Physics 504, Lecture 18
April 4, 2011

We now begin with Chapter 12, on the relativistic dynamics of charged particles in interaction with electromagnetic fields. I am permuting the order in Jackson and will consider sections 2 and 3 before section 1, and we will skip section 4.

1 Motion of charged particles in fixed external fields

There are many applications which require an analysis of how charged particles move in static external (macroscopic) electromagnetic fields. These include the bending of beams to make circular accelerators for nuclear and particle physics, understanding plasmas in deep space and in attempts to make fusion energy devices, designing velocity and momentum separators for beams of particles, understanding van Allen belts which cause the auroras, and many other applications. Of course, all of these analyses are based on the general formula $dp^\alpha/d\tau = (q/c)F^{\alpha\beta}U_\beta$, or in nonrelativistic language,

$$\frac{d\vec{p}}{dt} = q\left(\vec{E} + \frac{1}{c}\vec{v}\times\vec{B}\right), \quad \frac{dE}{dt} = q\vec{v}\cdot\vec{E}.$$

1.1 Constant Uniform $\vec{B}$ Only

First let us consider a uniform constant $\vec{B}$ with no electric field. The energy is constant, so so are $|v|$ and γ, and we have

$$\frac{d\vec{v}}{dt} = \frac{1}{\gamma m}\frac{d\vec{p}}{dt} = \frac{q}{\gamma mc}\vec{v}\times\vec{B} = \vec{v}\times\vec{\omega}_B, \quad \text{where } \vec{\omega}_B = \frac{q}{\gamma mc}\vec{B} = \frac{qc\vec{B}}{E}.$$

Thus the component of $\vec{v}$ parallel to $\vec{B}$ is a constant, and the other two components rotate counterclockwise about the $\vec{B}$ direction if the charge is positive. The position component parallel to $\vec{B}$ grows linearly with time, while the motion transverse to that is in a circle with angular velocity ω_B. The radius a of this circle is determined from $v_\perp = \omega_B a$, so

$$a = \frac{v_\perp}{\omega_B} = \frac{p_\perp}{m\gamma} \Big/ \frac{qB}{\gamma mc} = \frac{p_\perp c}{qB}.$$

This can be used to determine the momentum of a particle by measuring the radius of curvature in a magnetic field, and has been used in particle detectors at all high energy accelerators ever since the field began. It is also the formula which tells us how strong the magnetics must be at the LHC at CERN to get 7 TeV protons to bend in a circle with a circumference of only 27 km.[1]

The fact that the frequency of revolution $\omega_B/2\pi = qB/2\pi mc\gamma$ is nearly constant independent of v as long as the particle is nonrelativistic ($\gamma \approx 1$) makes possible the continuous acceleration of particles in a cyclotron[2], which consists of two "D" shaped cavities with an oscillating voltage between them, with a constant magnetic field normal to their plane which provides the bending to make particles go in semicircles of ever larger radius, each in half a period of the oscillating voltage, until they reach the outside of the cyclotron and come out in a "high energy" beam. The first cyclotron was built by Lawrence (not Lorentz) in 1929, was 4 inches in diameter. Three years later an 11 inch one set the high energy record for protons at 1 MeV. Lawrence kept building bigger and bigger machines, ignoring warnings that $\gamma = 1$ was not an exact statement, so to get the 184 inch cyclotron to work, particles had to be accelerated in bunches with the frequency gradually decreased in synchonicity with the increasing particle energies and γ's.

1.2 Constant Uniform $\vec{E}$ and $\vec{B}$

Next. let's consider adding a fixed uniform electric field to our magnetic one. The electric field does do work on the particle, so we can no longer assume $|v|$ and γ are constant, and the situation is considerably more complicated. If the electric field is perpendicular to the magnetic field, however, we can use a Lorentz transformation to a more suitable frame to help. If we take $\vec{E}$ to

[1]With $B = P_\perp c/qR$ in gaussian units, but $B = P_\perp/qR$ in SI units. As $P_\perp \approx E/c$ and $E/q = 7 \times 10^{12}$ V, $R = 4300$ m, $B = 5.4$ T. Unfortunately the 1232 dipole magnets, each 14.3 m long, do not cover the whole circumference, but only 17.6 km, so the magnets need to be 8.3 T, which is considerably harder to maintain.

[2]A nice web page is `http://www.aip.org/history/lawrence/epa.htm`

be in the y direction and $\vec{B}$ in the z, we can apply a Lorentz transformation in the x direction, with $u_x = c\tanh\zeta$. Then the Lorentz transformation $A^\mu{}_\nu$ and the primed and unprimed $F^{\mu\nu}$ are

$$A^\mu{}_\nu = \begin{pmatrix} \cosh\zeta & \sinh\zeta & 0 & 0 \\ \sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$F^{\mu\nu} = \begin{pmatrix} 0 & 0 & -E_y & 0 \\ 0 & 0 & -B_z & 0 \\ E_y & B_z & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to F'^{\mu\nu} = \begin{pmatrix} 0 & 0 & -E'_y & 0 \\ 0 & 0 & -B'_z & 0 \\ E'_y & B'_z & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

with

$$\begin{aligned} E'_y &= \cosh\zeta\, E_y - \sinh\zeta\, B_z \\ B'_z &= \cosh\zeta\, B_z - \sinh\zeta\, E_y \end{aligned}$$

or, more generally, as long as $\vec{u} \perp \vec{B}$ and $\vec{u} \perp \vec{E}$,

$$\vec{E}' = \gamma(\vec{E} + \frac{\vec{u}}{c} \times \vec{B}), \qquad \vec{B}' = \gamma(\vec{B} - \frac{\vec{u}}{c} \times \vec{E}).$$

If we choose $\vec{u} = c\vec{E} \times \vec{B}/B^2$, we have

$$\begin{aligned} \vec{E}' &= \gamma\left(\vec{E} + (\vec{E} \times \hat{B}) \times \hat{B}\right) = \gamma\left(\vec{E} - \vec{E} + (\vec{E} \cdot \hat{B})\hat{B}\right) = 0 \\ \vec{B}' &= \gamma\left(\vec{B} - \frac{1}{B^2}(\vec{E} \times \vec{B}) \times \vec{E}\right) = \gamma\vec{B}\left(1 - \frac{E^2}{B^2}\right) = \frac{1}{\gamma}\vec{B}, \end{aligned}$$

as $\vec{u}^2/c^2 = E^2/B^2$. So in the $\mathcal{O}'$ frame, we have our previous situation: the particle spirals around the $\vec{B}'$ field, though more slowly than before[3]. But in the original $\mathcal{O}$ frame, there is an additional "$\vec{E} \times \vec{B}$ drift" velocity $\vec{u} = c\vec{E} \times \vec{B}/B^2$. Note that this velocity is in the same direction regardless of the sign of the charge of the particle, as it depends only on the fields $\vec{E}$ and $\vec{B}$, while the helical motion is reversed for particles of opposite charge.

There is an important special case, when the helical motion degenerates into constant motion along the $\vec{B}'$ field, so $\vec{v}'$ is a constant in the $\vec{B}$ direction,

[3]Note that the γ here, $\gamma(u) = \gamma(E/B)$, is not the same as the particle's γ which enters into the expression for ω.

but more importantly the drift velocity in the $\vec{E} \times \vec{B}$ direction is $u = cE/B$. Only particles with this value of the velocity component in that direction will travel in straight lines. and so a series of apertures can serve as a velocity selector. You learned all this as freshman, though then you assumed $\vec{u} \perp \vec{B}$, which we now see is not required.

Transforming to a frame moving at cE/B is not possible if $E > B$. Naïve use of the Lorentz transformation will give imaginary coordinates and fields. But there is a value of u which will annihilate the $\vec{B}$ field instead of the $\vec{E}$ field, so we choose $\vec{u} = c\vec{E} \times \vec{B}/E^2$ and we are left with a problem with a constant uniform $\vec{E}'$, no $\vec{B}'$, and a constant $d\vec{p}'/dt'$. Nonrelativistically this would give simple ballistic (parabolic) motion. Relativistically there are corrections due to the variation of γ in $\vec{p} = m\gamma\vec{v}$, so the motion is more complicated, but still analytically solvable (see Problem 12.3).

As you will see from homework, $E^2 - B^2$ and $\vec{E} \cdot \vec{B}$ are both invariants. This is why, for $\vec{E} \perp \vec{B}$, we had two distinct cases, depending on the sign of $E^2 - B^2$. This also shows that if $\vec{E}$ and $\vec{B}$ are not perpendicular in any frame, so $\vec{E} \cdot \vec{B} \neq 0$ in that frame, the same is true in any frame, and it is not possible to do a Lorentz transformation to a frame in which one of them vanishes. Still, the problem of uniform static $\vec{E}$ and $\vec{B}$ is solvable by brute force using cartesian coordinates.

There are many important situations in which charged particles move in a static but not uniform magnetic field. If there is a sense in which the field varies slowly, one may calculate the orbits as perturbations about the solution with uniform $\vec{B}$. Often, we find a helical motion, as for uniform $\vec{B}$, somewhat spiraling around the field lines, but with a drift slow compared to the helical velocity. This is explained in section 4, but ...

We will skip section 4, and go back now to section 1.

The last approximation we wish to consider uses the adiabatic invariance of the action. The action involved is $\oint \vec{P}_\perp \cdot d\vec{r}_\perp$ for the motion in the plane perpendicular to the field lines. But before we can discuss this, we need to know the **canonical** momentum $\vec{P}$ conjugate to $\vec{r}$, which **is not** the ordinary momentum $\vec{p} = m\gamma\vec{u}$. To find the **canonical momentum** we need to discuss the Lagrangian.

2 Lagrangian and Hamiltonian

In the first three pages of Jackson we learned the basic laws of electromagnetism: four Maxwell equations and one Lorentz force, and those managed to keep us busy for the next 500 pages. Now we have become more sophisticated, using four-dimensional notation, and we have reduced Maxwell's equations to 2, so we have a "complete" description of electromagnetism in only three equations. Do we really need to get any more sophisticated, and ask if we can rewrite things in Lagrangian formulation?

We might have asked the same question about Newtonian mechanics — after all, Lagrangian mechanics is nothing but a rewriting of $F = ma$, specifying lagrangians instead of force laws. But Lagrange invented his formulation for a reason— it helped him with complex solar system dynamics, and we would have a hard time doing quantum mechanics without the Hamiltonian. Also, these elegant reformulations are pretty! And finally, we will need a lagrangian formulation to make quantum field theory, and to develop concepts of gauge fields that are generalizable to non-Abelian field theories to give us the Standard Model of particle physics, which includes the gauge field theories called the Electro-Weak theory and Quantum Chromodynamics.

So, whatever the motivation, let us turn to the Lagrangian formulation of mechanics, starting with a free point particle. Recall that the way Hamilton's principle works is that for each conceivable motion from some specific initial position at an initial time to some specific final position at some final time, we associate an action, and the real motion is a path for which the action is an extremum, with small perturbations in the path $\vec{x}(t)$ producing no change in the action (to first order in the function $\delta x(t)$). Notice that this formulation treats space and time on the same footing, so it is a good starting point for a relativistic theory.

We usually write the action as an integral of the Lagrangian over time

$$A = \int L dt,$$

which sort of messes up the obvious relativity by picking out time. We expect the action to be a relativistic invariant, as its variation determines the physical path which must be physically the same regardless of which observer is describing it[4]. I will do things a bit differently from Jackson

[4]This is not a totally convincing argument, and in fact we are being a bit too demanding here, but the basic idea is right: A should be a scalar.

— let's begin by asking what is the simplest action that could describe a point particle? It is a functional of the path, and should be invariant, so the simplest possibility is the invariant length of the path, that is, the proper time. To get the units right we will multiply by $-mc^2$, so let's try

$$\begin{aligned} A &= -mc^2 \int d\tau = -mc \int \sqrt{dx^\mu dx_\mu} = -mc \int \sqrt{U^\alpha U_\alpha}\, d\tau \\ &= -mc^2 \int \sqrt{1 - \frac{\vec{u}^{\,2}}{c^2}}\, dt. \end{aligned}$$

So the Lagrangian is

$$L(\vec{x}, \vec{u}, t) = -\, mc^2 \big/ \gamma(\vec{u}) = -mc^2 \sqrt{1 - \frac{\vec{u}^{\,2}}{c^2}},$$

and we should note that it is not an invariant, but instead transforms so that Ldt is an invariant.

In three dimensional language this Lagrangian gives us a canonical momentum

$$\left(\vec{P}\right)_i = \frac{\partial L}{\partial u_i} = \frac{m u_i}{\sqrt{1 - \frac{\vec{u}^{\,2}}{c^2}}} = (\vec{p}\,)_i,$$

the same momentum we already associated with a relativistic particle. It also gives us, from the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial u_i} - \frac{\partial L}{\partial x_i} = 0,$$

the equation $p_i =$ constant, as x_i is an ignorable coordinate. This is, of course, the correct equation of motion for a free particle.

Shall we be a bit more ambitious? We would like to have electromagnetism enter. If our particle has a charge q, what invariant could we add to γL to produce an interaction with electromagnetism? The Lagrangian is supposed to depend only on positions and velocities, and if we want a relativistic (and translation) invariant Lagrangian, it can depend on position only through the value of the electromagnetic field at the point the particle is at. So the only quantity the particle can provide is the four-velocity U_α, which needs to be dotted into a vector. Electromagnetism provides only the 4-vector A^α and the field-strength $F^{\alpha\beta}$. We can't use $U_\alpha U_\beta F^{\alpha\beta}$ because it is identically zero, so if we want an interaction linear in the fields our only choice is

$$\gamma L_{\rm int} = -\frac{q}{c} U_\alpha A^\alpha, \quad \Longrightarrow L_{\rm int} = -q\Phi + \frac{q}{c}\vec{u}\cdot\vec{A}.$$

The first term looks like the right negative of the potential energy (recall L is often $T-V$) for the electrostatic field. If we now consider the full lagrangian

$$L(\vec{x},\vec{u},t) = -mc^2\sqrt{1-\frac{\vec{u}^{\,2}}{c^2}} + \frac{q}{c}\vec{u}\cdot\vec{A}(\vec{x},t) - q\Phi(\vec{x},t),$$

the canonical momentum becomes

$$\vec{P} = \partial L/\partial\vec{u} = \frac{m\vec{u}}{\sqrt{1-\frac{\vec{u}^{\,2}}{c^2}}} + \frac{q}{c}\vec{A}(\vec{x},t) = \vec{p} + \frac{q}{c}\vec{A},$$

so the canonical momentum is not the ordinary momentum $\vec{p} = m\gamma\vec{u}$, but has an extra piece proportional to the vector potential. The equations of motion[5] are now

$$\begin{aligned}
\frac{d}{dt}\underbrace{\frac{\partial L}{\partial u_i}}_{P_i} - \frac{\partial L}{\partial x_i} &= \frac{dp_i}{dt} + \frac{q}{c}\underbrace{\frac{d}{dt}\vec{A}_i}_{\left(\frac{\partial A_i}{\partial t} + u_j\partial_j A_i\right)} - \frac{q}{c}u_j\partial_i A_j + q\partial_i\Phi \\
&= \frac{dp_i}{dt} + \frac{q}{c}\frac{\partial \vec{A}_i}{\partial t} + q\partial_i\Phi + \frac{q}{c}\left(u_j\partial_j A_i - u_j\partial_i A_j\right) \\
= 0 &= \left(\frac{d\vec{p}}{dt} + \frac{q}{c}\frac{\partial\vec{A}}{\partial t} + q\vec{\nabla}\Phi - \frac{q}{c}\vec{u}\times\left(\vec{\nabla}\times\vec{A}\right)\right)_i \\
\text{so} \qquad \frac{d\vec{p}}{dt} &= q\vec{E} + \frac{q}{c}\vec{u}\times\vec{B}
\end{aligned}$$

so we see that this Lagrangian gives us the correct Lorentz force equation.

What is the Hamiltonian? $H = \vec{P}\cdot\vec{u} - L$, but reexpressed in terms of $\vec{P}$ rather than $\vec{u}$. As

$$\vec{u} = \vec{p}/m\gamma(u) = \frac{\vec{p}}{m}\sqrt{1-u^2/c^2} \quad \Longrightarrow \quad \vec{u} = \frac{c\vec{p}}{\sqrt{p^2+m^2c^2}},$$

[5]Note all vectors and indices here are three dimensional, without distinction of upper and lower indices. See footnote in section 2.2 for conversion to 4-D.

and $m\gamma(u) = \sqrt{p^2 + m^2c^2}/c$. Then we need to substitute $\vec{p} \to \vec{P} - q\vec{A}/c$. Thus

$$\begin{aligned} H &= \frac{\vec{P}\cdot\left(\vec{P} - q\vec{A}/c\right) + m^2c^2}{m\gamma(u)} - \frac{q}{cm\gamma(u)}\left(\vec{P} - q\vec{A}/c\right)\cdot\vec{A} + q\Phi \\ &= \frac{\left(\vec{P} - q\vec{A}/c\right)^2 + m^2c^2}{m\gamma(u)} + q\Phi = \sqrt{(c\vec{P} - q\vec{A})^2 + m^2c^4} + q\Phi. \end{aligned}$$

Note H is the total energy, the kinetic energy $p^0c + e\Phi$, so this just verifies $(p^0)^2 - \vec{p}^2 = m^2c^2$.

2.1 Adiabatic Invariance of Flux Through Particle Orbits

Before we continue with more formal developments of the Lagrangian and Hamiltonian presentations of electromagnetism, let us make use of the canonical momentum we have just found. From our mechanics course, we recall that if a system is such that it is a slowly varying perturbation on an integrable system, and if the motion in an action-angle pair is cyclic in the unperturbed system, the action will be approximately invariant, even over times where the motion changes considerably.

The application here is that the motion perpendicular to a uniform static magnetic field is cyclic, so the action $J = \oint \vec{P}_\perp \cdot d\vec{r}_\perp$ is an invariant. We need to use the **canonical** momentum $\vec{P} = \vec{p} + (q/c)\vec{A}$ here, rather than just $\vec{p} = m\gamma\vec{v}$. So the action is

$$J = \oint m\gamma\vec{v}_\perp \cdot d\vec{r}_\perp + \frac{q}{c}\oint \vec{A}\cdot d\vec{r}.$$

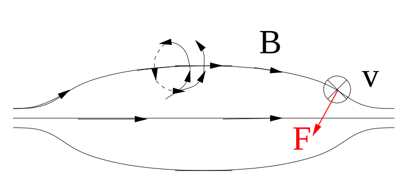
As the motion is a circle[6] with radius a, with $\vec{v}_\perp = -\vec{\omega}_B \times \vec{r}$, the first term is $-\int_0^{2\pi} m\gamma\omega_B a^2 d\theta = -2\pi m\gamma\omega_B a^2$. As $m\gamma\vec{\omega}_B = q\vec{B}/c$, this is just $-2q\Phi_B/c$, where Φ_B is the magnetic flux through the orbit. As Stokes theorem tells us $\oint \vec{A}\cdot d\vec{r} = \int_S \vec{\nabla}\times\vec{A} = \int_S \vec{n}\cdot\vec{B}$, the second term is just $q\Phi_B/c$, so

$$-J = \frac{q}{c}\Phi_B = \frac{q}{c}B\pi a^2 = \pi\frac{c}{q}\frac{p_\perp^2}{B},$$

[6]Signs here are a bit problematic. The usual description of a particle rotating about an axis with angular velocity $\vec{\omega}$ has $\vec{v} = \vec{\omega}\times\vec{r}$, and acceleration $\vec{a} = \vec{\omega}\times\vec{v} = \vec{\omega}\times(\vec{\omega}\times\vec{r}) = -\omega^2\vec{r}$. Notice, however, that (12.38) is the reverse of that, so $\omega = -\omega_B$. This accounts for the negative sign in $\vec{v}_\perp = -\vec{\omega}_B \times \vec{r}$.

and each of these expressions can be used as an approximate invariant if $\vec{B}$ varies slowly, compared to the gyroradius of the particle's motion.

For motion in a static purely magnetic field the speed is constant, and so is γ, but we see that the transverse speed is proportional to the square root of the magnetic field. Of course the constant speed squared is $v_\parallel^2 + v_\perp^2$, so if the particle drifts along a field line into a region of stronger magnetic field, its $v_\perp$ can grow until it consumes all the available speed, which means the motion along the field line must stop and reverse itself. This is called a magnetic mirror. This effect can be understood directly in terms of the Lorentz force by noting that, because $\vec{B}$ is divergenceless, if the field is getting stronger the field lines are converging, which means they have a radial component which produces a force on the circling charged particles which opposes their drift into this region.



2.2 Covariant description

Let us return to discussing the Lagrangian formalism. For the free particle our action could be described in completely covariant language as the proper time of the path taken. If we choose an arbitrary parameterization along the path, so $x^\mu(\lambda)$ is a path through spacetime, the infinitesimal proper time is

$$\frac{1}{c}\sqrt{\eta_{\mu\nu}dx^\mu dx^\nu} = \frac{1}{c}\sqrt{\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}\,d\lambda,$$

so we can write the action as

$$A = -mc\int\sqrt{\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}\,d\lambda.$$

We can now look for an extremal path $x^\mu(\lambda)$ in the usual way, with λ taking the role usually taken by time, and get the Euler-Lagrange equation

$$\frac{d}{d\lambda}\frac{\partial L}{\partial\frac{\partial x^\mu}{\partial\lambda}} = \frac{\partial L}{\partial x^\mu} = 0,$$

as x^μ is an ignorable coordinate. Evaluating the derivative in the left hand side, we find

$$\frac{d}{d\lambda}\left(\frac{\eta_{\mu\nu}\frac{dx^\nu}{d\lambda}}{\sqrt{\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}}\right) = 0,$$

or

$$\frac{dx^\mu}{d\lambda} = C^\mu \sqrt{\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}.$$

This equation determines less than meets the eye. It appears to be four differential equations determining the four function $x^\mu(\lambda)$. But, in addition to introducing four arbitary constants of integration, C^μ, in fact it determines only three independent functions of λ, as we can see from contracting it with itself. One equation is

$$\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = C^2\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda},$$

which gives us only that $C^2 = 1$ and not a differential equation helping to determine x^μ as a function of λ.

This lack of a deterministic equation should not be surprising if we recognize that the length of the path (the proper time) is completely independent of how the path is parameterized. The physics of $x^\mu(\lambda)$ is no different from the physics of $x^\mu(\sigma(\lambda))$, as long as σ is a monotonic function of λ. This inability to determine the future is a form of gauge invariance, though not the one we are used to (and will further discuss) of electrodynamics. But it is not a serious issue, for we can *choose* to take the proper time along the path as our parameter. Then $\eta_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = c^2$, and our equation of motion is

$$\frac{dx^\mu}{d\tau} = \frac{1}{m}p^\mu = \text{ constant},$$

as it should be.

What do we do about the contribution of L_{int} to the action,

$$A_{\text{int}} = \int \frac{-q}{c}\frac{dx^\mu}{d\tau}A_\mu\frac{1}{\gamma}dt = \int \frac{-q}{c}\frac{dx^\mu}{d\tau}A_\mu\, d\tau = \int \frac{-q}{c}A_\mu dx^\mu \quad ?$$

The last expression makes it clear — this involves completely covariant quantities. But to use the Euler Lagrange equations we go back to the penultimate expression, and write the equivalent of a Lagrangian for variation of $x^\mu(\lambda)$,

$$\tilde{L} = -mc\sqrt{\eta_{\alpha\beta}\frac{\partial x^\alpha}{\partial\lambda}\frac{\partial x^\beta}{\partial\lambda}} - \frac{q}{c}A_\alpha\frac{\partial x^\alpha}{\partial\lambda}$$

with $A = \int \tilde{L}d\lambda$. In looking at the Euler-Lagrange equations we need to recall that A_μ is a function of position, so again $d/d\tau$ of it is a stream derivative, so

$$\frac{d}{d\tau}A_\mu = U^\alpha \frac{\partial A_\mu}{\partial x^\alpha}.$$

Thus the equations corresponding to Euler-Lagrange give

$$\begin{aligned}\frac{d}{d\tau}\frac{\partial}{\partial U^\mu}\left(-mc\sqrt{\eta_{\nu\rho}U^\nu U^\rho} - \frac{q}{c}U^\nu A_\nu\right) &= \frac{\partial}{\partial x^\mu}\left(-mc\sqrt{\eta_{\nu\rho}U^\nu U^\rho} - \frac{q}{c}U^\nu A_\nu\right)\\ -m\frac{d}{d\tau}U_\mu - \frac{q}{c}U^\nu\frac{\partial A_\mu}{\partial x^\nu} &= -\frac{q}{c}U^\nu \frac{\partial A_\nu}{\partial x^\mu}.\end{aligned}$$

In the second line we imposed (after taking the derivative) the constraint $U^2 = c^2$. Now

$$m\frac{d}{d\tau}U_\mu = \frac{q}{c}U^\nu\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right) = \frac{q}{c}U^\nu F_{\mu\nu}, \tag{1}$$

the correct Lorentz force.

Note that the four dimensional **canonical** momentum can be defined as[7]

$$P_\alpha = -\frac{\partial\tilde{L}}{\partial\frac{\partial x^\alpha}{\partial\lambda}} = mU_\alpha + \frac{q}{c}A_\alpha,$$

where we have required our parameter λ to be c times the proper time.

Notice that we now have

$$\left(P_\alpha - \frac{q}{c}A_\alpha\right)\left(P^\alpha - \frac{q}{c}A^\alpha\right) = m^2 U_\alpha U^\alpha = m^2c^2.$$

This is one example of the minimum substitution principle, which states that electromagnetic interactions of other objects can be obtained from replacing p_α of the theory without electromagnetic interactions with $p_\alpha - (q/c)A_\alpha$ everywhere.

[7]Why the minus sign? In deriving the Lorentz force earlier, we used 3-D notation with $u_i = (\vec{u})_i$, but as part of the 4-vector U, $U^\alpha = (\gamma, \gamma u_i)$ but $U_\alpha = (\gamma, -\gamma u_i)$, $U_i = -u_i$. For covariance, when we differentiate L with respect to contravariant U^α, we need to get the *covariant*

$$P_\alpha = (E/c, -\vec{P}) \propto \frac{\partial L}{\partial U^\alpha}.$$

To get the sign right for the spatial components, we need the proportionality constant to be -1.

2.3 Action for the Electromagnetic Fields

We have written a Lagrangian which determines the dynamics of charged particles in the presence of a predetermined electromagnetic field, but of course this course has been devoted to a study of the mutual interactions of charged particles and electromagnetic fields. The Lorentz force and Maxwell's equations form a coupled set of equations which determine the evolution of both the particles and the fields. We need a Lagrangian that does both as well.

What will such a Lagrangian depend on? The fields are degrees of freedom at every point of space (-time). So we need the lagrangian formulation of a continuum[8], field, where the discrete variables q_i are replaced by fields, generally some set of $\phi_i(\vec{x}, t)$, and the velocities are replaced by $\partial_\mu \phi_i$. Rather than a sum over discrete degrees of freedom, the Lagrangian becomes an integral of a Lagrangian density $\mathcal{L}$ over space, and the action becomes its integral over space-time. Even for nonrelativistic physics the space and time derivatives enter the same way, and the Euler-Lagrange equations become

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial \phi_i / \partial x^\mu)} - \frac{\partial \mathcal{L}}{\partial \phi_i} = 0.$$

What are our fundamental fields? A lagrangian should depend on the fields ϕ_i and their first derivatives $\partial_\mu \phi_i$, and will give equations of motion with second order derivatives. Maxwell's equations involve only first derivatives of $\vec{E}$ and $\vec{B}$, or $F^{\mu\nu}$, but we know that $F^{\mu\nu}$ can be written in terms of first derivatives of A^μ, so we take the basic degrees of freedom to be the fields $A^\mu(x^\nu)$.

We have already seen that the action should contain $-(q/c) A_\mu dx^\mu$ if we have a single particle of charge q. That is, the Lagrangian has an interaction term $-q_i \Phi(\vec{x}_i) + \frac{q_i}{c} \vec{u}_i \cdot \vec{A}(\vec{x}_i)$. If we have many particles, the interaction term in L is

$$\begin{aligned} \sum_i \left(-q_i \Phi(\vec{x}_i) - \frac{1}{c} q_i \vec{u}_i \cdot \vec{A}(\vec{x}_i, t) \right) \quad &\rightarrow \quad \int d^3x \left(-\rho(\vec{x}) \Phi(\vec{x}) - \frac{1}{c} \vec{J}(\vec{x}) \cdot \vec{A}(\vec{x}) \right) \\ &= \quad -\frac{1}{c} \int d^3x A_\alpha(\vec{x}) J^\alpha(\vec{x}). \end{aligned}$$

---

[8]Those who have not seen the lagrangian formulation of field dynamics might want to look at my text in `www.physics.rutgers.edu/~shapiro/507/gettext.shtml` and look at chapter 8 (or get `book9_2.pdf` from the same location). Of course there are also many published books as well, in particular Landau and Lifshitz "The Classical Theory of Fields".

So we have a piece involving A^μ which will contribute a term J_μ to the Euler-Lagrange equation for A, but we need a pure electromagnetic field term to generate the left hand side of Maxwell's equations, which are linear in the fields, so the term in the Lagrangian should be quadratic in the fields, Lorentz invariant, and involve a total of two derivatives. Let us try

$$\mathcal{L} = -\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} - \frac{1}{c} J_\mu A^\mu,$$

where it is understood that $F_{\mu\nu}$ stands for $\partial_\mu A_\nu - \partial_\nu A_\mu$ and is not an independent field. Note the only contribution to $\partial\mathcal{L}/\partial A_\mu$, which is supposed to be taken with the derivative terms held fixed, is the $-J^\mu/c$ from the interaction term. We have

$$\frac{\partial F_{\mu\nu}}{\partial\left(\dfrac{\partial A_\rho}{\partial x^\sigma}\right)} = \delta^\sigma_\mu \delta^\rho_\nu - \delta^\sigma_\nu \delta^\rho_\mu,$$

so

$$\frac{\partial \mathcal{L}}{\partial\left(\dfrac{\partial A_\rho}{\partial x^\sigma}\right)} = -\frac{1}{4\pi} F_{\rho\sigma},$$

and the full Euler-Lagrange equation is

$$-\frac{1}{4\pi} \partial_\sigma F^{\sigma\mu} + \frac{1}{c} J^\mu = 0,$$

or

$$\partial_\sigma F^{\sigma\mu} = \frac{4\pi}{c} J^\mu.$$