Construction of Field Theories

Lecture 24

Physics 411
Classical Mechanics II

October 29th, 2007

We are beginning our final descent, and I'll take the opportunity to look at the freedom we have in designing rational (by which I mean relatively sane) field theories out of Lagrangians[1]. Starting with the scalar fields, there is not much to say: at quadratic order we have a mass term and quadratic derivative term that recovers massive Klein-Gordon scalars. For vector theories like E&M, there are really only two options – massive (with Lorentz gauge fixed) and massless (with gauge freedom), and we count the number of valid (slash useful) quadratic terms for vector fields, leading to a tabulation of "all" free vector fields that support superposition (and at most second derivatives in the field equations).

24.1 Available Terms for Scalar Fields

We want to create a free field, what constructs are available? If we demand the field equations support superposition (linearity) so that the sum of two free fields is also a field, then we must have no more than quadratic terms in the Lagrangian. If we want wave-like solutions, then there should be second derivatives in the field equations and hence quadratic first-derivatives in the Lagrangian.

Finally, to support general covariance, our $\bar{\mathcal{L}}$ must be a scalar, then $\mathcal{L} \equiv \sqrt{-g}\,\bar{\mathcal{L}}$ is, appropriately, a density. and the only option is

$$\bar{\mathcal{L}}_s = \alpha\,\phi_{,\mu}\,g^{\mu\nu}\,\phi_{,\nu} + \beta\,\phi^2. \tag{24.1}$$

[1]For a comprehensive discussion of terms and model-building for relativistic field theory, see N. Wheeler, *Introduction to the Principles of Relativistic Field Theory*, pp. 1–32 (and the rest of it!). The particular massless E&M setting is described in V. Rubakov, *Classical Theory of Gauge Fields.*

So the creation of scalars also implies the metric as a fundamental building block. From here, first order form follows naturally, and all of this is clearly just the Klein-Gordon equation for massive scalar fields.

24.2 Available Terms for Vector Fields

Vector fields allow more options. Our two available terms from the scalar case become, for vector A_μ: $A_\mu\, A^\mu$ (mass term), $A_{\mu,\nu}\, A^{\mu}{}_{,}{}^{\nu}$, $A_{\mu,\nu}\, A^{\nu}{}_{,}{}^{\mu}$ and $(A^{\alpha}_{,\alpha})^2$. So we can imagine a general Lagrangian of the form

$$\bar{\mathcal{L}}_v = \alpha\, A_\mu\, A^\mu + \beta\, A_{\mu,\nu}\, A^{\mu}{}_{,}{}^{\nu} + \gamma\, A_{\mu,\nu}\, A^{\nu}{}_{,}{}^{\mu} + \delta\, (A^{\mu}{}_{,\mu})^2. \tag{24.2}$$

The non-derivative term is pretty clearly a "mass" term, of the usual $m^2\, \phi^2$ variety, and we know how this works in the field equations – so set $\alpha = 0$ for now, and take the rest. The field equations, after variation, are

$$2\,(\gamma + \delta)\, \partial^\nu \partial^\mu\, A_\mu + 2\, \beta\, \partial^\mu \partial_\mu\, A^\nu = 0, \tag{24.3}$$

(here and in what follows, we take the usual Minkowski metric as our background, although this is not required, it makes the notation simpler) which seems reasonable.

There is a slight problem, or rather, we will *create* a slight problem for ourselves. At issue is the nature of objects of the form A_μ – we want to talk about "real" vectors, here, and not derivatives of scalars. From tensor notation, it is clear that there is a "derivative" (in both senses of the word) vector we can make, $\psi_{,\mu}$. This is really just a scalar, and our goal is a true vector field theory. So what we'd like to do is get rid of any potential (no deep significance intended) scalar derivative polluting our new theory – we already dealt with scalar fields above. Our target is a set of field equations that make no reference to the scalar portion of a vector at all.

Recall from three-dimensional vector calculus that a vector, like $\mathbf{E}$ is determined from its divergence and curl – generically $\mathbf{E} = \nabla \times \mathbf{F} + \nabla G$, and then $\nabla \times \mathbf{E} = \nabla^2 G$ so that $\nabla^2 G$ is the divergence of $\mathbf{E}$ and $\nabla \times \mathbf{E} = \nabla\,(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}$ is the curl of $\mathbf{E}$. This breakup persists in four-dimensional vector fields, and in particular, we are interested in the $\mathbf{F}$ part rather than the ∇G part. Just as $\nabla \cdot \mathbf{E} = 0$ implies a purely solenoidal potential in three dimensions, the requirement $\partial_\mu A^\mu = 0$ (Lorentz gauge) indicates that there is no generic scalar gradient lurking in A^μ.

We will take a projective approach – breaking the full A^μ into a portion that has no $\psi_{,\mu}$ and a pure $\psi_{,\mu}$ part, then require that the field equation be insensitive to the scalar, and work backwards to the form of the action that generates this. Start by assuming we can write $A_\mu = \bar{A}_\mu + \psi_{,\mu}$ (in general, we can) – then the field equation becomes

$$2\,(\gamma+\delta)\,\left[\partial^\nu\partial^\mu\,\bar{A}_\mu + \partial^\nu\partial^\mu\partial_\mu\,\psi\right] + 2\,\beta\,(\partial_\mu\partial^\mu\,\bar{A}^\nu + \partial^\mu\partial_\mu\partial^\nu\,\psi) = 0, \tag{24.4}$$

and if we want no reference to ψ, so that the field equation refers only to the $\bar{A}_\mu$ contribution, then the requirement is $\beta + \gamma = -\delta$. Again, the logic here is that we have no scalar field information in our initial Lagrangian, so we have not coupled (accurately or not) the scalar ψ to the vector $\bar{A}_\mu$, and we want to make this lack of information explicit in the field equations.

Then the original Lagrangian reads (again, ignoring the mass term)

$$\bar{\mathcal{L}}_v = \beta\,A_{\mu,\nu}\,A^{\mu}{}_{,}{}^{\nu} + \gamma\,A_{\mu,\nu}\,A^{\nu}{}_{,}{}^{\mu} - (\beta+\gamma)\,(A^{\mu}{}_{,\mu})^2. \tag{24.5}$$

If we assume the field A_μ falls off at infinity, then we can add any "pure divergence" to $\bar{\mathcal{L}}$ without changing the field equations. Remember that $\bar{\mathcal{L}}$ is part of the larger action structure:

$$\boxed{S = \int d\tau\,\sqrt{-g}\,\bar{\mathcal{L}} = \int d\tau\,\sqrt{-g}\,(\bar{\mathcal{L}} + \partial_\mu\,F^\mu)} \tag{24.6}$$

for any F^μ depending on A_μ, provided, again, that $A_\mu \to 0$ at infinity (so the surface term vanishes).

As a particular function, take $F^\mu = (A^{\mu}{}_{,\nu}\,A^\nu - A^\mu\,A^{\nu}{}_{,\nu})$, then from the point of view of $\bar{\mathcal{L}}$, $\partial_\mu\,F^\mu$ "=" 0. Formally, this divergence is

$$\begin{aligned}\partial_\mu\,F^\mu &= (\partial_\mu\partial_\nu\,A^\mu)\,A^\nu + \partial_\nu\,A^\mu\,\partial_\mu\,A^\nu - \partial_\mu\,A^\mu\,\partial_\nu\,A^\nu - A^\mu\,\partial_\mu\partial_\nu\,A^\nu \\ &= \partial_\nu\,A^\mu\,\partial_\mu\,A^\nu - \partial_\mu\,A^\mu\,\partial_\nu\,A^\nu\end{aligned} \tag{24.7}$$

and we can add this with impunity to the Lagrangian – suppose we add $-(\beta+\gamma)\,(A_{\mu,\nu}\,A^{\nu}{}_{,}{}^{\mu} - A^{\mu}{}_{,\mu}\,A^{\nu}{}_{,\nu})$ which is of the above form, to the revised $\bar{\mathcal{L}}$ in (24.5), the clear choice for killing off the $(A^{\alpha}{}_{,\alpha})^2$ term:

$$\bar{\mathcal{L}}_v = \beta\left(A_{\mu,\nu}\,A^{\mu}{}_{,}{}^{\nu} - A_{\mu,\nu}\,A^{\nu}{}_{,}{}^{\mu}\right) \tag{24.8}$$

and we have finally arrived at the antisymmetric derivative Lagrangian we began with in our study of E&M.

Evidently, the antisymmetry of the field-strength tensor is somehow related to its "pure vector" character. This procedure also introduces, explicitly, the gauge invariance of A_μ – its field equations and Lagrangian are totally independent of scalar gradients, and only the vector part contributes to the physics of electricity and magnetism.

Once we have (24.8), the conjugate momenta is clear – take $F^{\mu\nu} \equiv \frac{\partial \bar{\mathcal{L}}}{\partial A_{\mu,\nu}} = \beta \left(A^{\mu}{}_{,}{}^{\nu} - A^{\nu}{}_{,}{}^{\mu}\right)$, the usual. First order form follows from the Legendre transform as always.

24.2.1 Mass Term

One can, with the above Lagrangian in place, introduce a mass term in the usual way – take

$$\bar{\mathcal{L}}_{Vm} = \frac{1}{2}\,\alpha\,F^{\mu\nu}\left(A_{\nu,\mu} - A_{\mu,\nu}\right) - \beta\,F^2 - \frac{1}{2}\,m^2\,A^\mu\,A_\mu \tag{24.9}$$

then variation gives

$$\begin{aligned} \beta\,F_{\mu\nu} &= \alpha\left(A_{\nu,\mu} - A_{\mu,\nu}\right) \\ -\alpha\,F^{\mu\nu}{}_{,\nu} - m^2\,A^\mu &= 0. \end{aligned} \tag{24.10}$$

It is interesting that, by taking the divergence again, we get:

$$\partial_\mu \partial_\nu\,F^{\mu\nu} + m^2\,\partial_\mu\,A^\mu = m^2\,\partial_\mu\,A^\mu = 0 \tag{24.11}$$

which is precisely Lorentz gauge. For massive A^μ (a form of Proca's equation), we are required, for internal consistency, to have a divergenceless field. This is equivalent to the condition that there can be *no* scalar divergence associated with A_μ – just as $\nabla \cdot \mathbf{F} = 0$ implies that $\mathbf{F}$ is a pure curl with no $\nabla\psi$ admixture. So, far from releasing us from scalar dependence (by making no reference to it at all), the massive vector theory requires that the gradient be omitted.

It was precisely this gauge freedom for A_μ, we remember, that led to charge conservation when we coupled to a source j^μ – we have now lost that freedom – we are forced into Lorentz gauge, and the relevant field equations in that setting will be

$$F^{\mu\nu}{}_{,\nu} - \left(m^2\,A^\mu + \gamma\,j^\mu\right) = 0, \tag{24.12}$$

from which we learn only that the combination $\sim A^\mu + j^\mu$ must be divergenceless.

All of this is to say that massive vector fields are quite different from massless vector fields. In the language of field theory, we call massless A^μ a "gauge field", and it is responsible for light, for example. The massive vector A^μ is *not* a gauge field, as it lacks the gauge independence, and does not lead to the theory of E&M.

24.3 One Last Ingredient

There is a final building block, which I hesitate to mention. We know that for vector fields, for example, A_μ, $A_{\mu,\nu}$ and $g_{\mu\nu}$ are the major players, but there is a further element. The Levi-Civita tensor density can also be used to develop Lagrangians. We define, *in any coordinate system* the numerical symbol:

$$\epsilon^{\alpha\beta\gamma\delta} \equiv \begin{cases} 1 & \text{for indices in "even" order} \\ -1 & \text{for indices in "odd" order} \\ 0 & \text{for any repeated index} \end{cases} \tag{24.13}$$

where we take some canonical ordering (if $\{\alpha, \beta, \gamma, \delta\}$ are in $0 \rightarrow 3$, for example, we might take $\epsilon^{0123} = 1$) as the base ordering and count even and odd (number of flips) to get any given ordering. This is the antisymmetric "symbol", it doesn't transform as a fourth-rank contravariant tensor because we require it to be numerically identical in all coordinate systems. What, then, is it?

Suppose we treat $\epsilon^{\alpha\beta\gamma\delta}$ as a fourth-rank contravariant tensor density with weight p, then

$$\epsilon'^{\alpha\beta\gamma\delta} = \det\left(\frac{\partial x'}{\partial x}\right)^p \frac{\partial x'^\alpha}{\partial x^\rho} \frac{\partial x'^\beta}{\partial x^\sigma} \frac{\partial x'^\gamma}{\partial x^\mu} \frac{\partial x'^\delta}{\partial x^\nu} \epsilon^{\rho\sigma\mu\nu}, \tag{24.14}$$

and consider the object:

$$\frac{\partial x'^\alpha}{\partial x^\rho} \frac{\partial x'^\beta}{\partial x^\sigma} \frac{\partial x'^\gamma}{\partial x^\mu} \frac{\partial x'^\delta}{\partial x^\nu} \epsilon^{\rho\sigma\mu\nu}. \tag{24.15}$$

This, as an expression, is totally antisymmetric in $\{\alpha, \beta, \gamma, \delta\}$, and so must be proportional to $\epsilon^{\alpha\beta\gamma\delta}$. Let the proportionality constant be A, then to find the value for A, note that the determinant of a matrix M with elements M_{ij} can be written as

$$\det M = \epsilon^{i_1 i_2 i_3 i_4}\, m_{1i_1}\, m_{2i_2}\, m_{3i_3}\, m_{4i_4} \tag{24.16}$$

in four dimensions. We only need one value for the object in (24.15) to set the constant A, so take:

$$\frac{\partial x'^0}{\partial x^\rho} \frac{\partial x'^1}{\partial x^\sigma} \frac{\partial x'^2}{\partial x^\mu} \frac{\partial x'^3}{\partial x^\nu} \epsilon^{\rho\sigma\mu\nu} = \det\left(\frac{\partial x'}{\partial x}\right) \epsilon^{0123} = \det\left(\frac{\partial x'}{\partial x}\right), \tag{24.17}$$

and the transformation for ϵ reads

$$\epsilon'^{\alpha\beta\gamma\delta} = \det\left(\frac{\partial x'}{\partial x}\right)^p \left(\det\left(\frac{\partial x'}{\partial x}\right) \epsilon^{\alpha\beta\gamma\delta}\right), \tag{24.18}$$

and if we insist on $\epsilon'^{\alpha\beta\gamma\delta} = \epsilon^{\alpha\beta\gamma\delta}$ then the Levi-Civita symbol must be a tensor density of weight $p = -1$.

I mention all of this because it is possible, on the vector side, to introduce a fourth invariant – the scalar density $\epsilon^{\alpha\beta\gamma\delta} A_{\beta,\alpha} A_{\delta,\gamma}$ – and we are honor bound to find its role in the Lagrangian and field equations. It suffices to look again to the massless theory.

$$\mathcal{L}_V = \alpha \sqrt{-g} \left(A_{\mu,\nu} \left(A^{\mu}{}_{,}{}^{\nu} - A^{\nu}{}_{,}{}^{\mu}\right)\right) + \beta \epsilon^{\alpha\beta\gamma\delta} A_{\beta,\alpha} A_{\delta,\gamma} \tag{24.19}$$

but the Levi-Civita term in this setting can be written as a total divergence – namely

$$\begin{aligned} \partial_\mu \left(\epsilon^{\mu\nu\alpha\beta} \left(A_\nu A_{\beta,\alpha}\right)\right) &= \epsilon^{\mu\nu\alpha\beta} \left(A_{\nu,\mu} A_{\beta,\alpha} + A_\nu A_{\beta,\alpha\nu}\right) \\ &= \epsilon^{\mu\nu\alpha\beta} \left(A_{\nu,\mu} A_{\beta,\alpha}\right), \end{aligned} \tag{24.20}$$

where the second term in the first line dies under $\alpha \leftrightarrow \nu$ symmetry of partials hitting the antisymmetry of $\epsilon^{\mu\nu\alpha\beta}$.

The Levi-Civita contribution, then, is not a contribution at all. We have gained nothing by its introduction in $D = 4$ – this is not always the case.