Chapter 1

Classical Action Principles

1.1 Classical Lagrange-Hamilton Principle

This is sometimes referred to as the principle of least action. Let the classical system under consideration be described by N generalized coordinates

$$q = \{q_i\}_{i=1,\dots N}, \tag{1.1a}$$

and the corresponding velocities

$$\dot{q} = \left\{\dot{q}_i \equiv \frac{d}{dt} q_i\right\}_{i=i,\dots N}. \tag{1.1b}$$

The dynamics of the system is specified by giving the Lagrangian, $L = L(q, \dot{q}, t)$. The action W is the time integral of the Lagrangian from some initial time t_2 to some final time t_1,

$$W_{12} = \int_2^1 dt\, L(q(t), \dot{q}(t), t). \tag{1.2}$$

The action principle states that under infinitesimal variations, the change in the action depends only on the endpoints, that is,

$$\delta W_{12} = G_1 - G_2, \tag{1.3}$$

where G_a is a function depending only on dynamical variables at time t_a. In other words, the action is *stationary* with respect to variations between 2 and 1. This stationary property picks out the physical trajectory connecting q_2, $\dot{q}_2$ and q_1, $\dot{q}_1$.

The Lagrangian for a nonrelativistic particle of mass m moving in a potential $V(\mathbf{r})$ is

$$L = T - V = \frac{1}{2} m \dot{\mathbf{r}}^2 - V(\mathbf{r}), \tag{1.4}$$

where the independent variables are $\mathbf{r}$ and t. The possible variations are a change in the path, $\delta\mathbf{r}$, and a change in the time of the endpoints δt_2, δt_1. However, for the latter it is more convenient to define a change in the time parameter $t \to t + \delta t(t)$ where $\delta t(t_1) = \delta t_1$, $\delta t(t_2) = \delta t_2$. Then

$$dt \to d(t + \delta t) = dt\left(1 + \frac{d\delta t}{dt}\right), \tag{1.5a}$$

$$\frac{d}{dt} \to \left(1 - \frac{d\delta t}{dt}\right)\frac{d}{dt}. \tag{1.5b}$$

Because of this change in t, the limits of integration in W_{12} are *not* changed.

We are now ready to compute the infinitesimal variation in W_{12}:

$$\begin{aligned}\delta W_{12} &= \int_2^1 dt\biggl\{m\frac{d\mathbf{r}}{dt}\cdot\frac{d}{dt}\delta\mathbf{r} - \delta\mathbf{r}\cdot\boldsymbol{\nabla}V(\mathbf{r}) \\ &\qquad + \frac{d\delta t}{dt}\left[\frac{1}{2}m\left(\frac{d\mathbf{r}}{dt}\right)^2 - V(\mathbf{r})\right] - m\left(\frac{d\mathbf{r}}{dt}\right)^2\frac{d\delta t}{dt}\biggr\} \\ &= \int_2^1 dt\biggl\{\frac{d}{dt}\left[m\dot{\mathbf{r}}\cdot\delta\mathbf{r} - \delta t\left(\frac{1}{2}m\dot{\mathbf{r}}^2 + V\right)\right] \\ &\qquad + \delta\mathbf{r}\cdot[-m\ddot{\mathbf{r}} - \boldsymbol{\nabla}V] + \delta t\frac{d}{dt}\left[\frac{1}{2}m\dot{\mathbf{r}}^2 + V\right]\biggr\}.\end{aligned} \tag{1.6}$$

Because $\delta\mathbf{r}$ and δt are independent variations, we conclude that

$$m\ddot{\mathbf{r}} = -\boldsymbol{\nabla}V, \tag{1.7a}$$

which is Newton's law, and

$$\frac{dE}{dt} = 0, \quad \text{where} \quad E = \frac{1}{2}m\dot{\mathbf{r}}^2 + V(\mathbf{r}) \quad \text{is the energy,} \tag{1.7b}$$

which is the statement of energy conservation. What is left of the variation comes only from the endpoints, so we infer the form of the "generators,"

$$G = \mathbf{p}\cdot\delta\mathbf{r} - E\delta t, \quad \mathbf{p} = m\dot{\mathbf{r}} = \text{momentum}. \tag{1.8}$$

Let us repeat this analysis for a general Lagrangian, $L(q_i, \dot{q}_i, t)$. In the following we will adopt a summation convention of summing over repeated indices i. We find

$$\begin{aligned}\delta W_{12} &= \int_2^1 dt\biggl\{\frac{\partial L}{\partial\dot{q}_i}\frac{d}{dt}\delta q_i + \frac{\partial L}{\partial q_i}\delta q_i + \frac{d\delta t}{dt}L - \frac{\partial L}{\partial\dot{q}_i}\dot{q}_i\frac{d\delta t}{dt} + \frac{\partial L}{\partial t}\delta t\biggr\} \\ &= \int_2^1 dt\biggl\{\frac{d}{dt}\left[\frac{\partial L}{\partial\dot{q}_i}\delta q_i + \delta t\left(L - \frac{\partial L}{\partial\dot{q}_i}\dot{q}_i\right)\right] \\ &\qquad + \delta q_i\left(-\frac{d}{dt}\frac{\partial L}{\partial\dot{q}_i} + \frac{\partial L}{\partial q_i}\right) - \delta t\frac{d}{dt}\left(L - \frac{\partial L}{\partial\dot{q}_i}\dot{q}_i\right) + \delta t\frac{\partial L}{\partial t}\biggr\}.\end{aligned} \tag{1.9}$$

From the interior terms we deduce

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \tag{1.10}$$

which is the Euler-Lagrange equation, and the equation of energy conservation (if there is no explicit time dependence)

$$\frac{d}{dt}H = \frac{\partial H}{\partial t}, \tag{1.11}$$

where the energy or Hamiltonian is

$$H = p_i \dot{q}_i - L, \tag{1.12}$$

in terms of the generalized momentum

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{1.13}$$

Finally, the generator has the form

$$G = p_i \delta q_i - H \delta t. \tag{1.14}$$

Let us now return to the simple one-particle system and write the Hamiltonian:

$$H(\mathbf{p}, \mathbf{r}) = \mathbf{p} \cdot \dot{\mathbf{r}} - L = \frac{1}{2}\frac{p^2}{m} + V(\mathbf{r}) = T + V. \tag{1.15}$$

We are now to regard $\mathbf{r}$, $\mathbf{p}$, and t as independent variables. Then the variation of the action is

$$\begin{aligned}
\delta W_{12} &= \delta \int_2^1 dt \left(\mathbf{p} \cdot \frac{d\mathbf{r}}{dt} - H \right) \\
&= \int_2^1 dt \left[\mathbf{p} \frac{d}{dt} \delta \mathbf{r} - \delta \mathbf{r} \cdot \frac{\partial H}{\partial \mathbf{r}} + \delta \mathbf{p} \cdot \frac{d\mathbf{r}}{dt} - \delta \mathbf{p} \cdot \frac{\partial H}{\partial \mathbf{p}} - \frac{d\delta t}{dt} H - \delta t \frac{\partial H}{\partial t} \right] \\
&= \int_2^1 dt \bigg\{ \frac{d}{dt} \left[\mathbf{p} \cdot \delta \mathbf{r} - H \delta t \right] + \delta \mathbf{r} \cdot \left[-\frac{d\mathbf{p}}{dt} - \frac{\partial H}{\partial \mathbf{r}} \right] \\
&\qquad + \delta \mathbf{p} \cdot \left[\frac{d\mathbf{r}}{dt} - \frac{\partial H}{\partial \mathbf{p}} \right] + \delta t \left(\frac{dH}{dt} - \frac{\partial H}{\partial t} \right) \bigg\}.
\end{aligned} \tag{1.16}$$

From this we infer the three Hamilton's equations,

$$\frac{d\mathbf{r}}{dt} = \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p}}{m} \quad \text{here}, \tag{1.17a}$$

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{r}} = -\boldsymbol{\nabla} V \quad \text{here}, \tag{1.17b}$$

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}. \tag{1.17c}$$

The generators are

$$G = \mathbf{p} \cdot \delta\mathbf{r} - H\,\delta t. \tag{1.18}$$

The generalization to $H(q_i, p_i)$ is immediate:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \tag{1.19a}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \tag{1.19b}$$

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}. \tag{1.19c}$$

The generators are

$$G = p_i \delta q_i - H\delta t. \tag{1.20}$$

Suppose we consider a function of the dynamical variables, $f(q_i, p_i, t)$. Its time derivative is

$$\begin{aligned}\frac{df}{dt} &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial f}{\partial p_i}\frac{dp_i}{dt} \\ &= \frac{\partial f}{\partial t} + \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial f}{\partial p_i}\end{aligned} \tag{1.21}$$

We define the *Poisson bracket* by

$$\{f, g\} \equiv \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}, \tag{1.22}$$

so we have

$$\frac{d}{dt}f = \frac{\partial f}{\partial t} + \{H, f\}. \tag{1.23}$$

Thus, if the following two conditions hold,

1. there is no explicit time dependence of f,

$$\frac{\partial f}{\partial t} = 0, \tag{1.24a}$$

   and

2. the Poisson bracket of f and H vanishes,

$$\{H, f\} = 0, \tag{1.24b}$$

then f is a constant of the motion.

It is sometimes useful to adopt a viewpoint intermediate between that of Lagrange and that of Hamilton. Let us write for the Lagrangian of our single particle system

$$\begin{aligned}L &= \mathbf{p} \cdot \left(\frac{d\mathbf{r}}{dt} - \mathbf{v}\right) + \frac{1}{2}m\mathbf{v}^2 - V(\mathbf{r}) \\ &= \mathbf{p} \cdot \frac{d\mathbf{r}}{dt} - H(\mathbf{r}, \mathbf{p}, \mathbf{v}),\end{aligned} \tag{1.25}$$

where

$$H(\mathbf{r},\mathbf{p},\mathbf{v}) = \mathbf{p}\cdot\mathbf{v} - \frac{1}{2}mv^2 + V(\mathbf{r}). \tag{1.26}$$

We are now to regard $\mathbf{r}$, $\mathbf{p}$, and $\mathbf{v}$ as independent variables. Then the variation of the action is

$$\begin{aligned}
\delta W_{12} &= \int_2^1 dt\biggl\{\mathbf{p}\cdot\frac{d}{dt}\delta\mathbf{r} - \delta\mathbf{r}\cdot\frac{\partial H}{\partial \mathbf{r}} \\
&\qquad + \delta\mathbf{p}\cdot\left[\frac{d\mathbf{r}}{dt} - \frac{\partial H}{\partial \mathbf{p}}\right] - \delta\mathbf{v}\cdot\frac{\partial H}{\partial \mathbf{v}} - \frac{d\delta t}{dt}H - \delta t\frac{\partial H}{\partial t}\biggr\} \\
&= \int_2^1 dt\biggl\{\frac{d}{dt}\left[\mathbf{p}\cdot\delta\mathbf{r} - H\,\delta t\right] \\
&\qquad - \delta\mathbf{r}\cdot\left[\frac{d\mathbf{p}}{dt} + \frac{\partial H}{\partial \mathbf{r}}\right] + \delta\mathbf{p}\cdot\left[\frac{d\mathbf{r}}{dt} - \frac{\partial H}{\partial \mathbf{p}}\right] \\
&\qquad - \delta\mathbf{v}\cdot\frac{\partial H}{\partial \mathbf{v}} + \delta t\left(\frac{d}{dt}H - \frac{\partial H}{\partial t}\right)\biggr\}.
\end{aligned} \tag{1.27}$$

This implies the four "equations of motion,"

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{r}} \quad (= -\boldsymbol{\nabla}V \quad \text{here})\,, \tag{1.28a}$$

$$\frac{d\mathbf{r}}{dt} = \frac{\partial H}{\partial \mathbf{p}} \quad (= \mathbf{v} \quad \text{here})\,, \tag{1.28b}$$

$$0 = \frac{\partial H}{\partial \mathbf{v}} \quad (= \mathbf{p} - m\mathbf{v} \quad \text{here})\,, \tag{1.28c}$$

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \tag{1.28d}$$

and the generator

$$G = \mathbf{p}\cdot\delta\mathbf{r} - H\,\delta t. \tag{1.28e}$$

1.1.1 Generators

The generators interrelate conservation laws and invariances of the system.

1. Suppose the action is invariant under a rigid displacement (translation) of the coordinate system:

$$\delta W_{12} = 0 = \mathbf{p}_1\cdot\delta\mathbf{r}_1 - \mathbf{p}_2\cdot\delta\mathbf{r}_2, \tag{1.29}$$

   where $\delta\mathbf{r}_1 = \delta\mathbf{r}_2$ for a rigid displacement. Then

$$\mathbf{p}_1 = \mathbf{p}_2, \tag{1.30}$$

   that is, momentum is conserved. By our equations of motion, this will be true, of course, only if V is constant. Conversely, if V is constant, W is invariant under a translation of the coordinate system.

2. If W is invariant under a rigid displacement in time (time translation, for which $\delta t_1 = \delta t_2$)
$$\delta W_{12} = 0 = -H_1 \delta t_1 + H_2 \delta t_2, \tag{1.31}$$
which implies
$$H_1 = H_2, \tag{1.32}$$
that is, energy is conserved. This is consistent with our equations of motion, unless H has *explicit* time dependence, in which case
$$\frac{dH}{dt} = \frac{\partial H}{\partial t}. \tag{1.33}$$

3. Suppose W is invariant under rigid rotations,
$$\delta \mathbf{r} = \delta \boldsymbol{\omega} \times \mathbf{r}. \tag{1.34}$$
Then
$$\begin{aligned} 0 = \delta W_{12} &= \mathbf{p}_1 \cdot \delta \mathbf{r}_1 - \mathbf{p}_2 \cdot \delta \mathbf{r}_2 \\ &= \delta \boldsymbol{\omega} \cdot (\mathbf{r}_1 \times \mathbf{p}_1 - \mathbf{r}_2 \times \mathbf{p}_2), \end{aligned} \tag{1.35}$$
which means that $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is conserved. This will be true here provided $V(\mathbf{r}) = V(|\mathbf{r}|)$.

1.2 Classical Field Theory

Let us move on to classical field theory by writing down the appropriate Lagrangian for *relativistic classical electrodynamics*[1]:

$$\begin{aligned} L = \sum_k \mathbf{p}_k \cdot \left(\frac{d\mathbf{r}_k}{dt} - \mathbf{v}_k \right) - m_{0k} c^2 \sqrt{1 - \frac{v_k^2}{c^2}} + \frac{e_k}{c} v_k^\mu A_\mu(\mathbf{r}_k) \\ + \int (d\mathbf{r}) \left[-\frac{1}{2} F^{\mu\nu} (\partial_\mu A_\nu - \partial_\nu A_\mu) + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right], \end{aligned} \tag{1.36}$$

where m_{0k} is the rest mass of the kth particle, which has velocity $\mathbf{v}_k$, position $\mathbf{r}_k$, and momentum $\mathbf{p}_k$. Appearing here is the four-velocity

$$v_k^\mu = (c, \mathbf{v}_k), \text{ that is } v_k^0 = c, v_k^i = \mathbf{v}_k^i, \tag{1.37}$$

where we have adopted the usual convention that Greek indices run over four values, $\mu = 0, 1, 2, 3$, while Latin indices take on only the three spatial values, $i = 1, 2, 3$. Note that

$$dt\, v^\mu = (c\, dt, d\mathbf{r}) \tag{1.38}$$

[1]For motivation, see J. Schwinger, L. L. DeRaad, Jr., K. A. Milton, and W.-y. Tsai, *Classical Electrodynamics* [Perseus (Westview Press), New York, 1998]

is a four-vector. The four-gradient is

$$\partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \boldsymbol{\nabla}\right), \tag{1.39}$$

and the four-vector potential is

$$A^\mu = (\phi, \mathbf{A}), \tag{1.40}$$

in terms of the usual scalar (ϕ) and vector ($\mathbf{A}$) potentials. $F^{\mu\nu}$ is the electromagnetic field strength tensor, which is antisymmetric,

$$F^{\mu\nu} = -F^{\nu\mu}, \tag{1.41}$$

and therefore has six distinct nonzero components which are the electric and magnetic field strengths,

$$F^{0i} = E^i, \quad F^{ij} = \epsilon^{ijk} B_k, \tag{1.42}$$

where the antisymmetric tensor (Levi-Civita symbol) is defined by

$$\epsilon^{123} = +1, \quad \epsilon^{ijk} = \epsilon^{jki} = \epsilon^{kij} = -\epsilon^{jik} = -\epsilon^{ikj} = -\epsilon^{kji}. \tag{1.43}$$

Indices are lowered with the metric tensor

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.44}$$

so for example

$$A_\mu = g_{\mu\nu} A^\nu = (-\phi, \mathbf{A}), \tag{1.45}$$

and the summation convention over repeated indices is used.

Let us work out the four independent variations of L with respect to particle variables: ($\boldsymbol{\nabla}_k = \partial/\partial \mathbf{r}_k$)

$$\delta \mathbf{r}_k : \quad \delta L = \frac{d}{dt}(\delta \mathbf{r}_k \cdot \mathbf{p}_k) + \delta \mathbf{r}_k \cdot \left(-\frac{d\mathbf{p}_k}{dt} + \frac{e_k}{c} v_k^\mu \boldsymbol{\nabla}_k A_\mu(\mathbf{r}_k)\right), \tag{1.46a}$$

$$\delta \mathbf{p}_k : \quad \delta L = \delta \mathbf{p}_k \cdot \left(\frac{d\mathbf{r}_k}{dt} - \mathbf{v}_k\right), \tag{1.46b}$$

$$\delta \mathbf{v}_k : \quad \delta L = \delta \mathbf{v}_k \cdot \left(-\mathbf{p}_k + \frac{m_{0k}\mathbf{v}_k}{\sqrt{1 - v_k^2/c^2}} + \frac{e_k}{c}\mathbf{A}(\mathbf{r}_k)\right), \tag{1.46c}$$

$$\delta t : \quad \delta L = \frac{d}{dt}(-H\delta t) + \delta t \frac{dH}{dt}, \tag{1.46d}$$

assuming no explicit time dependence, so the action principle implies

$$\mathbf{v}_k = \frac{d\mathbf{r}_k}{dt}, \tag{1.47a}$$

$$\mathbf{p}_k = \frac{m_{0k}\mathbf{v}_k}{\sqrt{1-v_k^2/c^2}} + \frac{e_k}{c}\mathbf{A}(\mathbf{r}_k), \tag{1.47b}$$

$$\frac{d\mathbf{p}_k}{dt} = \frac{e_k}{c}\boldsymbol{\nabla}_k v_k^\mu A_\mu(\mathbf{r}_k), \tag{1.47c}$$

$$\frac{dH}{dt} = 0, \tag{1.47d}$$

where the Hamiltonian has a particle and a field part,

$$H = \sum_k H_k + H_f, \tag{1.48}$$

where

$$\begin{aligned} H_k &= \mathbf{p}_k\cdot\mathbf{v}_k + m_{0k}c^2\sqrt{1-\frac{v_k^2}{c^2}} - \frac{e_k}{c}v_k^\mu A_\mu(\mathbf{r}_k) \\ &= \frac{m_{0k}c^2}{\sqrt{1-v_k^2/c^2}} + e_k\phi(\mathbf{r}_k), \end{aligned} \tag{1.49}$$

where Eq. (1.47b) was used to eliminate $\mathbf{p}_k$, and the field part will be given below.

We continue by working out the field variations of Eq. (1.36):

$$\delta A_\mu: \quad \delta L = \int (d\mathbf{r})\left[\delta A_\mu\left(j^\mu - \partial_\nu F^{\mu\nu}\right) + \partial_\nu\left(\delta A_\mu F^{\mu\nu}\right)\right], \tag{1.50}$$

where the current density is

$$j^\mu(\mathbf{r}) = \sum_k \frac{e_k}{c}v_k^\mu\delta(\mathbf{r}-\mathbf{r}_k) = (\rho, \mathbf{j}), \tag{1.51}$$

so that

$$\int (d\mathbf{r})A_\mu(\mathbf{r})j^\mu(\mathbf{r}) = \sum_k \frac{e_k}{c}v_{k\mu}A^\mu(\mathbf{r}_k). \tag{1.52}$$

The two remaining variations are

$$\delta F^{\mu\nu}: \quad \delta L = \int (d\mathbf{r})\delta F^{\mu\nu}\left[-\frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu) + \frac{1}{2}F_{\mu\nu}\right], \tag{1.53}$$

$$\delta t: \quad \delta W_f = \int dt(d\mathbf{r})\frac{d\delta t}{dt}\left[-F^{i\nu}\partial_i A_\nu + \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right]. \tag{1.54}$$

The action principle thus implies Maxwell's equations,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{1.55a}$$

$$\partial_\nu F^{\mu\nu} = j^\mu, \tag{1.55b}$$

and gives the field part of the Hamiltonian,

$$\begin{aligned} H_f &= \int (d\mathbf{r})\left[F^{i\nu}\partial_i A_\nu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right] \\ &= \int (d\mathbf{r})\left[\mathbf{E}\cdot\boldsymbol{\nabla}\phi + \mathbf{B}\cdot\boldsymbol{\nabla}\times\mathbf{A} + \frac{1}{2}(E^2 - B^2)\right]. \end{aligned} \tag{1.56}$$

But using the field equations that give the construction of the field strengths in terms of the potentials,

$$E_i = -F^{i0} = -\partial^i A^0 + \partial^0 A^i - -\nabla_i \phi - \frac{\partial A_i}{c\partial t}, \tag{1.57a}$$

$$B_i = \frac{1}{2}\epsilon_{ijk} F_{jk} = \epsilon_{ijk}\partial_j A_k = (\boldsymbol{\nabla} \times \mathbf{A})_i, \tag{1.57b}$$

the field part of the Hamiltonian becomes

$$H_f = \int (d\mathbf{r}) \left[\frac{1}{2}(E^2 + B^2) - \rho\phi\right], \tag{1.58}$$

since $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho$. Notice then that the total Hamiltonian, from Eqs. (1.58) and (1.49) is simply, from Eq. (1.51)

$$H = \sum_k m_{0k} c^2 (1 - v_k^2/c^2)^{-1/2} + \int (d\mathbf{r}) \frac{1}{2}(E^2 + B^2), \tag{1.59}$$

the sum of the free particle Hamiltonians and the pure field part. It appears that the interaction has disappeared. This, of course, is not the case, because $\mathbf{E}, \mathbf{B}$ depend on the particle positions and velocities.

What about the generators? From Eqs. (1.46a), (1.46d), and (1.50) we have

$$G = \sum_k \delta\mathbf{r}_k \cdot \mathbf{p}_k - \frac{1}{c}\int (d\mathbf{r}) \delta\mathbf{A} \cdot \mathbf{E} - H\delta t. \tag{1.60}$$

This says that just as $\mathbf{p}_k$ is canonically conjugate to $\mathbf{r}_k$, $-\mathbf{E}$ is cononically conjugate to $\mathbf{A}/c$. In fact, if we introduce the Lagrange density according to

$$L = \int (d\mathbf{r}) \mathcal{L}, \tag{1.61}$$

we have from Eq. (1.36),

$$c\frac{\partial \mathcal{L}}{\partial \dot{A}_i} = -F^{0i} = -E^i. \tag{1.62}$$

[Cf. Eq. (1.13).]

1.2.1 Field Momentum and Angular Momentum

Consider a rigid displacement of the origin of the coordinate system,

$$\mathbf{r} \to \mathbf{r} + \delta\mathbf{r}, \tag{1.63}$$

which is sketched in Fig. 1.1. A quantity F which is coordinate independent is a different function of the old and new coordinates:

$$F(\mathbf{r}) = \bar{F}(\mathbf{r} + \delta\mathbf{r}), \tag{1.64}$$

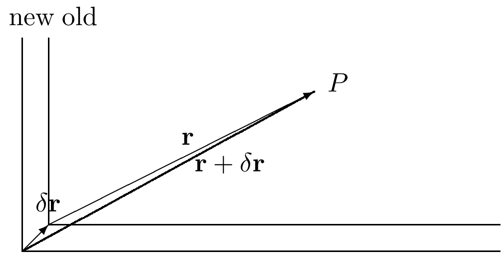

Figure 1.1: Description of a physical point P in two different coordinate systems, labeled “old” and “new,” that differ by a displacement $\delta\mathbf{r}$.

that is, the “new” function of the “new” coordinate is the same as the “old” function of the “old” coordinate. Because the change in coordinates is infinitesimal, the new function differs only slightly from the old function,

$$\bar{F}(\mathbf{r}) = F(\mathbf{r}) + \delta F(\mathbf{r}), \tag{1.65}$$

so

$$\delta F(\mathbf{r}) = F(\mathbf{r} - \delta\mathbf{r}) - F(\mathbf{r}) = -\delta\mathbf{r}\cdot\boldsymbol{\nabla}F(\mathbf{r}). \tag{1.66}$$

The field generator corresponding to this coordinate displacement thus is

$$\begin{aligned}
G_f &= -\frac{1}{c}\int(d\mathbf{r})\delta\mathbf{A}\cdot\mathbf{E} = \frac{1}{c}\int(d\mathbf{r})[(\delta\mathbf{r}\cdot\boldsymbol{\nabla})\mathbf{A}]\cdot\mathbf{E}\\
&= -\frac{1}{c}\int(d\mathbf{r})[\delta\mathbf{r}\times(\boldsymbol{\nabla}\times\mathbf{A}) - \delta\mathbf{r}\cdot(\boldsymbol{\nabla})\cdot\mathbf{A}]\cdot\mathbf{E}\\
&= \frac{1}{c}\int(d\mathbf{r})[(\mathbf{E}\times\mathbf{B})\cdot\delta\mathbf{r} - (\boldsymbol{\nabla}\cdot\mathbf{E})(\mathbf{A}\cdot\delta\mathbf{r})]\\
&= \frac{1}{c}\int(d\mathbf{r})(\mathbf{E}\times\mathbf{B})\cdot\delta\mathbf{r} - \sum_k\frac{e_k}{c}\mathbf{A}(\mathbf{r}_k)\cdot\delta\mathbf{r},
\end{aligned} \tag{1.67}$$

where we used $\boldsymbol{\nabla}\cdot\mathbf{E} = \rho$ and Eq. (1.51). The total generator corresponding to the coordinate displacement is

$$G = \sum_k G_k + G_f = \mathbf{P}\cdot\delta\mathbf{r}, \tag{1.68}$$

where the total momentum is, from Eqs. (1.60) and (1.67),

$$\begin{aligned}
\mathbf{P} &= \sum_k\left(\mathbf{p}_k - \frac{e_k}{c}\mathbf{A}(\mathbf{r}_k)\right) + \frac{1}{c}\int(d\mathbf{r})\mathbf{E}\times\mathbf{B}\\
&= \sum_k m_k\mathbf{v}_k + \frac{1}{c}\int(d\mathbf{r})\mathbf{E}\times\mathbf{B},
\end{aligned} \tag{1.69}$$

where we have used Eq. (1.47b) and introduced the relativistic mass

$$m_k = m_{0k}(1 - v_k^2/c^2)^{-1/2}. \tag{1.70}$$

The corresponding expression for the angular momentum is worked out in the homework:

$$\mathbf{J} = \sum_k \mathbf{r}_k \times m_k \mathbf{v}_k + \frac{1}{c}\int (d\mathbf{r})\, \mathbf{r} \times (\mathbf{E} \times \mathbf{B}). \tag{1.71}$$

1.3 Energy-Momentum Tensor

From this point on, we will adopt natural units, $c = 1$ (and when we move to quantum mechanics, $\hbar = 1$). These quantities have no meaning outside the particular scheme of choosing units, and have defined values in SI, so might as well be defined to be unity. $\hbar c$ is a particularly convenient unit conversion factor,

$$\hbar c = 1.97 \times 10^{-5}\mathrm{eV\,cm} = 197\mathrm{MeV\,fm}. \tag{1.72}$$

Now, let us consider how fields transform under four-dimensional (space-time) coordinate transformations. For a scalar field, the field is the same at the same physical point, so

$$\bar\phi(\bar x) = \phi(x), \tag{1.73}$$

where for an infinitesimal transformation

$$\bar x^\mu = x^\mu + \delta x^\mu, \tag{1.74}$$

so expanding the field,

$$\begin{aligned}\bar\phi(x) &= \phi(x) + \delta\phi(x) = \phi(x - \delta x)\\ &= \phi(x) - \partial_\mu \phi(x) \delta x^\mu,\end{aligned} \tag{1.75}$$

or

$$\delta\phi = -\delta x^\mu \partial_\mu \phi. \tag{1.76}$$

Take the derivative of this:

$$\partial_\mu \delta\phi = \delta(\partial_\mu\phi) = -\delta x^\nu \partial_\mu \partial_\nu \phi - (\partial_\mu \delta x^\nu)\partial_\nu \phi. \tag{1.77}$$

We will take this to be the rule for how a vector field transforms:

$$\delta A_\mu = -\delta x^\nu \partial_\nu A_\mu - A_\nu \partial_\mu \delta x^\nu. \tag{1.78}$$

A check of this last result is provided by considering a rigid spatial translation,

$$\delta x_\mu = (0, \delta\mathbf{r}), \quad \partial_\nu \delta x_\mu = 0. \tag{1.79}$$

Then the rule (1.78) implies correctly [cf. Eq. (1.66)]

$$\delta A_\mu = -\delta\mathbf{r} \cdot \boldsymbol{\nabla} A_\mu. \tag{1.80}$$

For a rigid rotation

$$\delta x_\mu = (0, \delta\boldsymbol{\omega} \times \mathbf{r}), \tag{1.81}$$

so

$$\partial_i \delta x_j = \partial_i \epsilon_{jkl} \delta\omega_k x_l = \epsilon_{jki}\delta\omega_k, \tag{1.82}$$

the transformation of a three-vector field is

$$\delta\mathbf{A} = -(\delta\mathbf{r}\cdot\boldsymbol{\nabla})\mathbf{A} + \delta\boldsymbol{\omega}\times\mathbf{A}. \tag{1.83}$$

The last term here says that $\mathbf{A}$, like $\mathbf{r}$, is a vector.

A tensor transforms by the obvious generalization of the transformation law for a vector:

$$\delta F_{\mu\nu} = -\delta x^\lambda \partial_\lambda F_{\mu\nu} - (\partial_\mu \delta x^\lambda) F_{\lambda\nu} - (\partial_\nu \delta x^\lambda) F_{\mu\lambda}, \tag{1.84}$$

which is consistent with the result found in the homework.

Now let us calculate the change in the field part of the electrodynamic Lagrangian (1.36)

$$L_f = \int (d\mathbf{r}) \left[-\frac{1}{2} F^{\mu\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu) + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right], \tag{1.85}$$

or better, the change in the corresponding action

$$W = \int dt\, L, \tag{1.86}$$

where now we take the integration to be over all time. Then the Lagrange density $\mathcal{L}$ is defined by

$$W = \int (dx)\mathcal{L}, \quad (dx) = dt(d\mathbf{r}). \tag{1.87}$$

If we substitute the field strength construction $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, the change of the Lagrange density under a coordinate transformation is

$$\begin{aligned} \delta\mathcal{L} &= -\frac{1}{2}\delta F_{\mu\nu} F^{\mu\nu} = \frac{1}{2}\delta x^\lambda (\partial_\lambda F_{\mu\nu}) F^{\mu\nu} + (\partial_\mu \delta x^\lambda) F_{\lambda\nu} F^{\mu\nu} \\ &= -\delta x^\lambda \partial_\lambda \mathcal{L} + F^{\mu\lambda} F^\nu{}_\lambda \partial_\mu \delta x_\nu \\ &= -\partial_\lambda(\delta x^\lambda \mathcal{L}) + t^{\mu\nu}\partial_\mu \delta x_\nu, \end{aligned} \tag{1.88}$$

where the electromagnetic energy-momentum or stress tensor is

$$t^{\mu\nu} = F^{\mu\lambda} F^\nu{}_\lambda + g^{\mu\nu}\mathcal{L}, \quad \mathcal{L} = -\frac{1}{4} F^{\alpha\beta} F_{\alpha\beta}. \tag{1.89}$$

Notice that the energy-momentum tensor is symmetric,

$$t^{\mu\nu} = t^{\nu\mu}. \tag{1.90}$$

When the region we are considering contains no charges, $\delta W = 0$ by the stationary action principle,

$$0 = \delta W = \int (dx) t^{\mu\nu} \partial_\mu \delta x_\nu \tag{1.91}$$

up to a surface term, so since the variation δx_ν is arbitrary at every point in spacetime,

$$\partial_\mu t^{\mu\nu} = 0. \tag{1.92}$$

This conservation law, which is the local statement of energy-momentum conservation, may be directly verified using Maxwell's equations. How this is modified when currents are present is also given in the homework.

Let us examine the explicit components of $t^{\mu\nu}$. The time-time component is the energy density:

$$\begin{aligned} t^{00} &= F^{0i}F^0{}_i + \frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} \\ &= E^2 - \frac{1}{2}(E^2 - B^2) = \frac{1}{2}(E^2 + B^2). \end{aligned} \tag{1.93}$$

The time-space components are the momentum density,

$$t^{0i} = t^{i0} = F^{0j}F^i{}_j = E^j \epsilon^{ijk} B_k = (\mathbf{E}\times\mathbf{B})_i. \tag{1.94}$$

The stress tensor, which measures the flux of the ith component of momentum crossing a surface perpendicular to the jth direction, is

$$\begin{aligned} t^{ij} &= F^{i0}F^j{}_0 + F^{ik}F^{jk} + \delta^{jk}\left(-\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta}\right) \\ &= -E_iE_j + \epsilon^{ikl}\epsilon^{jkm}B_lB_m + \frac{1}{2}\delta_{ij}(E^2 - B^2). \end{aligned} \tag{1.95}$$

If we use the identity

$$\epsilon_{ikl}\epsilon_{jkm} = \delta_{ij}\delta_{lm} - \delta_{im}\delta_{lj}, \tag{1.96}$$

we can write the result in dyadic notation

$$\mathbf{t} = -\mathbf{EE} - \mathbf{BB} + \frac{1}{2}\mathbf{1}(E^2 + B^2), \tag{1.97}$$

which is the familiar 3-dimensional stress tensor.

1.3.1 Scale Invariance

It is of some significance that the Maxwell stress tensor is traceless:

$$t \equiv t^\lambda{}_\lambda = F^{\alpha\beta}F_{\alpha\beta} + 4\left(-\frac{1}{4}\right)F^{\alpha\beta}F_{\alpha\beta} = 0. \tag{1.98}$$

This reflects the *scale* invariance of the Maxwell theory.

A scale transformation is a particular kind of coordinate transformation,

$$\delta x^\mu = \delta a\, x^\mu. \tag{1.99}$$

Under such a transformation, the action changes by

$$\delta W = \int (dx) t^{\mu\nu}\partial_\mu \delta x_\nu = \int (dx) t^{\mu\nu}\left(\delta a\, g_{\mu\nu} + x_\nu \partial_\mu \delta a\right), \tag{1.100}$$

which, because $t = 0$, indeed vanishes if δa is constant, and, generally, by the action principle implies

$$\partial_\mu \left(x_\nu t^{\mu\nu} \right) = 0. \tag{1.101}$$

The conserved current here,

$$c^\mu = x_\nu t^{\mu\nu} \tag{1.102}$$

is called the scale current.