

10.1 The Lagrangian for a Charged Particle

Consider now the equation of motion for a (relativistic) charged particle in an externally applied electromagnetic field. We have

$$\begin{aligned}\frac{d\mathbf{p}}{dt} &= q[\mathbf{E} + \mathbf{u} \times \mathbf{B}] \\ \frac{dW}{dt} &= q\mathbf{u} \cdot \mathbf{E}\end{aligned}$$

which implies

$$\frac{dU^\alpha}{d\tau} = \frac{q}{m} F^{\alpha\beta} U_\beta.$$

(We have now used W for the energy of the particle to avoid confusion with the magnitude of the electric field). We would like to obtain this equation from the principle of least action. Consider first a non-relativistic particle with kinetic energy $T = \frac{1}{2}m\mathbf{u}^2$ in a potential $V(\mathbf{r})$. The Lagrangian is $L = T - V = \frac{1}{2}m\mathbf{u}^2 - V(\mathbf{r})$. We can think of this as being defined for *any* path $\mathbf{r}(t)$ with $\mathbf{u}(t) = \dot{\mathbf{r}}(t)$,

$$L = L[\mathbf{r}(t), \dot{\mathbf{r}}(t)],$$

and then may define the *action* for any path $\mathbf{r}(t)$ connecting some initial point $\mathbf{r}_1 = \mathbf{r}(t_1)$ to some final point $\mathbf{r}_2 = \mathbf{r}(t_2)$;

$$A = \int_{t_1}^{t_2} L[\mathbf{r}(t), \dot{\mathbf{r}}(t)] dt.$$

The action principle states that the action is stationary under variations of the path about the *actual* path followed by the particle in its motion from \mathbf{r}_1 at t_1 to \mathbf{r}_2 at t_2 . The proof is

$$\begin{aligned}\delta A &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{r}} \cdot \delta \mathbf{r} + \frac{\partial L}{\partial \dot{\mathbf{r}}} \cdot \delta \dot{\mathbf{r}} \right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{r}} \cdot \delta \mathbf{r} + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{r}}} \cdot \delta \mathbf{r} \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{r}}} \right) \cdot \delta \mathbf{r} \right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{r}} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{r}}} \right) \right] \cdot \delta \mathbf{r} dt\end{aligned}$$

where at the last step we use $\delta \mathbf{r} = 0$ at the end-points. The vanishing of the integral for arbitrary $\delta \mathbf{r}$ satisfying the end-point condition is now equivalent to the statement of the Lagrange equations

$$\frac{d}{dt} \mathbf{p} = \frac{\partial L}{\partial \mathbf{r}}$$

where

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}}.$$

These are indeed Newton's equations for the case considered, since \mathbf{p} is the momentum and $\partial L / \partial \mathbf{r} = -\nabla V$ is the force.

We now seek a relativistic generalisation. Since the condition $\delta A = 0$ must determine the same trajectory independent of the reference frame, we require that A should be a Lorentz scalar. But if

$$\begin{aligned}A &= \int_{t_1}^{t_2} L dt \\ &= \int_{\tau_1}^{\tau_2} L \frac{dt}{d\tau} d\tau\end{aligned}$$

is to be a scalar, it follows that $L \frac{dt}{d\tau} = \gamma L$ must be a scalar. For a *free* particle this must furthermore be independent of the position. But the only Lorentz scalar one can make out of the velocity alone is $U^\alpha U_\alpha = c^2$. So

$$\gamma L_{\text{free}} = \text{const.}$$

To get the correct non-relativistic limit, this has to be $-mc^2$, and then

$$L_{\text{free}} = \frac{-mc^2}{\gamma(\mathbf{u})} = -mc^2 \sqrt{1 - \frac{\mathbf{u}^2}{c^2}}.$$

This indeed gives

$$\mathbf{p} = \frac{\partial L_{\text{free}}}{\partial \mathbf{u}} = m\mathbf{u}\gamma(u)$$

and the equation of motion $\dot{\mathbf{p}} = 0$.

For a slowly moving charged particle, the interaction with the electromagnetic field introduces a term $V = q\Phi = qcA^0$ to the potential, and so leads to

$$L_{\text{int}}^{\text{NR}} = -qcA^0$$

as the expression for the non-relativistic approximation to the interaction part of the Lagrangian. To go to the relativistic case, we try to find a Lorentz scalar to which this is an approximation, and then write $\gamma L_{\text{int}} =$ that Lorentz scalar. It is not hard to see that this has to be $-qU_\mu A^\mu$, so that we are led to consider

$$\gamma L = -mc^2 - qU_\mu A^\mu,$$

or

$$L = -mc^2 \sqrt{1 - \frac{\mathbf{u}^2}{c^2}} - q(cA^0 - \mathbf{u} \cdot \mathbf{A}).$$

The momentum *canonically conjugate* to \mathbf{r} is

$$\mathbf{P} = \frac{\partial L}{\partial \mathbf{u}} = m\gamma\mathbf{u} + q\mathbf{A}.$$

So in terms of the *mechanical* momentum \mathbf{p} we have

$$\mathbf{p} = \mathbf{P} - q\mathbf{A}.$$

One may then check that the Lagrange equations

$$\frac{d}{dt}P^i = \frac{\partial L}{\partial r^i}$$

do in fact give the correct equations of motion:

$$\frac{d}{dt}P^i = -qc \frac{\partial A^0}{\partial r^i} + q \sum_j u^j \frac{\partial A^j}{\partial r^i},$$

so that

$$\dot{p}^i + q \left(\frac{\partial A^i}{\partial t} + \sum_j \frac{\partial A^i}{\partial r^j} \frac{\partial r^j}{\partial t} \right) = -q \frac{\partial \Phi}{\partial r^i} + q \sum_j u^j \frac{\partial A^j}{\partial r^i},$$

which gives

$$\begin{aligned} \dot{\mathbf{p}} &= q[\mathbf{E} + \mathbf{u} \times (\nabla \times \mathbf{A})] \\ &= q[\mathbf{E} + \mathbf{u} \times \mathbf{B}]. \end{aligned}$$

10.2 The Hamiltonian

The Hamiltonian is defined by

$$H = \mathbf{P} \cdot \mathbf{u} - L,$$

in which \mathbf{u} has to be eliminated in favour of \mathbf{P} . This leads to

$$H = (\mathbf{P} - q\mathbf{A}) \cdot \mathbf{u} + mc^2/\gamma + qcA^0$$

with

$$\mathbf{u} = \frac{\mathbf{P} - q\mathbf{A}}{m\gamma}$$

and

$$\gamma = \left[\frac{m^2 c^2 + (\mathbf{P} - q\mathbf{A})^2}{m^2 c^2} \right]^{\frac{1}{2}}$$

which leads directly to

$$\begin{aligned} H &= [(\mathbf{P} - q\mathbf{A})^2 c^2 + m^2 c^4]^{\frac{1}{2}} + qcA^0 \\ &= [\mathbf{p}^2 c^2 + m^2 c^4]^{\frac{1}{2}} + qcA^0 \end{aligned}$$

from which one sees that the introduction of the electromagnetic field leads to the changes

$$\begin{aligned} \mathbf{p} &\rightarrow \mathbf{P} = \mathbf{p} + q\mathbf{A} \\ p^0 &= H/c \rightarrow P^0 = p^0 + qA^0. \end{aligned}$$

10.3 The Lagrangian for the Electromagnetic Field

For a single particle moving in a potential $V(\mathbf{r})$, the equation of motion can be obtained from the principle of stationary action

$$\delta A = 0$$

where the action is

$$A = \int_{t_1}^{t_2} L dt$$

and the Lagrangian is $L[\mathbf{r}(t), \dot{\mathbf{r}}(t)] = T - V$, T being the kinetic energy expressed as a function of the particle velocity $\dot{\mathbf{r}}$. For the principle of stationary action gives

$$\frac{\delta A}{\delta \mathbf{r}(t)} \equiv \frac{\partial L}{\partial \mathbf{r}(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{r}}(t)} = 0,$$

which is equivalent to Newton's equation of motion. This generalises directly to a system of (finitely) many degrees of freedom, with generalised coordinates $q^i(t)$ and velocities $\dot{q}^i(t)$, when

$$L = L[q^i(t), \dot{q}^i(t)],$$

$$A = \int_{t_1}^{t_2} L[q^i, \dot{q}^i] dt,$$

$$\frac{\delta A}{\delta q^i(t)} \equiv \frac{\partial L}{\partial q^i(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i(t)} = 0.$$

Our first problem will be to understand how this may be further generalised to the situation of a *field* theory in which one has a continuous infinity of degrees of freedom (one – or a finite number – for every point in space). A way to approach this problem is suggested by a familiar model of a solid, in which we think of a ‘crystal lattice’ of atoms interacting with their nearest neighbours through some sort of elastic force, and for good measure also acted on by some other potential. For simplicity this can be taken to start with in

just one dimension. If we suppose that the displacement of the i th atom is ϕ_i , the potential energy, kinetic energy and Lagrangian are then

$$\begin{aligned} V &= \sum_i \left[\frac{1}{2} k (\phi_{i+1} - \phi_i)^2 + v(\phi_i) \right] \\ &= \sum_i \left[\frac{1}{2} k a^2 \left(\frac{\Delta \phi_i}{a} \right)^2 + v(\phi_i) \right]. \\ T &= \sum_i \frac{1}{2} m \dot{\phi}_i^2. \\ L &= T - V \\ &\rightarrow \int dx \left[\frac{1}{2} \frac{m}{a} \left(\frac{\partial \phi}{\partial t} \right)^2 - \frac{1}{2} k a \left(\frac{\partial \phi}{\partial x} \right)^2 - \frac{1}{a} v(\phi) \right]. \end{aligned}$$

At the last step we have suggested how to go to a continuum limit, where now we let $a \rightarrow 0$, keeping finite $\frac{m}{a}$, ka and $\frac{v}{a}$. This is in fact the appropriate way to model a continuous elastic medium. It also suggests that for a *field theory* we take as the Lagrangian an integral over what is called the *Lagrangian density*,

$$L = \int d^3x \mathcal{L} \left[\phi(\mathbf{x}, t), \nabla \phi, \frac{\partial \phi}{\partial t} \right].$$

We then have for the action

$$A = \int dt \int d^3x \mathcal{L}$$

and the principle of stationary action $\delta A = 0$ gives the *Euler-Lagrange field equation*:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) = \frac{\partial \mathcal{L}}{\partial \phi} - \nabla \cdot \left(\frac{\partial \mathcal{L}}{\partial \nabla \phi} \right),$$

which we can rewrite in a suggestive fashion by noting that since $\partial^\mu = (\frac{1}{c} \frac{\partial}{\partial t}, \nabla)$, with $\mathcal{L} = \mathcal{L}(\phi, \partial^\mu \phi)$ the Euler-Lagrange equation becomes

$$\partial^\mu \left\{ \frac{\partial \mathcal{L}}{\partial (\partial^\mu \phi)} \right\} = \frac{\partial \mathcal{L}}{\partial \phi}.$$

This then is the *field equation*. For example, if

$$\mathcal{L} = \frac{1}{2} (\partial^\mu \phi) (\partial_\mu \phi) - \frac{\kappa^2}{2} \phi^2,$$

as might be suggested by our model of an elastic medium, the field equation is

$$\partial^\mu (\partial_\mu \phi) = -\kappa^2 \phi$$

or

$$\square \phi = -\kappa^2 \phi,$$

which is the *Klein-Gordon equation*. Note that if ϕ is a Lorentz scalar, \mathcal{L} is also a Lorentz scalar, and the resulting field equation is Lorentz invariant. Conversely, to get a Lorentz invariant field equation, we must start from a Lorentz invariant *action*, and this means in turn that \mathcal{L} must be a Lorentz scalar.

Let us turn now to the electromagnetic field. The field variable will then not be the scalar field ϕ of our previous example, but will be the 4-vector A^α . The Lagrangian density depends on the derivatives of the fields, so instead of $\partial^\mu \phi$ it will involve $\partial^\beta A^\alpha$. In the previous example we had introduced $(\partial^\mu \phi) (\partial_\mu \phi)$ as a scalar, quadratic in the field gradients. On general grounds we expect that there should be a term in the Lagrangian of this form, and so are led to seek a Lorentz scalar quadratic in the derivatives $\partial^\beta A^\alpha$. Just as a Lorentz scalar Lagrangian density leads to Lorentz invariance of the field equations, so also *gauge* invariance

of the Lagrangian density will lead to gauge invariance of the field equations. So we need to find a gauge invariant Lorentz scalar quadratic in the field derivatives. We have already encountered two such, namely $F^{\alpha\beta}F_{\alpha\beta}$ and $F^{\alpha\beta} * F_{\alpha\beta}$; but the latter of these is in fact only a *pseudo*-scalar. This motivates the choice

$$\mathcal{L} = \text{const } F^{\alpha\beta}F_{\alpha\beta}$$

for the electromagnetic field in the absence of sources j^μ . If sources are present we expect to have to subtract from this the potential energy of the interaction between the sources and the electromagnetic field, so have

$$\mathcal{L} = \text{const } F^{\alpha\beta}F_{\alpha\beta} - j^\mu A_\mu,$$

(the interaction term being just the generalisation to a general source of what we have already discussed for the interaction of a single charged particle with the electromagnetic field) wherein $F_{\alpha\beta} \equiv \partial_\alpha A_\beta - \partial_\beta A_\alpha$. It might be objected that the interaction term is *not* gauge invariant; indeed under a gauge transformation it changes by $j^\mu \partial_\mu \chi$. But this differs from the total divergence $\partial_\mu(j^\mu \chi)$ (which makes no contribution to the field equations anyway, and so can be dropped) by $\chi \partial_\mu j^\mu$, which vanishes *because the current is conserved*. There is thus illustrated the very important and deep connection between the gauge invariance of the theory and the conservation law.

The Euler-Lagrange equations which follow from the choice $\mathcal{L} = kF^{\alpha\beta}F_{\alpha\beta} - j^\mu A_\mu$ are

$$\partial_\beta \left(\frac{\partial \mathcal{L}}{\partial A_{\alpha,\beta}} \right) = \frac{\partial \mathcal{L}}{\partial A_\alpha} = -j^\alpha.$$

(We have introduced the very convenient notation $f_{,\beta}$ for $\partial_\beta f = \frac{\partial f}{\partial x^\beta}$, so that $A_{\alpha,\beta}$ means $\partial_\beta A_\alpha = \frac{\partial A_\alpha}{\partial x^\beta}$). But

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial A_{\alpha,\beta}} &= k \frac{\partial}{\partial A_{\alpha,\beta}} (F_{\rho\sigma} \eta^{\rho\mu} \eta^{\sigma\nu} F_{\mu\nu}) \\ &= 2k F^{\mu\nu} \frac{\partial}{\partial A_{\alpha,\beta}} (A_{\nu,\mu} - A_{\mu,\nu}) \\ &= 2k F^{\mu\nu} (\delta_\nu^\alpha \delta_\mu^\beta - \delta_\mu^\alpha \delta_\nu^\beta) \\ &= -4k F^{\alpha\beta}, \end{aligned}$$

so we have

$$-4k F^{\alpha\beta}{}_{,\beta} = -j^\alpha,$$

which is to be compared with the field equation

$$F^{\alpha\beta}{}_{,\beta} = -\mu_0 j^\alpha.$$

This fixes the normalisation constant $k = -\frac{1}{4\mu_0}$. The Lagrangian density can also be written as

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4\mu_0} F^{\alpha\beta}F_{\alpha\beta} - j^\mu A_\mu \\ &= -\frac{1}{2\mu_0} \left(\mathbf{B}^2 - \frac{\mathbf{E}^2}{c^2} \right) - j^\mu A_\mu \\ &= \frac{1}{2} \left(\epsilon_0 \mathbf{E}^2 - \frac{1}{\mu_0} \mathbf{B}^2 \right) - j^\mu A_\mu. \end{aligned}$$

10.4 The Hamiltonian for the Electromagnetic Field

The Hamiltonian (which generalises $H = \sum_i p_i \dot{q}^i - L$) is

$$H = \int d^3x \frac{\partial \mathcal{L}}{\partial \dot{A}^\mu} \dot{A}^\mu - L$$

since the momentum conjugate to A^μ is $\frac{\partial \mathcal{L}}{\partial A^\mu}$. Note that since the time derivative of the component A^0 does not enter into the Lagrangian, the momentum conjugate to A^0 vanishes identically; this is a reflection of the fact that not all of the four components of A^μ are independent, and only three need be determined by the equations of motion, with A^0 being given from the gauge condition, for example $\partial_\mu A^\mu = 0$. The momentum conjugate to A^i , (where $i = 1, 2, 3$) is $\frac{\partial \mathcal{L}}{\partial A^i} = -\epsilon_0 E^i$, and we then find

$$\begin{aligned} H &= - \int d^3x \epsilon_0 \mathbf{E} \cdot \dot{\mathbf{A}} - L \\ &= \int d^3x \{ \epsilon_0 [\mathbf{E} \cdot (\mathbf{E} + \nabla \Phi)] - \frac{1}{2} \epsilon_0 [\mathbf{E}^2 - c^2 \mathbf{B}^2] + \rho \Phi - \mathbf{A} \cdot \mathbf{J} \} \\ &= \int d^3x \{ \frac{1}{2} \epsilon_0 [\mathbf{E}^2 + c^2 \mathbf{B}^2] - \mathbf{A} \cdot \mathbf{J} - \Phi [\nabla \cdot (\epsilon_0 \mathbf{E}) - \rho] + \nabla \cdot (\epsilon_0 \mathbf{E} \Phi) \}. \end{aligned}$$

The last term is a divergence which makes no contribution to the (Hamilton) equations of motion. The only other place where Φ occurs is in the penultimate term, and the corresponding Hamilton equation simply ensures the vanishing of the expression which multiplies Φ , namely the Maxwell equation $\nabla \cdot (\epsilon_0 \mathbf{E}) - \rho = 0$. If this equation is imposed as a *constraint*, Φ may be eliminated altogether, and we have

$$H = \int d^3x \{ \frac{1}{2} \epsilon_0 [\mathbf{E}^2 + c^2 \mathbf{B}^2] - \mathbf{A} \cdot \mathbf{J} \}.$$

In these remaining terms $\mathbf{B} = \nabla \times \mathbf{A}$, and the field variable is \mathbf{A} , with conjugate momentum as previously given, $-\epsilon_0 \mathbf{E}$. So the Hamiltonian has been expressed entirely in terms of the fields and the conjugate momenta as is required. Hamilton's equations, the analogues of $\dot{p} = -\frac{\partial H}{\partial q}$, $\dot{q} = \frac{\partial H}{\partial p}$, are

$$-\epsilon_0 \dot{\mathbf{E}} = \mathbf{J} - c^2 \epsilon_0 \nabla \times \mathbf{B},$$

$$\dot{\mathbf{A}} = -\mathbf{E} - \nabla \Phi.$$

And these are again Maxwell's equations. [You might notice that the constraint equation $\nabla \cdot (\epsilon_0 \mathbf{E}) - \rho = 0$ is consistent with the first of these because the sources satisfy the conservation equation $\dot{\rho} + \nabla \cdot \mathbf{J} = 0$, as we have already remarked]

10.5 The Canonical Stress Tensor

We have introduced the Lagrangian density

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4\mu_0} F^{\alpha\beta} F_{\alpha\beta} - j^\mu A_\mu \\ &= \frac{1}{2} \epsilon_0 (\mathbf{E}^2 - c^2 \mathbf{B}^2) - j^\mu A_\mu \end{aligned}$$

with

$$L = \int \mathcal{L} d^3x \quad A = \int L dt = \frac{1}{c} \int \mathcal{L} d^4x.$$

Similarly we have

$$H = \int \mathcal{H} d^3x$$

with

$$\begin{aligned} \mathcal{H} &= \frac{\partial \mathcal{L}}{\partial A_{\mu,0}} A_{\mu,0} - \mathcal{L} \\ &= \frac{1}{2} \epsilon_0 (\mathbf{E}^2 + c^2 \mathbf{B}^2) + \nabla \cdot (\epsilon_0 \Phi \mathbf{E}) - \mathbf{j} \cdot \mathbf{A}. \end{aligned}$$

This suggests the definition

$$\begin{aligned} T^\nu{}_\lambda &= \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} A_{\mu,\lambda} - \delta_\lambda^\nu \mathcal{L} \\ &= \frac{1}{\mu_0} F^{\mu\nu} A_{\mu,\lambda} - \delta_\lambda^\nu \mathcal{L}, \end{aligned}$$

which is a *tensor* of which \mathcal{H} is a component. This tensor is called the *canonical stress tensor*. We have

$$T^{0\lambda} = (\mathcal{H}, \mathbf{\Pi}),$$

with $T^{00} = \mathcal{H}$ as given above, and

$$\Pi^i = T^{0i} = \frac{1}{\mu_0 c} \{ (\mathbf{E} \times \mathbf{B})^i + \nabla \cdot (\mathbf{E} A^i) - \frac{\rho}{\epsilon_0} A^i \}.$$

Let us recall that the energy density in the electromagnetic field is

$$\begin{aligned} u &= \frac{1}{2} (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}) \\ &= \frac{1}{2} \epsilon_0 (\mathbf{E}^2 + c^2 \mathbf{B}^2) \end{aligned}$$

and the Poynting vector, which is the flux vector for electromagnetic energy, is

$$\begin{aligned} \mathbf{S} &= \mathbf{E} \times \mathbf{H} \\ &= \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}. \end{aligned}$$

Then

$$\begin{aligned} \Pi^i &= \frac{1}{c} S^i + \nabla \cdot (c \epsilon_0 \mathbf{E} A^i) - j^0 A^i, \\ \mathcal{H} &= u + \nabla \cdot (c \epsilon_0 \mathbf{E} A^0) - j^0 A^0 + j^\mu A_\mu. \end{aligned}$$

Suppose now that the electromagnetic fields are localised in some region, and that there are no sources. Then integrating over that region

$$\begin{aligned} \int T^{00} d^3x &= \int [u + \nabla \cdot (c \epsilon_0 \mathbf{E} A^0)] d^3x \\ &= \int u d^3x \\ &= U_{\text{field}}; \\ \int T^{0i} d^3x &= \int \left[\frac{1}{c} S^i + \nabla \cdot (c \epsilon_0 \mathbf{E} A^i) \right] d^3x \\ &= \frac{1}{c} \int S^i d^3x \\ &= c P_{\text{field}}^i. \end{aligned}$$

Here U_{field} and $\mathbf{P}_{\text{field}}$ are respectively the total energy and momentum of the electromagnetic field in the region considered.

Recall the conservation law

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} = -\mathbf{j} \cdot \mathbf{E}.$$

Since u and \mathbf{S} are *not* components of a 4-vector, this is not in itself a covariant conservation law; but it does indicate how we might find a covariant extension of it. We consider

$$\begin{aligned} \partial_\alpha T^{\alpha\beta} &= \partial_\alpha \left\{ \frac{\partial \mathcal{L}}{\partial (\partial_\alpha A_\mu)} \partial^\beta A_\mu - \eta^{\alpha\beta} \mathcal{L} \right\} \\ &= \partial_\alpha \left[\frac{\partial \mathcal{L}}{\partial (\partial_\alpha A_\mu)} \right] \partial^\beta A_\mu + \frac{\partial \mathcal{L}}{\partial (\partial_\alpha A_\mu)} \partial_\alpha \partial^\beta A_\mu - \partial^\beta \mathcal{L}. \end{aligned}$$

Now use the Euler-Lagrange equations

$$\partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A_\mu)} = \frac{\partial \mathcal{L}}{\partial A_\mu},$$

to obtain

$$\partial_\alpha T^{\alpha\beta} = \frac{\partial \mathcal{L}}{\partial A_\mu} \partial^\beta A_\mu + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A_\mu)} \partial^\beta (\partial_\alpha A_\mu) - \partial^\beta \mathcal{L}$$

which vanishes identically (use the chain rule for differentiation). Thus the covariant conservation law

$$\partial_\alpha T^{\alpha\beta} = 0$$

is a consequence of the equation of motion.

If this is integrated over the region containing the fields,

$$\begin{aligned} \int \partial_\alpha T^{\alpha\beta} d^3x &= 0 \\ &= \int (\partial_0 T^{0\beta} + \partial_i T^{i\beta}) d^3x \\ &= \partial_0 \int T^{0\beta} d^3x + \text{surface term.} \end{aligned}$$

The surface term vanishes (we have supposed the fields to be localised in some region), so what we have obtained is

$$\begin{aligned} \frac{d}{dt} U_{\text{field}} &= 0, \\ \frac{d}{dt} \mathbf{P}_{\text{field}} &= 0, \end{aligned}$$

the conservation of the energy and of the momentum for localised fields in the absence of sources.

10.6 The Symmetric Stress Tensor

This is very pretty, but there are a number of objections:

(i) The electromagnetic energy and momentum ought to be covariantly defined as parts of a 4-vector; we have implicitly been using a frame in which the observer is at rest.

(ii) \mathcal{H} and $\mathbf{\Pi}$ differ from u and \mathbf{S} even in the absence of sources (albeit by a divergence).

(iii) The tensor $T^{\alpha\beta}$ is not symmetrical. The significance of this arises from the wish to incorporate into a covariant conservation law the conservation of the *angular momentum* of the electromagnetic field. Thus

$$\begin{aligned} \mathbf{M}_{\text{field}} &= \frac{1}{c} \int \mathbf{x} \times (\mathbf{E} \times \mathbf{H}) d^3x \\ &= \frac{1}{c\mu_0} \int \mathbf{x} \times (\mathbf{E} \times \mathbf{B}) d^3x \end{aligned}$$

is conserved, and a local covariant generalisation of this would be

$$\partial_\alpha M^{\alpha\beta\gamma} = 0$$

with

$$M^{\alpha\beta\gamma} = T^{\alpha\beta} x^\gamma - T^{\alpha\gamma} x^\beta.$$

This requires

$$\partial_\alpha (T^{\alpha\beta} x^\gamma - T^{\alpha\gamma} x^\beta) = 0,$$

i.e.,

$$(\partial_\alpha T^{\alpha\beta})x^\gamma + T^{\alpha\beta}\delta_\alpha^\gamma - (\partial_\alpha T^{\alpha\gamma})x^\beta - T^{\alpha\gamma}\delta_\alpha^\beta = 0,$$

or using the previous result

$$T^{\gamma\beta} - T^{\beta\gamma} = 0.$$

(iv) On general grounds it is expected that $T^\alpha_\alpha = 0$. This is related to the fact that the quanta of the electromagnetic field, the photons, have zero mass, and likewise to the scale-invariance of electromagnetism — these are technical points outside the scope of this course.

(v) If $T^{\alpha\beta}$ is to be of direct physical significance, it ought to be gauge invariant.

We seek to remedy these defects by modifying the tensor:

$$T^{\alpha\beta} \rightarrow \Theta^{\alpha\beta} = T^{\alpha\beta} - T_D^{\alpha\beta}$$

in such a way that

- a) $\Theta^{\alpha\beta} = \Theta^{\beta\alpha}$ Θ is symmetric
- b) $\Theta^\alpha_\alpha = 0$ Θ is traceless
- c) Θ is gauge-invariant
- d) $\partial_\alpha \Theta^{\alpha\beta} = 0$ Θ is conserved

but

$$e) \quad \int \Theta^{0\beta} d^3x = \int T^{0\beta} d^3x.$$

In the absence of sources, we had

$$\begin{aligned} T^\nu_\lambda &= \frac{1}{\mu_0} F^{\mu\nu} A_{\mu,\lambda} - \delta_\lambda^\nu \mathcal{L} \\ &= \frac{1}{\mu_0} \{ F^{\mu\nu} (A_{\mu,\lambda} - A_{\lambda,\mu}) + (F^{\mu\nu} A_\lambda)_{,\mu} - F^{\mu\nu}{}_{,\mu} A_\lambda \} - \delta_\lambda^\nu \mathcal{L} \\ &= -\frac{1}{\mu_0} F^{\mu\nu} F_{\mu\lambda} - \delta_\lambda^\nu \mathcal{L} + \frac{1}{\mu_0} (F^{\mu\nu} A_\lambda)_{,\mu}. \end{aligned}$$

So let us define

$$T_D^\nu{}_\lambda = -\frac{1}{\mu_0} (F^{\mu\nu} A_\lambda)_{,\mu}.$$

We then have

$$\Theta^{\nu\lambda} = -\frac{1}{\mu_0} [F^{\mu\nu} F^{\sigma\lambda} \eta_{\mu\sigma} - \frac{1}{4} \eta^{\nu\lambda} F^{\alpha\beta} F_{\alpha\beta}]$$

which is clearly

- a) symmetrical
- b) traceless (use $\delta^\alpha_\alpha = 4$)
- c) gauge invariant, since it depends only on $F^{\mu\nu}$
- d) conserved in the absence of sources, since

$$\begin{aligned} T_D^\nu{}_{\lambda,\nu} &= \partial_\nu \left[-\frac{1}{\mu_0} (F^{\mu\nu} A_\lambda)_{,\mu} \right] \\ &= -\frac{1}{\mu_0} (F^{\mu\nu} A_\lambda)_{,\mu,\nu}, \end{aligned}$$

which vanishes because $F^{\mu\nu}$ is antisymmetrical in μ, ν whilst the partial derivatives are symmetrical. Thus

$$\Theta^\nu{}_{\lambda,\nu} = T^\nu{}_{\lambda,\nu} = 0.$$

and

e) for localised fields, the space integrals of $\Theta^{0\beta}$ and $T^{0\beta}$ are equal, since

$$\begin{aligned} \int \Theta^{0\beta} d^3x - \int T^{0\beta} d^3x &= \int T_D^{0\beta} d^3x \\ &= -\frac{1}{\mu_0} \int (F^{\mu 0} A^\beta)_{,\mu} d^3x \\ &= -\frac{1}{\mu_0} \int (F^{i0} A^\beta)_{,i} d^3x \\ &= -\frac{1}{\mu_0 c} \int \nabla \cdot (\mathbf{E} A^\beta) d^3x \end{aligned}$$

which gives a surface integral which vanishes because $\mathbf{E} = 0$ on the boundary, the fields being localised.

We are thus motivated to define the *symmetric stress tensor* $\Theta^{\alpha\beta}$, even when there are sources present, by

$$\Theta^{\alpha\beta} \equiv -\frac{1}{\mu_0} [F^{\lambda\alpha} F_{\lambda}{}^\beta - \frac{1}{4} \eta^{\alpha\beta} F^{\mu\nu} F_{\mu\nu}]$$

for which

$$\begin{aligned} \Theta^{00} &= \frac{1}{2} \epsilon_0 (\mathbf{E}^2 + c^2 \mathbf{B}^2) = u \\ \Theta^{0i} &= \frac{1}{c \mu_0} (\mathbf{E} \times \mathbf{B})^i = \frac{1}{c} S^i = c g^i \\ \Theta^{ij} &= -\epsilon_0 \{ E^i E^j + c^2 B^i B^j - \frac{1}{2} \delta^{ij} (\mathbf{E}^2 + c^2 \mathbf{B}^2) \} \\ &= -[\text{the Maxwell stress tensor}]. \end{aligned}$$

This tensor thus combines the energy density u , the Poynting vector \mathbf{S} (or the momentum density vector \mathbf{g}), and the Maxwell stress tensor (a *three*-tensor which gives the mechanical stress present in the electromagnetic field, which is responsible for the ‘repulsion of lines of force’ described by Faraday) into a Lorentz covariant tensor.

10.7 The Conservation Laws

To see what happens to the conservation law in the presence of sources, consider

$$\begin{aligned} \partial_\alpha \Theta^{\alpha\beta} &= -\frac{1}{\mu_0} [F^{\lambda\alpha}{}_{,\alpha} F_{\lambda}{}^\beta + F^{\lambda\alpha} F_{\lambda}{}^\beta{}_{,\alpha} - \frac{1}{2} F^{\lambda\mu} F_{\lambda\mu,}{}^\beta] \\ &= j^\lambda F_{\lambda}{}^\beta - \frac{1}{2\mu_0} [2F^{\lambda\alpha} F_{\lambda}{}^\beta{}_{,\alpha} - F^{\lambda\mu} F_{\lambda\mu,}{}^\beta]. \end{aligned}$$

Thus

$$\begin{aligned} \partial_\alpha \Theta^{\alpha\beta} - j^\lambda F_{\lambda}{}^\beta &= -\frac{1}{2\mu_0} [2F_{\lambda\mu} F^{\lambda\beta, \mu} - F_{\lambda\mu} F^{\lambda\mu, \beta}] \\ &= -\frac{1}{2\mu_0} F_{\lambda\mu} [F^{\lambda\beta, \mu} - F^{\mu\beta, \lambda} - F^{\lambda\mu, \beta}] \\ &= \frac{1}{2\mu_0} F_{\lambda\mu} [F^{\beta\lambda, \mu} + F^{\mu\beta, \lambda} + F^{\lambda\mu, \beta}] \\ &= 0, \end{aligned}$$

where we make repeated use of the antisymmetry of $F_{\mu\nu}$ and at the last stage of Maxwell's homogeneous equation.

Thus we have derived

$$\partial_\alpha \Theta^{\alpha\beta} = j^\lambda F_{\lambda}{}^\beta \equiv -f^\beta,$$

where we have introduced the *Lorentz force density* $f^\beta = F^{\beta\lambda} j_\lambda = (f^0, \mathbf{f})$. We have

$$f^i = F^{i\lambda} j_\lambda = F^{i0} j^0 - \sum_k F^{ik} j^k = E^i \rho + (\mathbf{j} \times \mathbf{B})^i,$$

or

$$\mathbf{f} = \rho \mathbf{E} + \mathbf{j} \times \mathbf{B};$$

and

$$f^0 = F^{0\lambda} j_\lambda = F^{0i} j_i = \left(-\frac{\mathbf{E}}{c}\right) \cdot (-\mathbf{j}) = \frac{\mathbf{E} \cdot \mathbf{j}}{c}.$$

These should be compared with

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

and

$$\mathbf{F} \cdot \mathbf{v} = \mathbf{E} \cdot (q\mathbf{v});$$

respectively the Lorentz force (i.e., the rate of change of momentum) of a charged particle and the rate at which the field does work (i.e., the rate of change of the energy) on a charged particle. Thus the 4-vector f^β gives the rate of change of the energy and the momentum of the *sources*. In other words

$$\int f^\beta d^3x = \frac{d}{dt} P_{\text{matter}}^\beta.$$

What this means is that

$$\begin{aligned} \int (\partial_\alpha \Theta^{\alpha\beta} + f^\beta) d^3x &= 0 \\ &= \int (\partial_0 \Theta^{0\beta} + \partial_i \Theta^{i\beta}) d^3x + \frac{d}{dt} P_{\text{matter}}^\beta \\ &= \int \partial_0 \Theta^{0\beta} d^3x + \int \Theta^{i\beta} n_i d^2S + \frac{d}{dt} P_{\text{matter}}^\beta \\ &= \frac{d}{dt} [P_{\text{field}}^\beta + P_{\text{matter}}^\beta] + \text{surface term}, \end{aligned}$$

where

$$\begin{aligned} P_{\text{field}}^\beta &= \int \frac{1}{c} \Theta^{0\beta} d^3x \\ &= \int \left(\frac{u}{c}, \mathbf{g}\right) d^3x \\ &= \int \left(\frac{1}{c}(\text{energy density}), \text{momentum density}\right) d^3x, \end{aligned}$$

and the surface term vanishes for a closed system. At the differential level, the conservation equation gives (with $\beta = 0$)

$$\frac{1}{c} \left(\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} \right) = -\frac{\mathbf{E} \cdot \mathbf{j}}{c},$$

which is Poynting's equation; and (with $\beta = i$)

$$\frac{\partial g^i}{\partial t} = (T^{\text{Maxwell}})^{ij}{}_{,j} - (\rho \mathbf{E} + \mathbf{j} \times \mathbf{B})^i,$$

which expresses the fact that the rate of change of the density of momentum equals a contribution from the Maxwell stress exerted by the neighbouring field and a term which is the ‘equal and opposite’ reaction to the force exerted *by* the field *on* the sources.

10.8 The Field as an Ensemble of Oscillators

Consider an enclosure filled with electromagnetic radiation, but devoid of sources. The Hamiltonian is

$$H = \frac{1}{2}\epsilon_0 \int d^3x [\mathbf{E}^2 + c^2(\nabla \times \mathbf{A})^2].$$

In a radiation gauge ($\text{div}\mathbf{A} = 0, \Phi = 0$), we have

$$H = \frac{1}{2}\epsilon_0 \int d^3x [\dot{\mathbf{A}}^2 + c^2(\nabla \times \mathbf{A})^2].$$

We may analyse the field into a superposition of *normal modes*

$$\mathbf{A}(\mathbf{x}, t) = \sum_{\lambda} \frac{1}{\sqrt{\epsilon_0}} q_{\lambda}(t) \mathbf{A}_{\lambda}(\mathbf{x}).$$

The field equation $\square\mathbf{A} = 0$ implies for the normal mode potentials

$$\nabla^2 \mathbf{A}_{\lambda} + \frac{\omega_{\lambda}^2}{c^2} \mathbf{A}_{\lambda} = 0,$$

whilst the amplitudes q_{λ} satisfy

$$\ddot{q}_{\lambda}(t) + \omega_{\lambda}^2 q_{\lambda}(t) = 0;$$

the normal mode frequencies ω_{λ} are as usual constants of separation of the t - from the \mathbf{x} - variables. The gauge choice implies that $\text{div}\mathbf{A}_{\lambda} = 0$, and the factor $\frac{1}{\sqrt{\epsilon_0}}$ in the normal mode expansion is chosen for later convenience. The functions $\mathbf{A}_{\lambda}(\mathbf{x})$ can be chosen to satisfy the orthonormality condition

$$\int d^3x \mathbf{A}_{\lambda}(\mathbf{x}) \cdot \mathbf{A}_{\mu}(\mathbf{x}) = \delta_{\lambda\mu}.$$

In terms of the normal modes, we have

$$\begin{aligned} H = \frac{1}{2}\epsilon_0 \int d^3x & \left[\left(\sum_{\lambda} \frac{1}{\sqrt{\epsilon_0}} \dot{q}_{\lambda} \mathbf{A}_{\lambda} \right) \cdot \left(\sum_{\mu} \frac{1}{\sqrt{\epsilon_0}} \dot{q}_{\mu} \mathbf{A}_{\mu} \right) \right. \\ & \left. + c^2 \left(\sum_{\lambda} \frac{1}{\sqrt{\epsilon_0}} q_{\lambda} \nabla \times \mathbf{A}_{\lambda} \right) \cdot \left(\sum_{\mu} \frac{1}{\sqrt{\epsilon_0}} q_{\mu} \nabla \times \mathbf{A}_{\mu} \right) \right]. \end{aligned}$$

The last term may be manipulated as follows:

$$\begin{aligned} (\nabla \times \mathbf{A}_{\lambda}) \cdot (\nabla \times \mathbf{A}_{\mu}) &= \mathbf{A}_{\lambda} \cdot (\nabla \times (\nabla \times \mathbf{A}_{\mu})) + \text{a divergence which integrates to zero} \\ &= \mathbf{A}_{\lambda} \cdot [\nabla(\nabla \cdot \mathbf{A}_{\mu}) - \nabla^2 \mathbf{A}_{\mu}] \\ &= \mathbf{A}_{\lambda} \cdot (-\nabla^2 \mathbf{A}_{\mu}) \quad \text{since } \text{div } \mathbf{A}_{\mu} = 0 \\ &= \mathbf{A}_{\lambda} \cdot \frac{\omega_{\mu}^2}{c^2} \mathbf{A}_{\mu}. \end{aligned}$$

Putting this into the previous expression for H gives $H = \frac{1}{2} \sum_{\mu} [\dot{q}_{\mu}^2 + \omega_{\mu}^2 q_{\mu}^2]$, or better to write

$$H = \frac{1}{2} \sum_{\mu} [p_{\mu}^2 + \omega_{\mu}^2 q_{\mu}^2],$$

which is immediately recognised as the Hamiltonian for a system of simple harmonic oscillators. This was a result familiar to Planck in 1900!