $S(\eta)$ because

1. the Lagrangian is evaluated with the field $\eta'(x)$ rather than $\eta(x)$, producing a change

   $$\delta_1 S = \int \left(\frac{\partial \mathcal{L}}{\partial \eta_i}\mathfrak{d}\eta_i + \frac{\partial \mathcal{L}}{\partial \partial_\mu \eta_i}\mathfrak{d}\partial_\mu\eta_i\right) d^4x,$$

   where the variation with respect to the fields is now in terms of $\mathfrak{d}\eta_i(x) := \eta_i'(x) - \eta_i(x)$, at the same argument x.

2. Change in the region of integration, $\mathcal{R}'$ rather than $\mathcal{R}$,

   $$\delta_2 S = \left(\int_{\mathcal{R}'} - \int_{\mathcal{R}}\right) \mathcal{L}(\eta_i, \partial_\mu\eta_i, x)\, d^4x.$$

If we define dS_μ to be an element of the three dimensional boundary $\partial\mathcal{R}$ of $\mathcal{R}$, with outward-pointing normal in the direction of dS_μ, the difference in the regions of integration may be written as an integral over the surface,

$$\left(\int_{\mathcal{R}'} - \int_{\mathcal{R}}\right) d^4x = \int_{\partial\mathcal{R}} \delta x^\mu \cdot dS_\mu.$$

Thus

$$\delta_2 S = \int_{\partial\mathcal{R}} \mathcal{L}\delta x^\mu \cdot dS_\mu = \int_{\mathcal{R}} \partial_\mu \left(\mathcal{L}\delta x^\mu\right) \tag{8.17}$$

by Gauss' Law (in four dimensions).

As $\mathfrak{d}$ is a difference of two functions at the same values of x, this operator commutes with partial differentiation, so $\mathfrak{d}\partial_\mu\eta_i = \partial_\mu\mathfrak{d}\eta_i$. Using this in the second term of $\delta_1 S$, and using $A\partial_\mu B = \partial_\mu(AB) - B\partial_\mu A$, we have

$$\delta_1 S \;=\; \int_{\mathcal{R}} d^4x \left[\partial_\mu\left(\mathfrak{d}\eta_i \frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\right) + \mathfrak{d}\eta_i\left(\frac{\partial\mathcal{L}}{\partial\eta_i} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\right)\right]$$

Thus altogether $S' - S = \delta_1 S + \delta_2 S = \int_{\mathcal{R}} d^4x\delta\mathcal{L}$, with $\delta\mathcal{L}$ given by (8.16). This completes our alternate derivation that $S' - S = \int_{\mathcal{R}} d^4x\delta\mathcal{L}$, and Eq. (8.16).

Note that $\delta\mathcal{L}$ is a divergence plus a piece which vanishes if the dynamical fields obey the equation of motion, quite independent of whether or not the infinitesimal variation we are considering is a symmetry. As we mentioned, to be a symmetry, $\delta\mathcal{L}$ must be a divergence for all field configurations, not just those satisfying the equations of motion, so that the *variations over configurations* will give the correct equations of motion.

We have been assuming the variations δx and $\delta\eta$ can be treated as infinitesimals. This is appropriate for a continuous symmetry, that is, a symmetry group[13] described by a (or several) continuous parameters. For example, symmetry under displacements $x^\mu \to x^\mu + c^\mu$, where c^μ is any arbitrary fixed 4-vector, or rotations through an arbitrary angle θ about a fixed axis. Each element of such a group lies in a one-parameter subgroup, and can be obtained, in the limit, from an infinite number of applications of an infinitesimal transformation. If we call the parameter ϵ, the infinitesimal variations in x^μ

[13]Symmetries always form a group. Continuous symmetries form a *Lie group*, whose elements can be considered exponentials of linear combinations of generators. The generators form a *Lie algebra*.

and η_i are given by derivatives of $x'(\epsilon, x)$ and η' with respect to the parameter ϵ. Thus

$$\delta x^\mu = \epsilon \left. \frac{dx'^\mu}{d\epsilon} \right|_{x^\nu}, \qquad \delta\eta_i = \epsilon \left. \frac{d\eta_i'(x')}{d\epsilon} \right|_{x^\nu}.$$

The divergence must also be first order in ϵ, so $\delta\mathcal{L} = \epsilon\partial_\mu\Lambda^\mu$ if we have a symmetry.

We define the **current** for the transformation

$$J^\mu \;=\; -\frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\frac{d\eta_i'}{d\epsilon} + \frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\partial_\nu\eta_i\frac{dx'^\nu}{d\epsilon} - \mathcal{L}\frac{dx'^\mu}{d\epsilon} + \Lambda^\mu. \tag{8.18}$$

Recalling that $\mathfrak{d}\eta_i = \delta\eta_i - (\delta x^\nu)\partial_\nu\eta_i$, we can rewrite (8.16)

$$\begin{aligned}\delta\mathcal{L} \;&=\; \frac{\partial}{\partial x^\mu}\left(\delta x^\mu\mathcal{L} + \delta\eta_i\frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i} - \delta x^\nu(\partial_\nu\eta_i)\frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\right)\\ &\qquad +\mathfrak{d}\eta_i\left(\frac{\partial\mathcal{L}}{\partial\eta_i} - \frac{\partial}{\partial x^\mu}\frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\right)\end{aligned}$$

and see that

$$\begin{aligned}\epsilon\partial_\mu J^\mu \;&=\; \frac{\partial}{\partial x^\mu}\left(-\frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\delta\eta_i + \frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\partial_\nu\eta_i\delta x^\nu\right) - \frac{\partial}{\partial x^\mu}\left(\mathcal{L}\delta x^\mu\right) + \delta\mathcal{L}\\ &=\; \mathfrak{d}\eta_i\left(\frac{\partial\mathcal{L}}{\partial\eta_i} - \frac{\partial}{\partial x^\mu}\frac{\partial\mathcal{L}}{\partial\partial_\mu\eta_i}\right)\end{aligned}$$

Thus we have

$$\partial_\mu J^\mu = 0 \qquad \text{for a symmetry, when the fields obey the equations of motion.}$$

This condition is known as **current conservation**. Associated with each such current, we may define the charge enclosed in a constant volume V

$$Q_V(t) = \int_V d^3x J^0(\vec{x}, t).$$

If we evaluate the time derivative of the charge, we have

$$\begin{aligned}\frac{d}{dt}Q_V(t) \;&=\; \int_V d^3x\partial_0 J^0(\vec{x}, t) \approx -\int_V d^3x \sum_{i=1,3}\partial_i J^i(\vec{x}, t) = -\int_V d^3x\vec{\nabla}\cdot\vec{J}(\vec{x}, t)\\ &=\; -\int_{\partial V}\vec{J}\cdot d\vec{S},\end{aligned}$$

where ∂V is the boundary of the volume and $d\vec{S}$ an element of surface area. We have used the conservation of the current and Gauss' Law. If, as can usually be assumed, the current vanishes as we move infinitely far way from the region of interest, the surface integral vanishes if we take V to be all of space, and we find that the total charge is conserved, $dQ/dt = 0$, in the same sense that equations of motion are satisfied. The assumption about asymptotic behavior is not always valid, and we must consider whether we have grounds for it in particular applications. We will see later that in some circumstances there are "anomolies" when this assumption is not justified.

It should be mentioned that, because we are only considering infinitesimal transformations, it is possible to describe the symmetry without relating new fields at new points to old fields at the old points. We could simply consider whether the transformation of fields $\eta_i(x) \to \eta'_i(x) = \eta_i(x) + \mathfrak{d}\eta_i(x)$ is a symmetry, where $\mathfrak{d}\eta_i(x) = \delta\eta_i(x) - (\delta x^\nu)\partial_\nu\eta_i$ includes not only the natural variation $\delta\eta$ (that is, zero for a scalar and an orthogonal transformation for a vector), but also the derivative piece. The derivation then need not consider change of integration region, but will in general require a nonzero choice of Λ to compensate. This is not necessary in simple applications using the method described here. Another disadvantage of starting with $\mathfrak{d}$ is that it obscures the local nature of the field dependence.

8.3.1 Applications of Noether's Theorem

Now it is time to use the very powerful though abstract formalism Noether developed for continuous symmetries to ask about symmetries we expect our theories to have. At the very least, in this class, we are going to deal only with theories which are invariant under

- spatial translations, $\vec{x} \to \vec{x}\,' = \vec{x} + \vec{c}$.
- time translations, $t \to t' = t + c^0$, or in four dimensional notation, $x^0 \to x'^{\,0} = x^0 + c^0$.
- rotations, $x^i \to x'^{\,i} = \sum_j R^i{}_j x^j$, with $R^i{}_j$ an orthogonal matrix.
- Lorentz boost transformations.

where $R^i{}_j$ is an orthogonal real matrix of determinant 1. The first two of these together are four dimensional translations,

$$x^\mu \to x'^{\,\mu} = x^\mu + c^\mu, \tag{8.19}$$

and the last two (actually Lorentz transformations already include both) can be written $x^\mu \to x'^\mu = \sum_\nu \Lambda^\mu{}_\nu x^\nu = \Lambda^\mu{}_\nu x^\nu$, (using the Einstein summation convention), where the matrix Λ is a real matrix satisfying the pseudo-orthogonality condition

$$\Lambda^\mu{}_\nu \eta_{\mu\rho} \Lambda^\rho{}_\tau = \eta_{\nu\tau},$$

which is required so that the length of a four-vector is preserved, $x'^2 := x'^\mu x'_\mu = x^2$.

All together, this symmetry group is called the inhomogeneous Lorentz group, or Poincaré group.

Translation Invariance

First, let us consider the conserved quantities generated by translation invariance, for which $\delta x^\mu = c^\mu$. All fields we will deal with are invariant, or transform as scalars, under translations, so $\delta\eta_\ell = 0$. From (8.18) the conserved current is

$$J_c^\mu = c^\nu \frac{\partial \mathcal{L}}{\partial \partial_\mu \eta_\ell} \partial_\nu \eta_\ell - \mathcal{L} c^\mu = c^\nu T_\nu{}^\mu,$$

so the four conserved currents are nothing but the energy-momentum tensor whose conservation we found in (8.3) directly from the equations of motion. The conserved charges from this current are

$$P_\mu = \int_V T_\mu{}^0(\vec{x}, t) d^3x,$$

with $P_0 = H$, P_j the total momentum for $j = 1, 2, 3$.

Lorentz Transformations

Now consider an infinitesimal Lorentz transformation, with

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu = \left(\delta^\mu_\nu + \epsilon L^\mu{}_\nu\right) x^\nu, \quad \text{or } \delta x^\mu = \epsilon L^\mu{}_\nu x^\nu.$$

The pseudo-orthogonality of Λ requires

$$\eta_{\mu\nu} L^\mu{}_\rho \delta^\nu{}_\sigma + \eta_{\mu\nu} \delta^\mu{}_\rho L^\nu{}_\sigma = 0 = L_{\sigma\rho} + L_{\rho\sigma},$$

so the infinitesimal generator, when its indices are lowered, is antisymmetric. The fields may transform is many ways. A scalar field[14] will have $\xi'(x') =$

[14] Now that our fields may be developing space-time indices, we will change their name from η to ξ to avoid confusion with $\eta_{\mu\nu}$.

$\xi(x)$, with $\delta\xi = 0$, but a field ξ might transform like a contravariant vector, $\delta\xi^\mu = \epsilon L^\mu{}_\nu \xi^\nu$, or in an even more complex fashion such as a tensor or a spinor. Whatever the change in ξ_ℓ is, it will be proportional to $L^\mu{}_\nu$, so $\delta\eta_\ell = \epsilon L^\mu{}_\nu \Delta_\mu{}^\nu{}_\ell$, and the current generated is

$$J^\mu = L_{\rho\nu}\mathcal{M}^{\mu\rho\nu} = -\frac{\partial\mathcal{L}}{\partial\partial_\mu\xi_\ell} L^\rho{}_\sigma \Delta_\rho{}^\sigma{}_\ell + \frac{\partial\mathcal{L}}{\partial\partial_\mu\xi_\ell}\partial_\tau\xi_\ell L^\tau{}_\kappa x^\kappa - \mathcal{L}L^\mu{}_\nu x^\nu.$$

As $L_{\rho\nu}$ is antisymmetric under $\rho \leftrightarrow \nu$, there are six independant infinitesimal generators which can produce currents. Only the part antisymmetric under $\rho \leftrightarrow \nu$ in $\mathcal{M}^{\mu\rho\nu}$ enters in this expression, so we take $\mathcal{M}^{\mu\rho\nu}$ and $\Delta^{\rho\nu}{}_\ell$ to be antisymmetric under $\rho \leftrightarrow \nu$, and thus

$$\mathcal{M}^{\mu\rho\nu} = -\frac{\partial\mathcal{L}}{\partial\partial_\mu\xi_\ell}\Delta^{\rho\nu}{}_\ell + \frac{1}{2}\frac{\partial\mathcal{L}}{\partial\partial_\mu\xi_\ell}\left[\eta^{\rho\tau}(\partial_\tau\xi_\ell)x^\nu - \eta^{\nu\tau}(\partial_\tau\xi_\ell)x^\rho\right] - \frac{1}{2}\mathcal{L}\left(\eta^{\mu\rho}x^\nu - \eta^{\mu\nu}x^\rho\right).$$

Of course the six currents $\mathcal{M}^{\mu\rho\nu}$ are conserved only if the action is invariant, which will be the case only if the lagrangian density transforms like a scalar under lorentz transformations. This will be assured if all the vector indices of the fields are contracted correctly, one up and one down. Note that part of the current $\mathcal{M}^{\mu\rho\nu}$ is related to the energy-momentum tensor,

$$\mathcal{M}^{\mu\rho\nu} = \frac{1}{2}\left(x^\nu T^{\rho\mu} - x^\rho T^{\nu\mu}\right) - \frac{\partial\mathcal{L}}{\partial\partial_\mu\xi_\ell}\Delta^{\rho\nu}{}_\ell.$$

As $T^{\rho\mu}$ is a 4-current of the 4-momentum, we see that the first term is the 4-current of the four dimensional version of orbital angular momentum. The last term is then the contribution of the spin to the current of the total angular momentum.

8.4 Examples of Relativistic Fields

As we mentioned, Noether's theorem will generate conserved generators of Lorentz transformations if the lagrangian density transforms as a scalar under Poincaré transformations. For convenience we will take $c = 1$. The easiest example to consider is a single scalar field, with what is called the Klein-Gordon Lagrangian:

$$\mathcal{L} = \frac{1}{2}\left(-\eta^{\mu\nu}\frac{\partial\phi}{\partial x^\mu}\frac{\partial\phi}{\partial x^\nu} - m^2\phi^2\right) = \frac{1}{2}\left(\dot\phi^2 - (\vec\nabla\phi)^2 - m^2\phi^2\right).$$

The canonical momentum field is $\pi = \frac{\partial\mathcal{L}}{\partial\dot\phi} = \dot\phi$, and

$$T_\mu{}^\nu = \frac{\partial\mathcal{L}}{\partial\phi_{,\nu}}\phi_{,\mu} - \mathcal{L}\delta^\nu_\mu = -\phi^{,\nu}\phi_{,\mu} + \frac{1}{2}\delta^\nu_\mu\left(-\dot\phi^2 + (\vec\nabla\phi)^2 + m^2\phi^2\right).$$

The Hamiltonian is

$$H = \int T_0{}^0 d^3x = \frac{1}{2}\int\left[\dot\phi^2 + (\vec\nabla\phi)^2 + m^2\phi^2\right] d^3x,$$

the three-momentum is

$$(\vec P)_j = \int T_j{}^0\, d^3x = \int \dot\phi(\vec\nabla\phi)_j\, d^3x \qquad \text{or } \vec P = \int \pi\vec\nabla\phi\, d^3x.$$

The equation of motion (8.1) is

$$\left(\eta^{\mu\nu}\partial_\mu\partial_\nu - m^2\right)\phi = 0, \quad \text{or } \ddot\phi - \nabla^2\phi + m^2\phi = 0,$$

which has solutions which are waves, decomposable into plane waves $\phi(\vec x, t) \propto e^{i(\vec k\cdot\vec x - \omega t)}$, with $\omega^2 = k^2 + m^2$. Identifying k with the momentum and ω with the energy, as one would in quantum mechanics (with $\hbar = 1$) gives the relation one would expect for a particle of mass m: $E^2 = p^2 + m^2$ (as we have set $c = 1$ also. $E^2 = p^2c^2 + m^2c^4$ if you want to put c back in).

The only relativistic field we are familiar with from classical (non-quantum) mechanics is the electromagnetic field. We are familiar with $\vec E$ and $\vec B$ as fields defined throughout space and also functions of time. But $\vec E$ and $\vec B$ satisfy constraint equations. Maxwell's equations (in free space and SI units) are

$$\vec\nabla\cdot\vec E = \rho/\epsilon_0 \tag{8.20}$$
$$\vec\nabla\cdot\vec B = 0 \tag{8.21}$$
$$\vec\nabla\times\vec E + \frac{1}{c}\frac{\partial\vec B}{\partial t} = 0 \tag{8.22}$$
$$\vec\nabla\times\vec B - \frac{1}{c^2}\frac{\partial\vec E}{\partial t} = \mu_0\vec j \tag{8.23}$$

Notice that (8.20) and (8.21) are not equations of motion, as they do not involve time derivatives. Instead they are equations of constraint, best implemented by solving them in terms of independent degrees of freedom. As we saw in section (2.7), these equations allow us to consider $\vec E$ and $\vec B$ as derivatives of the magnetic vector potential $\vec{A}(\vec{x},t)$ and the electrostatic potential $\phi(\vec{x},t)$. Then we have $\vec{B}=\vec{\nabla}\times\vec{A}$, and $\vec{E}=-\vec{\nabla}\phi-\frac{1}{c}\frac{\partial\vec{A}}{\partial t}$. We also saw that the interaction of these fields with a charged particle could be given in terms of a potential

$$U(\vec{r},\vec{v})=q\left(\phi(r,t)-(\vec{v}/c)\cdot\vec{A}(\vec{r},t)\right)=-\frac{q}{c}\frac{dx^\mu}{dt}A_\mu,$$

if I take $A_0=-\phi=-A^0$. This is the first step in writing electromagnetism in relativistic notation[15].

The connection to $\vec{E}$ and $\vec{B}$ is best understood if we define a 1-form from A and its exterior derivative:

$$\mathbf{A}:=A_\mu(x^\nu)dx^\mu,\quad \mathbf{F}:=d\mathbf{A}=\frac{\partial A_\mu}{\partial x^\nu}dx^\nu\wedge dx^\mu=\frac{1}{2}F_{\mu\nu}dx^\mu\wedge dx^\nu.$$

Examining the components, we have

$$F_{0j} = \frac{1}{c}\dot{A}_j+\frac{\partial\phi}{\partial x_j}=-E_j=-F_{j0}, \tag{8.24}$$

$$F_{ij} = \frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}=\sum_k\epsilon_{ijk}(\vec{\nabla}\times\vec{A})_k=\sum_k\epsilon_{ijk}B_k. \tag{8.25}$$

As $\mathbf{F}:=d\mathbf{A}$ we know that $d\mathbf{F}=0$. $d\mathbf{F}$ is a 3-form,

$$d\mathbf{F}=\frac{1}{6}(d\mathbf{F})_{\mu\nu\rho}dx^\mu dx^\nu dx^\rho=\frac{1}{6}\epsilon_{\mu\nu\rho\sigma}V^\sigma dx^\mu dx^\nu dx^\rho,$$

where $V^\sigma=(-1/6)\epsilon^{\mu\nu\rho\sigma}(d\mathbf{F})_{\mu\nu\rho}$. As we saw in three dimensions in section (6.5), a k-form ω in D dimensions can be associated not only with an antisymmetric tensor of rank k, but also with one of rank $D-k$. That tensor is associated with a $(D-k)$-form, called the **Hodge dual**[16] of ω, written $*\omega$. On the basis vectors we define

$$*(dx^{\mu_1}\wedge\cdots dx^{\mu_k})=\frac{1}{(D-k)!}\epsilon^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_{D-k}}dx^{\nu_1}\wedge\cdots dx^{\nu_{D-k}}.$$

---

[15] Note U is not an invariant, nor should it be, as it is part of the energy. Therefore it is expected that it should transform like d/dt of a scalar.

[16] Warning: the dual of the dual of a k-form ω is $\pm\omega$, with the sign depending on D and k.

In particular, if ω is a 2-form in four dimensional Minkowski space,

$$\begin{aligned}
\omega &= \frac{1}{2}\omega_{\mu\nu}dx^\mu \wedge dx^\nu \\
*\,\omega &= \frac{1}{2}\left(\frac{1}{2}\epsilon^{\mu\nu}{}_{\rho\sigma}\omega_{\mu\nu}\right)dx^\rho \wedge dx^\sigma \\
d\omega &= \frac{1}{2}\omega_{\mu\nu,\rho}dx^\rho \wedge dx^\mu \wedge dx^\nu \\
*\,d\omega &= \frac{1}{2}\epsilon^{\mu\nu\rho}{}_\sigma\omega_{\mu\nu,\rho}dx^\sigma \\
\,d\omega &= \epsilon^{\kappa\rho\sigma}{}_\tau\left(\frac{1}{4}\epsilon^{\mu\nu}{}_{\rho\sigma}\omega_{\mu\nu,\kappa}\right)dx^\tau = \omega_{\tau\kappa,\kappa}dx^\tau.
\end{aligned}$$

In particular for our 2-form $\mathbf{F}$, the fact that $d\mathbf{F} = 0$, and thus $*d\mathbf{F} = 0$ tells us the vector $V^\sigma = (-1/6)\epsilon^{\mu\nu\rho\sigma}F_{\nu\rho,\mu} = 0$. The $\sigma = 0$ component of this

$$0 = 3V^0 = \frac{1}{2}\epsilon^{ijk}F_{jk,i} = \frac{1}{2}\epsilon^{ijk}\epsilon_{jk\ell}B_{\ell,i} = \delta_{i\ell}B_{\ell,i} = \vec{\nabla}\cdot\vec{B},$$

giving us the constraint equation (8.21). For the spatial component,

$$\begin{aligned}
0 &= -3V^i = \frac{1}{2}\sum_{\mu,\nu,\rho=0}^{3}\epsilon^{\mu\nu\rho i}F_{\nu\rho,\mu} = \frac{1}{2}\sum_{j,k=1}^{3}\left(\epsilon^{jki}F_{jk,0} + 2\epsilon^{jki}F_{k0,j}\right) \\
&= \frac{1}{2}\sum_{j,k=1}^{3}\left(\epsilon^{jki}\frac{1}{c}\epsilon_{jk\ell}\dot{B}_\ell + 2\epsilon^{jki}\partial_j E_k\right) \\
&= \left(\frac{1}{c}\dot{\vec{B}} + \vec{\nabla}\times\vec{E}\right)_i,
\end{aligned}$$

which gives us the constraint (8.22). So the two constraint equations among Maxwell's four are

$$d\mathbf{F} = 0. \tag{8.26}$$

What are the two dynamical equations? If we evaluate $*\,d*F = F_{\mu\nu,\nu}dx^\mu =: V_\mu dx^\mu$, we see the zeroth component contains only $F_{0j} = -E_j$, with $V_0 = \sum_j \partial F_{0j}/\partial x^j = -\vec{\nabla}\cdot\vec{E}$, which Maxwell tells us is $-\rho/\epsilon_0$. The spatial component is $V_i = F_{i0,0} + \sum_j F_{ij,j} = \dot{E}_j/c + \epsilon_{ijk}\partial_j B_k = \left(\dot{\vec{E}}/c + \vec{\nabla}\times\vec{B}\right)_i$ which Maxwell tells us is (modulo c) $\mu_0(\vec{j})_i$. This encourages us to define the 4-vector $J^\mu = (\rho, \vec{j})$ and its accompanying 1-form $\mathbf{J} = J_\mu dx^\mu$, and to write the two dynamical equations as

$$*\,d*\mathbf{F} = -\mathbf{J} \quad \text{or } d*\mathbf{F} = *\mathbf{J}. \tag{8.27}$$

How should we write the lagrangian density for the electromagnetic fields? As the dynamics is determined by the action, the integral of $\mathcal{L}$ over four-dimensional space-time, we should expect $\mathcal{L}$ to be essentially a 4-form, which needs to be made out of the 2-form $\mathbf{F}$. Our first idea might be to try $\mathbf{F}\wedge\mathbf{F}$, which is a 4-form, but unfortunately it is a closed 4-form, for $d(\mathbf{F}\wedge\mathbf{F}) = (d\mathbf{F})\wedge\mathbf{F} + (\mathbf{F})\wedge(d\mathbf{F})$, and $d\mathbf{F} = dd\mathbf{A} = 0$. Because we are working on a contractable space, $\mathbf{F}\wedge\mathbf{F}$ is thereform exact, and an exact form is useless as a lagrangian density because $\int d\omega d^4x = \int_S \omega$ which depends only on the boundaries, both in space and time, but this is exactly where variations of the dynamical degrees of freedom are kept fixed in determing the variation of the action.

There is another 2-form available, however, $*F$, so we might consider

$$\begin{aligned}
\mathcal{L}dtd^3x &= -\frac{1}{2}\mathbf{F}\wedge *\mathbf{F} = -\frac{1}{2}\cdot\frac{1}{2}F_{\mu\nu}dx^\mu\wedge dx^\nu\cdot\frac{1}{4}\epsilon^{\kappa\lambda}{}_{\rho\sigma}F_{\kappa\lambda}\wedge dx^\rho\wedge dx^\sigma \\
&= -\frac{1}{16}\epsilon^{\kappa\lambda}{}_{\rho\sigma}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\kappa\lambda}dx^0\wedge dx^1\wedge dx^2\wedge dx^3 \\
\mathcal{L} &= -\frac{c}{16}\epsilon^{\kappa\lambda}{}_{\rho\sigma}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\kappa\lambda} = -\frac{c}{8}(F^{\mu\nu}F_{\mu\nu} - F^{\mu\nu}F_{\nu\mu}) \\
&= -\frac{c}{4}F^{\mu\nu}F_{\mu\nu} = -\frac{c}{2}(-F_{0j}^2 + \frac{1}{2}\epsilon_{ijk}B_k\epsilon_{ij\ell}B_\ell = \frac{1}{2}(E^2 - B^2)
\end{aligned}$$

Exercises

8.1 The Lagrangian density for the electromagnetic field in vacuum may be written

$$\mathcal{L} = \frac{1}{2}\left(\vec{E}^{\,2} - \vec{B}^{\,2}\right),$$

where the dynamical degrees of freedom are not $\vec{E}$ and $\vec{B}$, but rather $\vec{A}$ and ϕ, where

$$\begin{aligned}
\vec{B} &= \vec{\nabla}\times A \\
\vec{E} &= -\vec{\nabla}\phi - \frac{1}{c}\dot{\vec{A}}
\end{aligned}$$

a) Find the canonical momenta, and comment on what seems unusual about one of the answers.
b) Find the Lagrange Equations for the system. Relate to known equations for the electromagnetic field.

8.2 A tensor transforms properly under Lorentz transformations as specified by equation (8.6), with each index being contracted with a suitable $L^{\cdot}{}_{\cdot}$ or $L_{\cdot}{}^{\cdot}$ as appropriate.
(a) The Minkowsky metric $\eta_{\mu\nu}$ should then be transformed into a new tensor by contracting with two $L_{\cdot}{}^{\cdot}$'s. Show that the new tensor η' is nonetheless the same as η. [That is, each element still has the same value].
(b) The Levi-Civita symbol in one reference frame is defined by $\epsilon^{0123} = 1$ and $\epsilon^{\mu\nu\rho\sigma}$ is antisymmetric under any interchange of two indices. Being a four-index contravariant tensor, it will transform with four $L^{\cdot}{}_{\cdot}$'s. Show that the transformed tensor still has the same values under proper[17] Thus both $\eta_{\mu\nu}$ and $\epsilon^{\mu\nu\rho\sigma}$ are both invariant and transform co- or contra-variantly.
(c) Show that if $T^{\rho_1\ldots\rho_j\,\mu}{}_{\sigma_1\ldots\sigma_k}$ transforms correctly, the tensor $T^{\rho_1\ldots\rho_j}{}_{\mu\,\sigma_1\ldots\sigma_k} := \eta_{\mu\nu}T^{\rho_1\ldots\rho_j\,\nu}{}_{\sigma_1\ldots\sigma_k}$ transforms correctly as well.
(d) Show that if two indices, one upper and one lower, are contracted, that is, set equal and summed over, the resulting object transforms as if those indices were not there. That is, $W^{\mu_1\ldots\mu_j}{}_{\rho_1\ldots\rho_k} := T^{\nu\mu_1\ldots\mu_j}{}_{\nu\rho_1\ldots\rho_k}$ transforms correctly.

[17]Proper Lorentz transformations are those that can be generated continuously from the identity. That is, they exclude transformations that reverse the direction of time or convert a right-handed coordinate system to a left-handed one.