Noether's Theorem

Classical mechanics for a set of classical variables \( q_i \) begins with a Lagrangian \( L \) and an action \( S \):

\[
S = \int dt L(q_i, \dot{q}_i, \ddot{q}_i)
\]

and

\[
\delta S = 0 \Rightarrow \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i = 0
\]

Classical mechanics for fields begins with the fields themselves, defined for all points in space \( x \), time \( t \), generally \( \phi(x, t) \), and their derivatives \( \frac{\partial \phi}{\partial x} \) and \( \frac{\partial \phi}{\partial t} \). Generally, \( \delta \phi \).

The analogy of \( L \) is the Lagrange density \( \mathcal{L} \):

\[
\mathcal{L} = \int \mathcal{L} \, d^3x
\]

and the action \( S \) becomes \( d^4x \cdot \mathcal{L} \):

\[
S = \int d^4x \mathcal{L} (\phi, \partial \phi)
\]

The bottom line is \( \delta S = 0 \):

\[
\partial \mu \left[ \frac{\partial \mathcal{L}}{\partial (\partial \phi)} \right] - \frac{\partial \mathcal{L}}{\partial \phi} = 0
\]

So our starting point. Let's do an example to refresh memory.
\[ L = \sum_{j=1}^{N} \frac{1}{2} m_j \dot{y}_j - \frac{1}{2} k \left( y_j - y_{j+1} \right)^2 - V(y_j) \]

Equation of Motion:
\[
\begin{bmatrix}
\frac{d}{dt} \frac{\partial L}{\partial \dot{y}_j} - \frac{\partial L}{\partial y_j} = 0
\end{bmatrix}
\]

Now suppose \( y \) varies smoothly with \( x \):
\[
y_{j+1} = y_j + \frac{dy_j}{dx} \bigg|_{x=x_0} + \frac{d^2 y_j}{dx^2} \bigg|_{x=x_0} \text{...}
\]

\[
m \ddot{y}_j + k \alpha^2 \frac{d^2 y_j}{dx^2} + \frac{\partial V}{\partial y_j} = 0
\]

\[
\Phi(x, t) = \frac{y}{k \alpha^2} - \frac{1}{c^2} \frac{1}{k \alpha^2} \sqrt{\frac{m}{k \alpha^2}} \left[ \frac{\partial^2 y_j}{\partial x^2} - \frac{1}{k \alpha^2} \frac{\partial^2 V}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 V}{\partial t^2} \right]
\]

\[
\dot{\Phi} = \frac{\dot{V}}{k \alpha^2}
\]

\[
\phi(x) \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial V}{\partial \phi} = 0
\]

\[
\phi(x) \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial \phi}{\partial x} - \frac{\partial V}{\partial \phi} = 0
\]

\[
\phi(x) \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial \phi}{\partial x} = 0
\]

\[
\Phi(x) = \frac{1}{\alpha} \left( \frac{\partial}{\partial x} \left( \frac{\partial \phi}{\partial x} \right) - \frac{\partial \phi}{\partial x} \right)
\]

\[
\phi(x) = \frac{1}{\alpha} \left( \frac{\partial}{\partial x} \left( \frac{\partial \phi}{\partial x} \right) - \frac{\partial \phi}{\partial x} \right)
\]
Now back to $\phi$. Suppose we imagine only a change in the field variable

$$\phi(x,t) \to \phi(x,t) + \delta \phi(x,t)$$

and at the same time

$$\partial_m \phi(x,t) \to \partial_m \phi(x,t) + \partial_m \delta \phi(x,t)$$

which is independent from $\phi$.

The change in $\mathcal{L}$ is

$$\delta \mathcal{L} = \int \left( \frac{\partial \mathcal{L}}{\partial \phi} \delta \phi + \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \partial_\mu \delta \phi \right) dx$$

Recall the equation of motion

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0$$

So

$$\delta \mathcal{L} = \int \left[ \partial_m \left( \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \right) \delta \phi_m + \frac{\partial \mathcal{L}}{\partial \phi} \partial_m \delta \phi \right] dx$$

$$= \partial_m \left[ \sum \frac{\partial \mathcal{L}}{\partial \phi_i} \delta \phi_i \right]$$

$$= \partial_m \left[ \sum \frac{\partial \mathcal{L}}{\partial \phi_i} \delta \phi_i \right]$$

Define \([\ ]\) as a current.

Now if it happens that $\delta \mathcal{L} = 0$, the Lagrangian is invariant under the change of $\phi$'s.

Then

$$\partial_m J^m = 0$$
Symmetry implies conservation law

The "charge" associated with the current is also conserved

\[ Q = \int d^3x \, j_0 \]

Changes in \( Q \) could be "internal" - re-define \( Q \) at every space-time point. Consider \( \mathcal{O}(\mathbf{x},t) \)

\[ \mathcal{O}(\mathbf{x}+\mathbf{e},t) \rightarrow \mathcal{O}(\mathbf{x},t) + \mathcal{O}(\mathbf{e},t) \]

\[ \mathcal{O}(\mathbf{x}+\mathbf{e},t) = \text{constant} \]

An example of a global symmetry transform

\( \mathcal{O}(\mathbf{x},t) \) varies w.r.t. \( \mathbf{x} \) t: "Local symmetry transformation"

Self could involve changes in coordinates. Latter ones lead to "conserved currents" associated with energy & momentum.

Internal ones are more interesting for the moment
Example

Let us suppose for the sake of argument that we have a classical physical system described by a complex field \( \Psi(x) = \Psi_1(x) + i \Psi_2(x) \) (or we could think of it as a pair of real fields). A Bose-Einstein condensate, or the other way around, be it of atoms or of liquid helium. Suppose that its Lagrange density does not have any explicit dependence on space and time—instead, it only depends on \( x \) and \( t \) through \( \Psi \) and \( \nabla \Psi \). That would be the case for a stretched string where \( \Psi(x,t) \) is the real displacement of the string, and you recall \( \sqrt{c^2} \).

\[
\boxed{L = \left( \frac{\partial \Psi}{\partial t} \right)^2 - (\nabla \Psi)^2 = (\partial_t \Psi)^2}
\]

A complex field would have a pair of \( \Psi \)’s so imagine

\[
\begin{align*}
\Psi_1 & \quad \Psi_2 \\
\end{align*}
\]

\* turns out to construct relativistically

consistent theories, you have to do this
Now suppose that we imagine making a global change in the rotation in the field variables

\[
\begin{pmatrix}
\Phi_1' \\
\Phi_2'
\end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}
\begin{pmatrix} \Phi_1 \\
\Phi_2
\end{pmatrix}
\]

or \( \Phi' = \Phi \circ \Theta \) and we call \( \Phi' = (\Phi_1', \Phi_2') \)

and suppose that under this change the \( \mathcal{L} \) is invariant. For the case of the string, this is just a rotation in the field space \( (\Phi_1, \Phi_2) \).

Let's suppose that \( \Theta \) is infinitesimal, \( \Theta = \varepsilon \varepsilon \)

\( \Phi' = \Phi + \varepsilon \varepsilon \)

\( \mathcal{L}' = -i \varepsilon \varepsilon \mathcal{L} \)

\varepsilon is the same everywhere in space. How does \( \mathcal{L} \) change?

\[
\mathcal{L}' = \frac{\partial}{\partial \Phi} \mathcal{L} + \varepsilon \left( \frac{\partial}{\partial \Phi} \mathcal{L} \right)
\]

but \( \mathcal{L}' = -i \varepsilon \varepsilon \mathcal{L} \)

\( \mathcal{L}(\Phi + \varepsilon \varepsilon) = -i \varepsilon \varepsilon \mathcal{L}(\Phi) \)

\( \mathcal{L} \) is the equation of motion

\( \frac{\partial}{\partial \Phi} \mathcal{L} = 0 \)

\[
\mathcal{L}' = -i \varepsilon \varepsilon \left[ \frac{\partial}{\partial \Phi} \left( \frac{\partial \mathcal{L}}{\partial \Phi^*} \right) \varepsilon + \frac{\partial}{\partial \Phi} \frac{\partial \mathcal{L}}{\partial \Phi^*} \varepsilon \right]
\]

and

\[
\mathcal{L}' = -i \varepsilon \varepsilon \left[ \frac{\partial}{\partial \Phi} \left( \frac{\partial \mathcal{L}}{\partial \Phi^*} \right) \varepsilon + \frac{\partial}{\partial \Phi} \frac{\partial \mathcal{L}}{\partial \Phi^*} \varepsilon \right]
\]
\[\mathcal{L} = \frac{1}{2} \left[ (\partial_{\mu} \varphi_1)^2 + (\partial_{\mu} \varphi_2)^2 \right] - V(\varphi_1, \varphi_2)\]

\[
\left[ \begin{array}{c}
\varphi_1' \\
\varphi_2'
\end{array} \right] = \left[ \begin{array}{cc}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{array} \right] \left[ \begin{array}{c}
\varphi_1 \\
\varphi_2
\end{array} \right] \sim \text{small } \left[ \begin{array}{cc}
1 & \theta \\
-\theta & 1
\end{array} \right] \left[ \begin{array}{c}
\varphi_1 \\
\varphi_2
\end{array} \right]
\]

1.e. \[\mathcal{L} \varphi_1 = \Theta \varphi_2 \quad \text{and} \quad \mathcal{L} \varphi_2 = -\Theta \varphi_1\]

\[\mathcal{L} = \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi_1)} + \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi_2)} + \partial_{\mu} \mathcal{L}_{\varphi_1} + \partial_{\mu} \mathcal{L}_{\varphi_2}
\]

Suppose \[V(\varphi_1, \varphi_2)\] is a function only of \[\varphi_1^2 + \varphi_2^2\] - obviously, the potential part of \[\mathcal{L}\] is unchanged under this transformation and last two terms give zero. Kinetic term

\[\mathcal{L} = \left[ \begin{array}{c}
\partial_{\mu} \varphi_1 \\
\partial_{\mu} \varphi_2
\end{array} \right] \left[ \begin{array}{c}
\Theta \varphi_2 \\
-\Theta \varphi_1
\end{array} \right] + \partial_{\mu} \mathcal{L}_{\varphi_1} + \partial_{\mu} \mathcal{L}_{\varphi_2}
\]

\[= 0, \text{triv!}\]

Yes, it's a symmetry. What is the associated conserved current?

\[J^\mu = \sum_j \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi_j)} = (\partial_{\mu} \varphi_1) \Theta \varphi_2 + (\partial_{\mu} \varphi_2) (-\Theta \varphi_1)
\]

\[= (\varphi_2 \partial_{\mu} \varphi_1 - \varphi_1 \partial_{\mu} \varphi_2) \text{ up to a small constant.}
\]

A lot like Schrödinger eqn ...
So for we considered a variation of $\phi$

$$\delta \phi = i q \epsilon \phi$$

with $\epsilon$ a constant over space, and we saw that

if $L(\phi') = L(\phi)$, with $\phi' = i \epsilon + \phi$, then the system

has a conserved current. What if we now let

e vary from point to point in space? That is, suppose that we asked for the following transformation

to be a symmetry:

$$\delta \phi(x) = i q \epsilon(x) \phi(x)$$

in

$$\phi(x') = e^{i \phi(x) - \phi(x')}$$

at each $x = 0$

We know that the variation in $L$ is still:

$$\delta L = \frac{\delta L}{\delta \phi} \delta \phi + \frac{\delta L}{\delta (\partial \phi)} \delta (\partial \phi)$$

but now $\delta (\partial \phi) = \partial \delta \phi$

$$\delta L = i q \epsilon \phi \partial \phi + i q \epsilon \phi \epsilon \phi$$

The equations of motion

$$\partial \phi \frac{\partial \delta \phi}{\partial \phi} = \partial \phi \frac{\partial \delta \phi}{\partial (\partial \phi)}$$

again allow us to write

$$\delta L = \epsilon(x) \partial \phi \delta \phi(x) + \frac{\partial L}{\delta \phi} \delta \phi$$

with $J^m = i q \epsilon \phi \epsilon \phi$. It is the same current as before — it is conserved.

$$\delta L = \epsilon(x) \partial \phi \delta \phi(x) + i q \epsilon \phi \epsilon \phi$$

What about the second term?
The only way to make it zero is to introduce new fields into \( E \), whose variation cancels the current. The current is a 4-vector, so let's add in a vector field \( A(x) \), whose variation is

\[
SA(x) = \partial_x A(x) + \partial_t x
\]

1.e.

\[
\bar{A}(x) = A(x) + \partial_t x
\]

\[
A(x,t) = A(x,t) + \partial_t e \frac{c}{c^2}
\]

Now

\[
S^2 = S_x^2 \text{ mtfm} + \left( \frac{\partial}{\partial A^i} S_{A^i} - \frac{2}{3} \partial_t A \right)
\]

\[
S^2 = S_x^2 + \frac{2}{3} \partial_t \partial_x x + \frac{2}{3} \partial_t \partial_x x
\]

\[
\partial_x (\partial_x x) \partial_t x
\]

This is zero as long as

\[
\frac{\partial x}{\partial A^i} = -\partial_x A^i \rightarrow \text{unique specification of new field to current}
\]

\[
\bar{A} = \frac{1}{2} - \partial_x A^i \partial_t x
\]

and

\[
\frac{\partial x}{\partial (\partial_x A^i)} = -\frac{\partial x}{\partial (\partial_x A^i)} \rightarrow \text{so that the}
\]

\[
\frac{\partial x}{\partial (\partial_x A^i)} - \frac{1}{2} \partial_x A^i \partial_t x
\]

\[
\partial_x \partial_x x \partial_t x
\]

\[
\partial_x \partial_x x \partial_t x
\]

\[
\partial_x \partial_x x \partial_t x
\]

The new condition also not give conservation laws, but instead conditions on the couplings of fields.
1) Suppose a global symmetry
\[ S \phi \rightarrow e^{i\theta} \phi \]
leave invariant. This is a conserved current
\[ \partial \mu J^\mu = 0 \]

2) Replace global symmetry by local symmetry
\[ S (x, t) \rightarrow e^{i\theta (x, t) \phi (x, t)} \]
force \[ \delta L = 0 \] also current
as new new gauge field is 4-vector \[ A^\mu (x) \]

Coupling b) \[ L_I = - \frac{1}{2} \epsilon^{\mu \nu \rho \sigma} F_{\mu \nu} F_{\rho \sigma} \]
the same conserved current as in c)

\[ c) \quad \text{instead} \quad 2 \mu A^\mu - \partial \nu A^\nu = F_{\mu \nu} \]
Symmetry transformation is

\[ \Phi'(x') = e^{i \gamma \phi(x)} \Phi(x) \]

and \( A' = A + \gamma \) for \( E \) invariance of \( \Phi' \).

A way to encode this transformation in \( \mathbb{R}^3 \) is to replace the usual derivative \( \partial_x \Phi \) by the covariant derivative \( \nabla_x \Phi \) and to write \( \Phi' \) as a function of \( (\nabla_x \Phi)' \).

Let \( D_x \Phi = \partial_x \Phi - i \gamma A_x \Phi \)

and write \( L \) as a function of \( (D_x \Phi)' \).

\[ \begin{align*}
D_x \Phi' &= \partial_x \Phi' - i \gamma A_x \Phi' \\
&= \partial_x \left[ e^{i \gamma \phi} \right] \\
&= e^{i \gamma \phi} \left( \partial_x \Phi + i \gamma A_x \Phi \right)
\end{align*} \]

\[ \begin{align*}
&= e^{i \gamma \phi} \left( \partial_x \Phi' - i \gamma A_x \Phi' \right) \\
&= e^{i \gamma \phi} \left[ D_x \Phi' \right]
\end{align*} \]

\[ L = \left( D_x \Phi' \right)^* (D_x \Phi') - V(\Phi')^2 \]

\[ = \left( e^{i \gamma \phi} \left( D_x \Phi \right)^* \left( D_x \Phi \right) - V(\Phi')^2 \right) \]

So invariant,

\[ |(\Phi')^2 = |e^{i \gamma \phi} \Phi|^2 \]
And finally, what would be a good candidate $L$ for the new electromagnetic degrees of freedom? To get a scalar, contract indices would make it a tensor:

$$L = c_1 \left[ F_{\mu\nu}^2 \right] + c_2 \left[ F_{\mu\nu} \right]^2 + \cdots$$

$$F_{\mu\nu} = F_{\mu\nu} F^{\mu\nu}$$

and indeed, the Lagrange density which produces Maxwell's equations is "the simplest thing you can write down." 

$$L = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} \, \mathcal{J}_\mu A^\mu + \text{terms independent of } A, \mathcal{J}.$$ 

Obviously all constants made arbitrary by construction.

The coefficients and signs are chosen arbitrarily so that the field equations have CGS conventions.

Why isn't there more? Let's derive that for a while. First, show how this $L$ produces Maxwell's equations as its equations of motion.

Write

$$L = -\frac{1}{16\pi} \int d^4x g^{\mu\nu} \left[ \partial_\mu A^\sigma - \partial_\sigma A^\mu \right] \left[ \partial_\nu A_\lambda - \partial_\lambda A_\nu \right] - \frac{1}{2} \, \mathcal{J}_\mu A^\mu$$

$$\frac{\partial L}{\partial (DA^d)} = -\frac{1}{16\pi} \int d^4x g^{\mu\nu} \left[ S_\rho \, ^{\mu \nu} \delta_a F^{\rho \lambda} - S_\rho \, ^{\mu \nu} \delta_a F_{\lambda}^{\rho} + S_\rho \, ^{\mu \nu} \delta_a F_{\rho \lambda} - S_\rho \, ^{\mu \nu} \delta_a F^{\rho \lambda} \right]$$

$$= -\frac{4}{16\pi} \Gamma_{\beta \rho} = \frac{1}{4\pi} F_{\beta \rho}$$

we need $\frac{\partial L}{\partial (DA^d)} = \frac{\partial L}{\partial DA^d} = 0$
The covariant transformation
\[ A'_\mu = A_\mu + \partial_\mu A(\Phi) \]
(or)

is called a **local gauge transformation**. Theories which are invariant under local gauge transformations are called "gauge theories" and include

**Electrodynamics**

and actually under a matrix generalization of

\[ \Phi(x) \rightarrow e^{i G(\Phi)} \Phi(x) \rightarrow \Phi \rightarrow R \Phi \]

**QCD** (strong interactions)

Weinberg-Salam model (weak + electromagnetic interactions)

i.e. all nature gauge invariance for \( G : \) coordinate representation.

How to satisfy local gauge invariance?

\[ L = \mathcal{L}(\Phi, \partial \Phi, F_{\mu \nu}) \]

\[ F_{\mu \nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \]

\[ D_\mu \Phi = \partial_\mu \Phi - ig A_\mu \Phi \quad \text{covariant derivative} \]

For then:
\[ \frac{d\mathcal{L}}{dA_\mu} = \frac{\partial \mathcal{L}}{\partial (D_\mu \Phi)} (\rightarrow i g) = \frac{\partial \mathcal{L}}{\partial (D_\mu \Phi)} (\rightarrow i g) \rightarrow J_\mu \]
\[ \text{and} \quad \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{1}{2} \mathbf{J}_d \]

\[ \mathbf{D} = \mathbf{E} + \frac{\mathbf{J}_d}{\varepsilon_0} \]

Now the field strength tensor \( \mathbf{F}^{\mathbf{D}} = \nabla \mathbf{A} - \nabla \mathbf{A}^T \) is nothing more than the \( E \) and \( B \) fields:

\[ \mathbf{B} = \nabla \times \mathbf{A} = \mathbf{E} \times \mathbf{B} = \varepsilon_0 \mu_0 \mathbf{J}_d \]

\[ \mathbf{E} = -\nabla \mathbf{A} - \frac{1}{2} \frac{\partial \mathbf{A}}{\partial t} \]

\[ E_i = -\varepsilon_0 \mu_0 \varepsilon_0 \mathbf{A}_0 = \varepsilon_0 \mathbf{A} \]

\[ F^{\mathbf{D}} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix} \]

\[ F^{\mathbf{B}} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{bmatrix} \]

For the inhomogeneous field equation is fine \( (\epsilon \nabla)^2 = (\epsilon \nabla)^2 \)

\[ \nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \]

\[ \nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J}_d \]

2 \% 4 Maxwell equations

The homogeneous equation vanish by construction

\( \nabla \cdot \mathbf{B} = 0 \rightarrow \mathbf{B} = \nabla \times \mathbf{A} \)

\( \nabla \times \mathbf{A} \times \mathbf{B} \)

In convenient coordinate language they are

\[ \frac{d}{dt} F_{\mathbf{d}} + \frac{\partial F_{\mathbf{d}}}{\partial x} + \frac{\partial F_{\mathbf{d}}}{\partial y} + \frac{\partial F_{\mathbf{d}}}{\partial z} = 0 \quad \mathbf{d} = \mathbf{B} + \mathbf{A} \]
where $\mathbf{\sigma}^{ab}$ = "stress field strength tensor"

$$\mathbf{\sigma}^{ab} = \frac{1}{2} \epsilon^{abcd} F_{cd} = \begin{bmatrix} 0 & -B_x & -B_y & -B_z \\ -B_x & 0 & E_z & -E_y \\ -B_y & -E_z & 0 & E_x \\ -B_z & E_y & -E_x & 0 \end{bmatrix}$$

$$\nabla_a \epsilon^{ab} = \frac{1}{2} \epsilon^{abcd} \nabla_b F_{cd} = \frac{1}{2} \epsilon^{abcd} \nabla_b (\partial_k A^k - \partial^k A_k) = 0 \quad \text{because for each term, } \epsilon \text{ is antisymmetric}$$

While we're at it, let's think about the Lagrangian Hamiltonian.

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

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

look at 4x4 tensors, sum up

$$\mathcal{L} = \frac{1}{8\pi} (E^2 - B^2)$$

Hamiltonian? Write $L = -\frac{1}{8\pi} \left[ \frac{\partial A_k}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^k} \right] \frac{\partial A^k}{\partial x^\nu}$

Need canonical coordinates - obvious choice the 4 $A_k$'s

Canonical momenta?

$$k = 1, 2, 3 \quad \Pi^k = \frac{\partial L}{\partial \dot{A}_k} = -\frac{1}{8\pi} \left[ \frac{\partial A_k}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^k} \right] \frac{\partial A^k}{\partial x^\nu} = \frac{E_k}{4\pi}$$

$$\Pi^0 = \frac{\partial L}{\partial \dot{A}_0} = 0 \quad \Rightarrow \quad L = \frac{1}{2} \sum_{k=1}^3 \Pi^k A^k - L = \frac{1}{4\pi} \left[ \frac{1}{2} (E^2 + B^2) \right]$$
At this point things get a bit tricky.

Note (first of all) that there is no momentum conjugate to $A_0$: $A_0$ has no independent dynamics.

Second, in the absence of sources, Maxwell's equations (which we could as well derive as a set of Hamilton equations) require

$$\mathbf{\nabla} \cdot \mathbf{E} = 0$$

so the 3 $E$'s are not independent.

The system has constraints. They are related to gauge symmetry; not all the $A_0$'s are independent variables.

To the best of my knowledge, this is not a problem for the classical theory, but it is a problem in constructing the quantum theory using "canonical quantization":

1) write down classical $L$

2) construct classical $H$; identify $p_i$'s

3) Impose canonical quantization condition

$$\{B_0, p_0\} = i\hbar \delta_{ij}$$

4) Construct quantum $H$ by replacing coordinates $q_i$ momenta $p_i$ with comm. operators.

To proceed beyond point 3) it is necessary to fix a gauge and in order to make $H$ time independent the gauge must be non-covariant, like Coulomb gauge. But then the formulas all look non-covariant (there is an instantaneous Coulomb interaction, for example) and then it is a big wish to show that the quantum theory is covariant all the non-covariant parts cancel.
Anyway, back to Lagrangians. The 4-vectors form a
four-integer and transform under Lorentz transformations
like a 4-vector. In matrix notation
where the L.T. is
\[
\Lambda = \exp(\alpha \omega \cdot \beta - \beta \omega)
\]
\[
\begin{pmatrix}
A'_0 \\
A'_1 \\
A'_2 \\
A'_3
\end{pmatrix} = \Lambda \begin{pmatrix}
A_0 \\
A_1 \\
A_2 \\
A_3
\end{pmatrix}
\]

E and B are elements of the second rank tensor \( F^{\alpha \beta} \).
It transforms as
\[
F' = \frac{\partial x'^i}{\partial x^\alpha} \frac{\partial x'^j}{\partial x^\beta} F^{ij}
\]

or
\[
F' = \Lambda F \Lambda^T \quad \text{(matrix notation)}
\]

Let's multiply out for a boost along the \( \gamma \) axis:

\[
A = \begin{pmatrix}
\gamma & \beta \gamma & 0 & 0 \\
\beta \gamma & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

\[
E'_1 = E_1, \quad B'_1 = B_1
\]
\[
E'_2 = \gamma (E_2 - \beta B_3), \quad B'_2 = \gamma (B_2 + \beta E_3)
\]
\[
E'_3 = \gamma (E_3 + \beta B_1), \quad B'_3 = \gamma (B_3 - \beta E_2)
\]

\[
E' = \gamma (E + \beta \times B) - \frac{\gamma}{\gamma + 1} \beta (\beta \cdot E)
\]
\[
B' = \gamma (B - \beta \times E) - \frac{\gamma}{\gamma + 1} \beta (\beta \cdot B)
\]

These I checked

This I did not.
Example: point charge at origin in frame $K''$, point charge moving with $\vec{v} = \hat{x}v$ in frame $K$. 

In $K''$, observer is at $(x_1', y_1', z_1') = (0, 0, 0)$ but in $K$, observer is at $(x', y', z') = (-vt', b, 0)$. 

Along $\vec{x}'$, $\vec{B}' = 0$, $\vec{E}' = \frac{q \hat{y}'}{(v')^2}$ 

\[ E_1' = -\frac{q \sqrt{b^2 + (v')^2}}{(v')^2}, \quad E_2' = -\frac{q \beta b}{(v')^3}, \quad E_3' = 0 \]

and $(v')^2 = b^2 + v^2 t^2$.

We have to convert both the coordinates and the fields. To convert the fields, $x' = x t^2$. 

\[ E_1' = \beta \left( -\frac{\sqrt{b^2 + (v't)^2}}{(v')^2} \right), \quad E_2' = \frac{\beta b}{\left[ b^2 + (v't)^2 \right]^{3/2}} \]

Now we transform the fields

\[ E_1 = E_1' = -\frac{q \sqrt{b^2 + (v't)^2}}{(v')^2}, \quad E_2 = \frac{\beta b}{\left[ b^2 + (v't)^2 \right]^{3/2}}, \quad E_3 = \frac{\beta \sqrt{b^2 + (v't)^2}}{\left[ b^2 + (v't)^2 \right]^{3/2}} \]

\[ B_3 = \beta x E_2' = \beta E_2 \]
Notes transverse $\mathbf{B}$ \((B_3)\) and as \(v \to c \) \(\mathbf{B}_3 \to 1E_2\).

Note NR limit: \(\frac{B}{b} = \frac{\nu x}{c} \left(\frac{\nu}{v} \cos \theta\right)\) \((\text{Biot-Savart: } \frac{\nu}{v} = \frac{1}{xL})\).

Note extreme relativistic limit \(E_2 \propto y\) but \(E_1 \sim 1\).

\[
E_2 \propto \frac{b \nu}{b^2} \quad \text{at} \quad \nu t = b \quad \Rightarrow \quad E_1 = \frac{\partial}{\partial t} \left(\frac{b \nu}{b^2}\right) \sim \frac{b}{b^2}
\]

Fields become a pulse of plane wave.

\[
\begin{array}{c}
\uparrow \quad \downarrow \\
\leftarrow \quad \rightarrow \\
\uparrow \downarrow \uparrow \downarrow
\end{array}
\]

\(b\) at rest

Indeed, if detector averages over time \(T > \frac{b}{\nu}\), \(\langle E_1 \rangle = 0\).

[Future: can exploit analogy between field of rel particle and plane wave = “Weisshenker-Willis approximation.”]
So why Maxwell's equations? Why is

\[ L = -\frac{1}{8\pi} F_{\mu\nu} F^{\mu\nu} + \text{nothing else?} \]

In pure classical case, only justification is observation of symmetry.

The answer involves dynamics (and the answer is really, there is something else).

Let's work in units with \( \hbar = c = 1 \), so in dimensions

\[ \text{Energy} = \frac{1}{\text{[Length]^4}} \quad (\text{[EC} = \text{energy} \times \text{length}^2)\]

A Lagrange density \( L \) is \( \frac{\text{[Energy]}}{\text{[Length]^3}} \)

\[ F_{\mu\nu} \sim \frac{\text{[E]}}{\text{[Length]}} \]

so \( \text{[E]} \sim \frac{1}{\text{[Length]^2}} \)

and \( \text{[E]} \) is dimensionless \( \left[ \frac{\text{E}^2}{\text{hc}^2} = \frac{1}{137} \right] \)

\[ L = (\text{dimensionless } #) \cdot F_{\mu\nu} F^{\mu\nu} + \ldots \]

\[ \frac{1}{\text{[Length]^4}} \]

What about other terms? Gauge invariance + Lorentz invariance require \( L = L \left( F_{\mu\nu} F^{\mu\nu} \right) \) or \( F F \)

and so if we imagine \( L \) is a polynomial in \( F \), each new term must be multiplied by a parent 

of dimension \( \text{[Length]^4} \) or \( \frac{1}{\text{[some mass]^4}} = \frac{1}{M^4} \]

What is the mass? If we are thinking about electrodynamics, the "mass" is the scale at which new physics modifies Maxwell's equations. The simplest new physics is the fundamentally quantum process of the exchange of virtual particles as intermediate...
strays more (production of \( e^+ e^- \) pairs) and so \( \lambda \sim m_e \). Now in wave number/momentum space,
\[
\mathcal{F}^2 = \frac{1}{\Lambda^4} \sim \frac{k^4}{\Lambda^4}
\]
(see classical \( \lambda = 2\pi k \approx k \Lambda \))

so for "typical momenta" \( k \ll \Lambda \) (= \( m_e \approx 1/2 \text{ MeV} \)) such processes are completely negligible. And of course at \( k \sim m_e \), these new processes turn on, and a classical description fails completely. Of course it has already failed far below that - in the actual radiation processes in atoms.

Indeed a calculation to support this statement was done by 2 of Heisenberg's students, Euler and Kockel, in 1935 - they started with QED (photons + electrons) and integrated the electrons out to derive an effective theory of photons.

\[
\mathcal{L} = \frac{1}{2} (E^2 - B^2) + \frac{e^4}{360 \pi^2 m_e^4} \left[ \frac{(E^2 - B)^2}{m_e^2} + 7(EB)^2 \right] + \ldots
\]

valid for frequencies \( h \omega \ll m_e c^2 \) (so \( \omega \ll m_e c^2/m_e^2 \)), which is \((\text{MeV to GeV})\) big! The new term includes "photon splitting" \( 4 \times 2 \) interactions as explicit nonlinear terms in \( \mathcal{L} \).

QED is itself an incomplete description of reality at very high energy, too - it is only a part of a continuum interaction "electroweak" interaction.