Maxwell's coastal battle, Dirac's quirks, and Bergmann's algorithm

Stefanus Koesno[1]

[1] Somewhere in California

San Jose, CA 95133 USA

(Dated: January 30, 2016)

Abstract

Students usually learn about metric and covariant formulation at the end of an electrodynamics class. However, once students migrated to covariant notation things switch immediately to relativity, creating an unspoken dichotomy and very little effort is given on how this new notation relate to electrodynamics.

At the quantum front, it is seldom that Dirac's algorithm to deal with constraints, which is a great achievement, is mentioned at all. Details of its application are usually not shown, especially when applied to Dirac's equation. Usual applications involve Maxwell's equations but details are sparse and assumptions are often surresptitiously inserted.

This article is meant to fill in the gaps between covariant formulation and electrodynamics and explain calculational steps and implicit assumptions in applying Dirac-Bergmann algorithm.

I. PRELUDE

This article is a result of a failed research project. Despite the failure I learned a tremendous amount about electrodynamics. There are apparently many ambiguous things that instructors do not give much attention to and students take for granted. One example is the Dirac-Bergmann algorithm which is seldom discussed in class, including all of its incertitudes which we will discuss in details. Another is the simple choice of metric which we usually ignore once we switch from the traditional formulation of Maxwell's equations to the covariant formulation, which will be our first discussion topic so let the best coast battle continue.

II. EAST VS WEST

Which metric reigns supreme? Most people ignore this or at least think this is a non issue because who deals with minus sign anymore or a factor of two, but it does make a difference. Consider for example the covariant Maxwell's equations (in tensor notation)

$$\partial_\mu F^{\mu\nu} = J^\nu$$
$$\text{west coast} \rightarrow \partial_0 F_0{}^\nu - \partial_i F_i{}^\nu = J^\nu$$
$$\text{east coast} \rightarrow -\partial_0 F_0{}^\nu + \partial_i F_i{}^\nu = J^\nu$$

Using the wrong metric apparently generates the wrong minus sign. Now of course this is with the consensus that

$$x^\mu = \{x^0, x^i\}$$
$$\frac{\partial}{\partial x^u} = \partial_\mu$$

for both metrics. This metric conundrum apparently doesn't affect the Maxwell lagrangian $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ because we are contracting two indices. In the case of the scalar field, it does make a difference since $\partial_\mu \phi \partial^\mu \phi$ only contracts *one* index

$$\partial_\mu \phi \partial^\mu \phi = \ \partial_0 \phi \partial_0 \phi - \partial_i \phi \partial_i \phi \quad \text{west coast}$$
$$= -\partial_0 \phi \partial_0 \phi + \partial_i \phi \partial_i \phi \quad \text{east coast}$$

2

You might think that everything is nice and dandy in Maxwell's land then, but the source term $J^\mu A_\mu$ does introduce a metric ambiguity. Again we have to agree that the source is $J^\mu = \overline{\psi}\gamma^\mu\psi$ for both metrics, sneaking minus signs is not allowed, *i.e.* you can't say that $J^\mu = \overline{\psi}\gamma^\mu\psi$ in the west and $J^\mu = -\overline{\psi}\gamma^\mu\psi$ in the east. On this note, if we can also think about the current as simply given by $J^\mu = \{\rho, \vec{J}\}$.

This might sound a bit harsh but the concept of vector spaces is independent of the concept of metric, in fact we add the metric concept into the vector space we want to work on to generate the concept of distance but vector spaces by themselves work alright without any concept of distance.

We first choose what vector spaces we want to work on and then add the concept of distance on top of it, so we first agree on the set of vectors despite the choice of metric later on.

Personally, east coast metric makes more sense because any three dimensional vector can be written as a four vector with the time component set to zero, $\vec{A} = \{0, A^i\}$, and the length of a 3d vectors in the east is naturally positive $\vec{A} \cdot \vec{A} = -0^2 + A^i A^i$ which is intuitive.

Going back to our Maxwell Lagrangian, we need to set our convention straight and be consistent. For this article I will use the west coast metric, since I'm currently living on the west coast.

Starting from the beginning, the potential vector is given by

$$A^\mu = \{\phi, \vec{A}\}$$
$$= \{\phi, A^i\}$$
$$A_\mu = \{\phi, -A^i\}$$

the next confusing thing is that the electric field is usually given in our beloved textbooks as

$$\vec{E} = -\partial_t \vec{A} - \nabla\phi$$

and neither the metric choice nor the contra/co varianness mentioned at all. Say we

choose

$$E_i = -\partial_0 A_i - \partial_i \phi$$

$$= -\partial_0 A_i - \partial_i A_0$$

we get the wrong result as we know that the $\vec{E}$ should be given by the electromagnetic tensor which is antisymmetric $F^{i0} = \partial^i A^0 - \partial^0 A^i = E^i$. The above situation can only be salvaged if we use both upper and lower indices in an inconsistent way, *e.g.*

$$E_i = -\partial_0 A^i - \partial_i A_0$$

$$E_i = \partial_0 A_i - \partial_i A_0$$

This is immediately obvoius since each term has one time and one spatial component and raising/lowering both indices either using both east and west metric will not change the sign of the whole term, we must have mixed upper and lower indices. Thus what we have been using all along is a case of a non tensorial equation masquerading as one (which will be very confusing for someone who has just learned GR when he looks back on EM).

We will then stick to the convention usually found in other literatures

$$F^{i0} = \partial^i A^0 - \partial^0 A^i = E^i$$

$$F^{ij} = \partial^i A^j - \partial^j A^i = -\varepsilon^{ijk} B_k$$

What this means is that the electric field is given by

$$E^i = -\partial_i A^0 - \partial^0 A^i \quad \text{west coast}$$

$$\to \nabla \phi = \partial_i A^0 \quad \partial_t \vec{A} = \partial^0 A^i$$

$$E^i = -\partial_i A_0 - \partial^0 A^i \quad \text{east coast}$$

$$\to \nabla \phi = \partial_i A_0 \quad \partial_t \vec{A} = \partial^0 A^i$$

and the magnetic field is

$$\partial^i A^j - \partial^j A^i = (\delta^i_m \delta^j_n - \delta^j_m \delta^i_n) \partial^m A^n$$

$$-\varepsilon^{ijk} B_k = -\varepsilon^{ijk} \varepsilon_{kmn} \partial^m A^n$$

$$\varepsilon_{kmn} \partial^m A^n = B_k$$

$$\to \varepsilon_{kmn} \partial^m A^n = (\nabla \times \vec{A})_k$$

The minus sign in line 2 on the right hand side is due to the fact that the levi civita symbol has three indices, so raising three of them using the west coast metric introduces an overal minus sign, therefore

$$\varepsilon^{123} = -1$$

$$\varepsilon_{123} = 1$$

$$-\varepsilon^{ijk}\varepsilon_{mnk} = (\delta^i_m \delta^j_n - \delta^j_m \delta^i_n)$$

$$-\varepsilon^{ijk}\varepsilon_{mjk} = 2\delta^i_m$$

$$-\varepsilon^{ijk}\varepsilon_{ijk} = 6$$

Apart from the minus sign up front the case for $\vec{B}$ is apparently easier since $F_{ij} = F^{ij}$ no matter what coast you are on since you are raising or lowering two free indices. An unavoidable conclusion one draws from this is that the relationship between the traditional convention used in textbooks and the covariant formulation of electromagnetic lagrangian can be adjusted at will, as long as one states it clearly and be consistent.

In conclusion, the choice of metric does matter, piling on top of this conundrum is tha fact that in the normal course of electrmagnetics we use a mishmash convention of upper and lower indices. This effect will be more pronounced when we do a massive photon lagrangian as we shall see later. For convenience the full dictionary between the electromagnetic tensor and the conventional electric and magnetic fields is given below

$$F^{i0} = E^i$$

$$F_{i0} = E_i$$

$$F^{ij} = -\varepsilon^{ijk}B_k$$

$$F_{ij} = -\varepsilon_{ijk}B^k$$

$$\varepsilon_{kmn}\partial^m A^n = (\nabla \times \vec{A})_k$$

$$\varepsilon^{kmn}\partial_m A_n = (\nabla \times \vec{A})^k$$

III. COVARIANT ELECTROMAGNETIC LAGRANGIANS AND THEIR EQUATIONS OF MOTION

What is the electromagnetic lagrangian? In the usual classical mechanics context the lagrangian is given by $T - V$, kinetic minus potential energy but for the EM field what are its kinetic and potential terms? They are actually just $\vec{E}$ and $\vec{B}$, the electric and magnetic fields themselves with $T = \frac{1}{2}\vec{E}^2$ and $V = \frac{1}{2}\vec{B}^2$. I have not found any satisfying argument showing why this is so except that they correctly reproduce Maxwell's equation. Let's see how they come about from the usual curvature, $F^{\mu\nu}$, formulation

$$
\begin{aligned}
\mathcal{L} &= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \\
&= -\frac{1}{4}\left(F^{0\nu}F_{0\nu} + F^{i\nu}F_{i\nu}\right) \\
&= -\frac{1}{4}\left(F^{0j}F_{0j} + \left(F^{i0}F_{i0} + F^{ij}F_{ij}\right)\right) \\
&= -\frac{1}{4}\left((-E^j)(-E_j) + \left((E^i)(E_i) + \varepsilon_{ijk}B^k\varepsilon^{ijl}B_l\right)\right) \\
&= -\frac{1}{4}\left(2E^jE_j - 2\delta^l_k B^k B_l\right) \\
&= -\frac{2}{4}\left(E^j(-E^j) - B^l(-B^l)\right) \\
&= -\frac{1}{2}\left(-\vec{E}\cdot\vec{E} + \vec{B}\cdot\vec{B}\right) \\
\mathcal{L} &= \frac{1}{2}\left(\vec{E}^2 - \vec{B}^2\right)
\end{aligned}
$$

again our choice here is the west coast metric with a non-conventional levi-civita symbol $\varepsilon_{ijk} = -\varepsilon^{ijk}$. One note about the dot product of a three vector, it is still given by the conventional $\vec{E}^2 = \vec{E}\cdot\vec{E} = E^i E^i$.

The above lagrangian will give us the usual hamiltonian, *i.e.* the sum of the electric and magnetic field strength, which we will see later. But before moving on to calculate its equations of motion, let us discuss other forms of the electromagnetic lagrangian. One popular form is to just use one term from $F_{\mu\nu}$. We can do this because we can split a tensor to its symmetric and antisymmetric terms

$$
\begin{aligned}
\partial_\mu A_\nu &= \left\{\frac{1}{2}(\partial_\mu A_\nu + \partial_\nu A_\mu) + \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu)\right\} \\
&= \partial_{\{\mu}A_{\nu\}} + \partial_{[\mu}A_{\nu]}
\end{aligned}
$$

and since $F^{\mu\nu}$ is antisymmetric

$$
\begin{aligned}
\partial_\mu A_\nu F^{\mu\nu} &= (\partial_{\{\mu} A_{\nu\}} + \partial_{[\mu} A_{\nu]}) F^{\mu\nu} \\
&= \partial_{[\mu} A_{\nu]} F^{\mu\nu} \\
&= \frac{1}{2} F_{\mu\nu} F^{\mu\nu}
\end{aligned}
$$

where we have used the fact that the contraction of symmetric and antisymmetric tensors vanish and that the antisymmetric part of $\partial_\mu A_\nu$ is just $\partial_{[\mu} A_{\nu]} = (1/2) F_{\mu\nu}$.

If we substitute this back into our original lagrangian we get

$$
\begin{aligned}
\mathcal{L} &= -\frac{1}{2} \partial_\mu A_\nu F^{\mu\nu} \\
&= -\frac{1}{2} \partial_\mu A_\nu \left(\partial^\mu A^\nu - \partial^\nu A^\mu \right)
\end{aligned}
$$

What is the advantage of writing the lagrangian this way? Mostly it's for quantum field theory use, the calculation of the propagator. A bigger advantage is that we get the equations of motion immediately, we only need to do an integration by parts such that

$$
\begin{aligned}
-\frac{1}{2} \partial_\mu A_\nu \left(\partial^\mu A^\nu - \partial^\nu A^\mu \right) &= \frac{1}{2} A_\nu \left(\partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu \right) \\
&= \frac{1}{2} A_\nu \left(\delta^\nu_\mu \partial^2 - \partial_\mu \partial^\nu \right) A^\mu \\
&= \frac{1}{2} A_\nu \left(g^{\nu\mu} \partial^2 - \partial^\mu \partial^\nu \right) A_\mu
\end{aligned}
$$

We can also employ this method to scalar fields

$$
\begin{aligned}
\frac{1}{2} \partial^\mu \phi \partial_\mu \phi - \frac{1}{2} m^2 \phi^2 &= \frac{1}{2} (-\phi \partial^2 \phi - m^2 \phi^2) \\
&= -\frac{1}{2} \phi (\partial^2 + m) \phi
\end{aligned}
$$

and whatever is in brackets is the Klein-Gordon equation, pops out of the lagrangian itself.

In fact, we can go backwards, starting from the propagator, to construct the lagrangian. This is the approach chosen by Zee in his book Quantum Field Theory in a Nutshell which he called bypassing Maxwell, it is somewhat of a shortcut although it is not too much of shortcut in my opinion. One thing you can do, if you want to go this route, is to remember that the difficulty with photons is gauge invariance and therefore the terms sandwiched

by the the A_μ's in the lagrangian must have no inverse if this is so. One candidate is of course

$$-k^2 g^{\mu\nu} + k^\mu k^\nu \rightarrow \partial^2 g^{\mu\nu} - \partial^\mu \partial^\nu$$

because in QFT it is always related to the momentum k^μ. And it doesn't have an inverse because

$$(-k^2 g^{\mu\nu} + k^\mu k^\nu)k_\nu = -k^2 k^\mu + k^\mu k^2$$
$$= 0$$

it has a non zero vector that it annihilates. The tricky thing is figuring out the minus sign up front which you can read all about in Zee's book.

Going back to our lagrangian, this particular form does have an advantage because it's somewhat easier to get the equations of motion, $i.e.$ whatever is sandwiched by the two A's

$$\delta S = \int d^4x \; \frac{1}{2}\delta A_\nu(g^{\nu\mu}\partial^2 - \partial^\mu\partial^\nu)A_\mu + \frac{1}{2}A_\nu(g^{\nu\mu}\partial^2 - \partial^\mu\partial^\nu)\delta A_\mu - J^\mu \delta A_\mu$$
$$= \int d^4x \; \frac{1}{2}\delta A_\mu(g^{\mu\nu}\partial^2 - \partial^\nu\partial^\mu)A_\nu + \frac{1}{2}A_\nu(g^{\nu\mu}\overleftarrow{\partial}^2 - \overleftarrow{\partial^\mu\partial^\nu})\delta A_\mu - J^\mu \delta A_\mu$$
$$= \int d^4x \; \delta A_\mu \left((g^{\mu\nu}\partial^2 - \partial^\nu\partial^\mu)A_\nu - J^\mu\right)$$

in the second line we have swapped $\mu \leftrightarrow \nu$ in the first term since they are just dummy indices and we did integration by parts on the second (the arrows on top of the derivatives show that they apply to the A on the left), the e.o.m can then be read off as

$$(g^{\mu\nu}\partial^2 - \partial^\nu\partial^\mu)A_\nu = J^\mu$$

Here we have included the source J^μ and chosen to add the term $-J^\mu A_\mu$ instead of $J^\mu A_\mu$ in the lagrangian. Now this is nothing more than

$$(g^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu)A_\nu = \partial^2 A^\mu - \partial^\nu\partial^\mu A_\nu$$
$$= \partial_\nu\partial^\nu A^\mu - \partial_\nu\partial^\mu A^\nu$$
$$= \partial_\nu(\partial^\nu A^\mu - \partial^\mu A^\nu)$$
$$J^\mu = \partial_\nu F^{\nu\mu}$$

8

which in turn is nothing more than

$$\partial_\nu F^{\nu\mu} = J^\mu$$

$$\partial_0 F^{0\mu} + \partial_i F^{i\mu} = J^\mu$$

$$(\partial_0 F^{00} + \partial_0 F^{0k}) + (\partial_i F^{i0} + \partial_i F^{ik}) = J^\mu, \quad F^{00} = 0$$

$$\partial_i F^{i0} + (\partial_0 F^{0k} + \partial_i F^{ik}) = J^0 + J^k$$

$$\rightarrow \partial_i F^{i0} = J^0$$

$$\rightarrow \partial_0 F^{0k} + \partial_i F^{ik} = J^k$$

Now we can check whether we have chosen the correct sign for $J^\mu A_\mu$ in the lagrangian. The first e.o.m related to J^0 is straightforward

$$\partial_i F^{i0} = J^0$$

$$\partial_i E^i = \rho$$

$$\rightarrow \nabla \cdot \vec{E} = \rho$$

which gauss's law. The next e.o.m is

$$\partial_0 F^{0k} + \partial_i F^{ik} = J^k$$

$$\partial_0(-E^k) + \partial_i(-\varepsilon^{ikm} B_m) = J^k$$

$$-\varepsilon^{ikm} \partial_i B_m = \partial_0 E^k + J^k$$

$$\varepsilon^{kim} \partial_i B_m = \partial_0 E^k + J^k$$

$$(\nabla \times \vec{B})^k = \partial_0 E^k + J^k$$

$$\rightarrow (\nabla \times \vec{B}) = \partial_0 \vec{E} + \vec{J}$$

which is ampere's law. So we added the correct source term in the lagrangian $-J^\mu A_\mu$.

Now there are no more equations of motion we can derive from the lagrangian, instead the rest of Maxwell's equations is given by bianchi identity

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = \partial_\lambda \partial_\mu A_\nu - \partial_\lambda \partial_\nu A_\mu + \partial_\mu \partial_\nu A_\lambda - \partial_\mu \partial_\lambda A_\nu + \partial_\nu \partial_\lambda A_\mu - \partial_\nu \partial_\mu A_\lambda$$

$$= (\partial_\lambda \partial_\mu A_\nu - \partial_\mu \partial_\lambda A_\nu) + (-\partial_\lambda \partial_\nu A_\mu + \partial_\nu \partial_\lambda A_\mu) + (\partial_\mu \partial_\nu A_\lambda - \partial_\nu \partial_\mu A_\lambda)$$

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0$$

note that none of the indices are summed. Let's start by setting $\lambda = 0, \mu = i, \nu = j$ and using our usual dictionary $F_{0i} = E^i = -E_i$, $F_{lm} = -\varepsilon_{ilm}B^i$, $F^{lm} = -\varepsilon^{ilm}B_i$

$$
\begin{aligned}
0 &= \partial_0 F_{ij} + \partial_i F_{j0} + \partial_j F_{0i} \\
&= -\partial_0 \varepsilon_{kij} B^k + \partial_i E_j - \partial_j E_i \\
&= -\varepsilon^{lij}\varepsilon_{kij}\partial_0 B^k + \varepsilon^{lij}(\partial_i E_j - \partial_j E_i) \\
&= 2\delta_k^l \partial_0 B^k + \varepsilon^{lij}\partial_i E_j - \varepsilon^{lij}\partial_j E_i \\
&= 2\partial_0 B^l + 2\varepsilon^{lij}\partial_i E_j \\
\rightarrow 0 &= \partial_0 B^l + \varepsilon^{lij}\partial_i E_j
\end{aligned}
$$

note that if $i = j$ line 2 will become $0 = 0$ thus $i \neq j \neq k$ is a must, we can rewrite this result as

$$
\begin{aligned}
0 &= \partial_0 B^k + (\nabla \times \vec{E})^k \\
\rightarrow -\partial_0 \vec{B} &= \nabla \times \vec{E}
\end{aligned}
$$

Next we set $\lambda = i, \mu = j, \nu = k$ with $i \neq j \neq k$

$$
\begin{aligned}
0 &= \partial_i F_{jk} + \partial_j F_{ki} + \partial_k F_{ij} \\
&= -\partial_i \varepsilon_{ljk} B^l - \partial_j \varepsilon_{lki} B^l - \partial_k \varepsilon_{lij} B^l \\
&= -\varepsilon^{ijk}(\partial_i \varepsilon_{ljk} B^l + \partial_j \varepsilon_{lki} B^l + \partial_k \varepsilon_{lij} B^l) \\
&= 2\delta_l^i \partial_i B^l + 2\delta_l^j \partial_j B^l + 2\delta_l^k \partial_k B^l \\
&= 6 \; \partial_l B^l \\
0 &= \partial_l B^l \\
\rightarrow 0 &= \nabla \cdot \vec{B}
\end{aligned}
$$

Thus we recover the other sourceless Maxwell's equation, Voila! and since these are just bianchi identity, they are always true regardless of the presence or absence of sources.

IV. MAXWELL'S HAMILTONIAN AND DIRAC'S QUIRKS

A. The subterfuge in the fields

We are now ready to discuss the electromagnetic hamiltonian. First we need to add the fermionic part to the lagrangian

$$\mathcal{L} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu + \frac{1}{2}\partial_\mu A_\nu \partial^\nu A^\mu + i\overline{\psi}\gamma^\mu \partial_\mu \psi - \overline{\psi}\left(e\gamma^\mu A_\mu + m\right)\psi$$

$$= -\frac{1}{2}\partial_\mu A_\nu F^{\mu\nu} + i\overline{\psi}\gamma^\mu \partial_\mu \psi - \overline{\psi}\left(e\gamma^\mu A_\mu + m\right)\psi$$

First we need to derive the conjugate momenta, they are given by the usual formula

$$p_A^\sigma = \frac{\delta\mathcal{L}}{\delta\partial_0 A_\sigma} = -\frac{1}{2}\delta_\mu^0 \delta_\nu^\sigma \partial^\mu A^\nu - \frac{1}{2}\partial_\mu A_\nu g^{0\mu} g^{\sigma\nu} + \frac{1}{2}\delta_\mu^0 \delta_\nu^\sigma \partial^\nu A^\mu + \frac{1}{2}\partial_\mu A_\nu g^{0\nu} g^{\sigma\mu}$$

$$= -\frac{1}{2}\partial^0 A^\sigma - \frac{1}{2}\partial^0 A^\sigma + \frac{1}{2}\partial^\sigma A^0 + \frac{1}{2}\partial^\sigma A^0$$

$$= -\partial^0 A^\sigma + \partial^\sigma A^0$$

$$\to p_A^\sigma = -F^{0\sigma} = F^{\sigma 0}, \quad F^{i0} = E^i$$

$$\to p_\psi = \frac{\delta\mathcal{L}}{\delta\partial_0 \psi} = i\overline{\psi}\gamma^0$$

$$\to p_{\overline{\psi}} = \frac{\delta\mathcal{L}}{\delta\partial_0 \overline{\psi}} = 0$$

we immediately see two problems, there is no conjugate momenta for A_0 (since $F^{00} = 0$) and none for $\overline{\psi}$. The proper way of dealing with this is to employ Dirac-Bergmann algorithm but before we do that we will just do it the traditional way. The purpose of this exercise is to clarify what the problem really is.

The hamiltonian (density) is then

$$\mathcal{H} = p_A^\sigma \partial_0 A_\sigma + p_\psi \partial_0 \psi - \mathcal{L}$$

$$= E^i \partial_0 A_i + i\overline{\psi}\gamma^0 \partial_0 \psi + \frac{1}{2}\partial_\mu A_\nu F^{\mu\nu} - i\overline{\psi}\gamma^\mu \partial_\mu \psi + \overline{\psi}\left(e\gamma^\mu A_\mu + m\right)\psi$$

$$\mathcal{H} = \mathcal{H}_{\text{ph}} + \mathcal{H}_{\text{fm}}$$

$$\mathcal{H}_{\text{ph}} = E^i \partial_0 A_i + \frac{1}{2}\partial_\mu A_\nu F^{\mu\nu} + J^\mu A_\mu, \quad J^\mu = e\overline{\psi}\gamma^\mu \psi$$

$$\mathcal{H}_{\text{fm}} = i\overline{\psi}\gamma^0 \partial_0 \psi - i\overline{\psi}\gamma^\mu \partial_\mu \psi + \overline{\psi}\left(e\gamma^\mu A_\mu + m\right)\psi$$

where we have separated the hamiltonian into the photonic and fermionic parts for convenience. We first tackle the photonic part, focusing on the first two terms for now

$$
\begin{aligned}
E^i \partial_0 A_i + \frac{1}{2} \partial_\mu A_\nu F^{\mu\nu} &= E^i \partial_0 A_i + \frac{1}{2} \partial_0 A_\nu F^{0\nu} + \frac{1}{2} \partial_i A_\nu F^{i\nu} \\
&= E^i \partial_0 A_i + \frac{1}{2} \partial_0 A_i (-E^i) + \frac{1}{2} \partial_i A_\nu F^{i\nu}, \quad F^{00} = 0 \\
&= \frac{1}{2} E^i \partial_0 A_i + \frac{1}{2} \partial_i A_\nu F^{i\nu} \\
&= \frac{1}{2} E^i \partial_0 A_i + \frac{1}{2} \left(\partial_i A_0 E^i + \partial_i A_j F^{ij} \right) \\
&= \frac{1}{2} \left(E^i \partial_0 A_i + E^i \partial_i A_0 \right) + \frac{1}{4} F_{ij} F^{ij} \\
&= \frac{1}{2} E^i \left(\partial_0 A_i - \partial_i A_0 + 2 \partial_i A_0 \right) + \frac{1}{4} F_{ij} F^{ij} \\
&= \frac{1}{2} E^i (-E_i) + E^i \partial_i A_0 + \frac{1}{4} F_{ij} F^{ij} \\
&= \frac{1}{2} E^i E^i + E^i \partial_i A_0 + \frac{1}{4} F_{ij} F^{ij}
\end{aligned}
$$

the hamiltonian is then given by

$$
\begin{aligned}
\mathcal{H}_{\mathrm{ph}} &= \frac{1}{2} E^i E^i + E^i \partial_i A_0 + \frac{1}{4} F_{ij} F^{ij} + J^\mu A_\mu \\
&= \frac{1}{2} E^i E^i - \partial_i E^i A_0 + \frac{1}{4} F_{ij} F^{ij} + J^\mu A_\mu
\end{aligned}
$$

where we have done integration by parts on the second term to go to the second line which is permitted since the hamiltonian density is integrated, $H = \int d^3 x \, \mathcal{H}$. While we are here we can massage this hamiltonian further

$$
\begin{aligned}
F_{ij} F^{ij} &= (-\varepsilon_{ijk} B^k)(-\varepsilon^{ijl} B_l) = \varepsilon_{ijk} \varepsilon^{ijl} B^k B_l \\
&= -2 \delta_k^l B^k B_l = -2 B^k B_k = 2 B^k B^k \\
&= 2 \vec{B} \cdot \vec{B} \\
\rightarrow \frac{1}{4} F_{ij} F^{ij} &= \frac{1}{2} \vec{B} \cdot \vec{B}
\end{aligned}
$$

thus the hamiltonian can be written in a more conventional way as advertised earlier

$$
\mathcal{H}_{\mathrm{ph}} = \frac{1}{2} (\vec{E}^2 + \vec{B}^2) - A_0 (\nabla \cdot \vec{E}) + J^\mu A_\mu
$$

Let's see what we can get out of this hamiltonian, first

$$
\frac{\delta \mathcal{H}_{\mathrm{ph}}}{\delta A_\rho} = -\partial_0 p_A^\rho
$$

12

Note that $p_A^0 = F^{00} = 0$ means that $\delta\mathcal{H}_{\mathrm{ph}}/\delta A_0 = -\partial_0(0) = 0$

$$\frac{\delta\mathcal{H}_{\mathrm{ph}}}{\delta A_0} = -\partial_i E^i + J^0$$

$$0 = -\partial_i E^i + J^0$$

$$\partial_i E^i = J^0$$

$$\to \nabla \cdot \vec{E} = \rho$$

Where $J^\mu = e\overline{\psi}\gamma^\mu\psi = (\rho, \vec{J})$, so we obtain the first of the Maxwell's equations, *i.e.* gauss's law. Next is $\delta\mathcal{H}_{\mathrm{ph}}/\delta A_k$. First, the terms we need are $\frac{1}{4}F_{ij}F^{ij} + J^i A_i$.

$$F_{ij}F^{ij} = (\partial_i A_j - \partial_j A_i)F^{ij}$$

$$F_{ij}F^{ij} = -A_j\partial_i F^{ij} + A_i\partial_j F^{ij}$$

$$F_{ij}F^{ij} = F_{ij}(\partial^i A^j - \partial^j A^i)$$

$$F_{ij}F^{ij} = -\partial^i F_{ij}A^j + \partial^j F_{ij}A^i$$

And we have done plenty of integration by parts (for spatial derivatives only), thus

$$\frac{\delta\left(\frac{1}{4}F_{ij}F^{ij}\right)}{\delta A_k} = \frac{1}{4}\left(-\partial_i F^{ik} + \partial_j F^{kj} - \partial^i F_i{}^k + \partial^j F^k{}_j\right)$$

$$= \frac{1}{4}\left(-\partial_i F^{ik} + \partial_j F^{kj} - \partial_i F^{ik} + \partial_j F^{kj}\right)$$

$$= \frac{1}{4}\left(-2\partial_i F^{ik} + 2\partial_j F^{kj}\right) = -\partial_i F^{ik}$$

The e.o.m is then

$$\frac{\delta\mathcal{H}_{\mathrm{ph}}}{\delta A_k} = -\partial_i F^{ik} + J^k$$

$$-\partial_0 p_A^k = -\partial_i F^{ik} + J^k$$

We can write it in a more familiar form by remembering that $p_A^k = E^k$ and

$$-\partial_i F^{ik} = -\partial_i(-\varepsilon^{ikm}B_m) = \varepsilon^{ikm}\partial_i B_m$$

$$= -\varepsilon^{kim}\partial_i B_m = -(\nabla \times \vec{B})^k$$

the e.o.m can then be written as

$$-\partial_0 p_A^k = -\partial_0 E^k = -(\nabla \times \vec{B})^k + J^k$$

$$\to (\nabla \times \vec{B}) = \partial_0\vec{E} + \vec{J}$$

13

which is ampere's law. At this point we can stop since we have recovered all of Maxwell's equations but let's see what the other hamiltonian equations bring us

$$\frac{\delta \mathcal{H}_{\text{ph}}}{\delta p_A^i} = \partial_0 A_i$$

$$\frac{\delta \mathcal{H}_{\text{ph}}}{\delta p_A^i} = \frac{\delta \mathcal{H}_{\text{ph}}}{\delta E^i} = \frac{1}{2} E^i + \frac{1}{2} E^i + \partial_i A_0$$

$$= F_{0i} + \partial_i A_0 = \partial_0 A_i - \partial_i A_0 + \partial_i A_0$$

$$\partial_0 A_i = \partial_0 A_i$$

which is just a tautology. Let's now tackle the fermionic hamiltonian. In this case, there is an equivocation. Since the conjugate momentum $p_\psi = i\gamma^0 \overline{\psi}$ contains no time derivative, we don't know what to substitute for p_ψ in the lagrangian. The fastest way to get the e.o.m's is to actually substitute all occurances $\overline{\psi}$ for p_ψ.

$$\mathcal{H}_{\text{fm}} = -p_\psi^0 \gamma^0 \gamma^i \partial_i \psi - iep_\psi^0 A_0 \psi - eip_\psi^0 \gamma^0 \gamma^i A_i \psi - imp_\psi^0 \gamma^0 \psi$$

$$= p_\psi^0 \gamma^i \gamma^0 \partial_i \psi - iep_\psi^0 A_0 \psi + eip_\psi^0 \gamma^i \gamma^0 A_i \psi - imp_\psi^0 \gamma^0 \psi$$

where $\gamma^0 \gamma^i = -\gamma^i \gamma^0$ from the anticommutation relation of the gamma matrices. Moving on with this hamiltonian, we can recover all of Dirac's e.o.m's

$$\frac{\delta \mathcal{H}_{\text{fm}}}{\delta p_\psi^0} = \partial_0 \psi$$

$$\frac{\delta \mathcal{H}_{\text{fm}}}{\delta p_\psi^0} = -\gamma^0 \gamma^i \partial_i \psi - ie A_0 \psi - ei\gamma^0 \gamma^i A_i \psi - im\gamma^0 \psi$$

$$\partial_0 \psi = -\gamma^0 \gamma^i \partial_i \psi - ie A_0 \psi - ei\gamma^0 \gamma^i A_i \psi - im\gamma^0 \psi$$

$$i\gamma^0 \partial_0 \psi = -i\gamma^i \partial_i \psi + e\gamma^0 A_0 \psi + e\gamma^i A_i \psi + m\psi$$

$$i\gamma^\mu \partial_\mu \psi = e\gamma^\mu A_\mu \psi + m\psi$$

The next e.o.m is easier to derive if we use the second line (with γ^i on the left of γ^0) of

14

the hamiltonian, $\mathcal{H}_{\text{fm}}$, above because we want to multiply by γ^0 from the right.

$$\frac{\delta \mathcal{H}_{\text{fm}}}{\delta \psi} = -\partial_0 p_\psi^0$$

$$\frac{\delta \mathcal{H}_{\text{fm}}}{\delta \psi} = -\partial_i p_\psi^0 \gamma^i \gamma^0 - i e p_\psi^0 A_0 + e i p_\psi^0 \gamma^i \gamma^0 A_i - i m p_\psi^0 \gamma^0$$

$$-\partial_0 p_\psi^0 \gamma^0 = -\partial_i p_\psi^0 \gamma^i - i e p_\psi^0 \gamma^0 A_0 + e i p_\psi^0 \gamma^i A_i - i m p_\psi^0$$

$$-i \partial_0 \overline{\psi} \gamma^0 \gamma^0 = -i \partial_i \overline{\psi} \gamma^0 \gamma^i + e \overline{\psi} \gamma^0 \gamma^0 A_0 - e \overline{\psi} \gamma^0 \gamma^i A_i + m \overline{\psi} \gamma^0$$

$$-i \partial_0 \overline{\psi} \gamma^0 \gamma^0 = i \partial_i \overline{\psi} \gamma^i \gamma^0 + e \overline{\psi} \gamma^0 \gamma^0 A_0 + e \overline{\psi} \gamma^i \gamma^0 A_i + m \overline{\psi} \gamma^0$$

$$-i \partial_0 \overline{\psi} \gamma^0 = i \partial_i \overline{\psi} \gamma^i + e \overline{\psi} \gamma^0 A_0 + e \overline{\psi} \gamma^i A_i + m \overline{\psi}$$

$$-i \partial_\mu \overline{\psi} \gamma^\mu = e \overline{\psi} \gamma^\mu A_\mu + m \overline{\psi}$$

from third to fourth line we have substituted p_ψ^0 for $i \overline{\psi} \gamma^0$ and from fourth to fifth line we have swapped the order of $\gamma^0 \gamma^i \to \gamma^i \gamma^0$ incurring a minus sign. Again, although the above derivation yields the correct e.o.m's they are fundamentally wrong! We need to employ Dirac-Bergmann algorithm to do it correctly.

Let's take a pause at this point and take stock. As the title of the subsection suggests, there seems to be no problem at all, we recovered all of Maxwell's and Dirac's equations. The problem is more pronounced in the fermonic case, it is

$$\frac{\delta \mathcal{H}_{\text{fm}}}{\delta p_{\overline{\psi}}} = \partial_0 \overline{\psi}$$

$$\to 0 = \partial_0 \overline{\psi}$$

which is the wrong e.o.m for $\overline{\psi}$. The same can *not* be said about photons

$$\frac{\delta \mathcal{H}_{\text{ph}}}{\delta p_A^0} = \partial_0 A_0$$

$$\to 0 = \partial_0 A_0$$

since A_0 can always be transformed away using gauge transformations. But the question remains, how can we recover all of Maxwell's equations with only half of the hamiltonian equations, $\dot{p} = -\partial H / \partial q$? The partial answer is that Maxwell's equation is quite different from the usual e.o.m's we are accustomed to.

Let's go back to a free particle in classical mechanics, $H = p^2/2m$, (or analogously the free scalar field). Newton told us that we need a second derivative and thus we need two

hamiltonian equations $\dot{q} = \partial H/\partial p$ and $\dot{p} = -\partial H/\partial q$. Since $p = m\dot{q}$ (which is obtained from $\dot{q} = \partial H/\partial p$), $\dot{p} = m\ddot{q} = -\partial H/\partial q$, we recover Newton's law. The two first order equations are needed to construct one second order equation.

Now this is where Maxwell's (and Dirac's) equations differ. Maxwell requires only derivatives of the momenta, $p_A^i = E^i$, we do *not* need derivatives of the "position" variables themselves which in this case are the vector potential A. This is because the actual observables in electrodynamics is the momenta E^i (and B^i) while the observables in our free particle example above (and classical mechanics in general) is the position of the particle, q. So we can safely discard the equations of motion giving first derivatives of A (which are just the electric and magnetic fields).

I guess this is the price we pay for wanting our lagrangians to be gauge invariant, *i.e.* by formulating our lagrangian in terms of A instead of E, B. Of course in quantum mechanics we cannot get the correct result without the vector potential A, *e.g.* the Aharonov-Bohm effect. Dirac's are more weird, which we will discuss next.

B. Dealing with Dirac's quirks

The Dirac lagrangian is one weird lagrangian. It's weird because Dirac set out to create a Schrodinger equation with first derivatives for both time and space and so he created a hamiltonian which is the "square root" of Klein-Gordon's. Due to this square root or only first derivative appearing in the hamiltonian things look weird, to see how weird it is let's start with the Dirac lagrangian

$$S = \int dt \, \mathcal{L} = \int d^4x \, i\overline{\psi}\gamma^\mu \partial_\mu \psi - m\overline{\psi}\psi$$

$$\delta S = \int d^4x \, i(\delta\overline{\psi})\gamma^\mu \partial_\mu \psi - m(\delta\overline{\psi})\psi + i\overline{\psi}\gamma^\mu(\delta\partial_\mu\psi) - m\overline{\psi}(\delta\psi)$$

$$= \int d^4x \, \delta\overline{\psi}\left(i\gamma^\mu\partial_\mu\psi - m\psi\right) + \left(-i\partial_\mu\overline{\psi}\gamma^\mu - m\overline{\psi}\right)\delta\psi$$

where we have done integration by parts on $(\delta\partial_\mu\psi)$ going from line 2 to line 3. Now you see the weirdness, variation with respect to $\overline{\psi}$ gives you the e.o.m of ψ and the variation w.r.t ψ gives you the e.o.m for $\overline{\psi}$. Hence, without $\overline{\psi}$ we cannot get the e.o.m of ψ!

16

In classical mechanics, we all know that a lagrangian of the form $L = \dot{q} + q$ is not allowed since

$$\partial_t \left(\frac{\partial L}{\partial \dot{q}} \right) = \frac{\partial L}{\partial q}$$
$$0 = 1$$

but that is exactly what the Dirac lagrangian looks like. The only way to salvage this is to add another degree of freedom, for simplicity let's call that d.o.f a such that

$$L = a\dot{b} - mab$$

however we still have a problem

$$p_b = a \qquad p_a = 0$$

there's no conjugate momentum for a, $p_a = 0$! You might not be too worried since we still can get the e.o.m's for a and b from L

$$\partial_t \left(\frac{\partial L}{\partial \dot{b}} \right) - \frac{\partial L}{\partial b} = \dot{a} + ma = 0$$
$$\partial_t \left(\frac{\partial L}{\partial \dot{a}} \right) - \frac{\partial L}{\partial a} = -\dot{b} + mb = 0$$

However, the hamiltonian gives the wrong result since

$$\dot{a} = \frac{\partial H}{\partial p_a} = \frac{\partial (p_b \dot{b} - a\dot{b} + mab)}{\partial p_a}$$
$$= 0$$
$$\dot{a} = \{a, H\} = \{a, p_b \dot{b} - a\dot{b} + mab\}$$
$$= 0$$

where we have calculated $\dot{a}$ in two ways and they both give the wrong result. At this point let's switch to another topic, something that we do not usually care about when doing legendre transformations going from the lagrangian to the hamiltonian.

As we have seen above, our conjugate momentum for a is zero. We can salvage this if we do an integration by parts to move the time derivative from $a\dot{b} \to -\dot{a}b$. At first sight

this is ok, after all variations at end points vanish by construction, *i.e.*

$$\int_{t_0}^{t_1} dt \ A(t)(\partial_t \delta q) = \int_{t_0}^{t_1} dt \ \partial_t(A(t)\delta q) - (\partial_t A(t))\delta q$$

$$= (A(t_1)\delta q(t_1)) - (A(t_0)\delta q(t_0)) - \int_{t_0}^{t_1} dt \ (\partial_t A(t))\delta q$$

$$= - \int_{t_0}^{t_1} dt \ (\partial_t A(t))\delta q$$

as $\delta q(t_0) = \delta q(t_1) = 0$ which is the rule of variations that we use to derive e.o.m's from the action S. The above usually applies to the lagrangian since $S = \int dt \ L$. However from the definition of the legrendre transformation we get

$$S = \int dt \ L$$

$$= \int dt \ (p\dot{q} - H(p,q))$$

so it seems that we can do integration by parts on H as well?

$$S \overset{?}{=} \int dt \ (-\dot{p}q - H(p,q))$$

well, we have to be careful here, for example

$$H(p,q) = p\dot{q} - L(q,\dot{q})$$

$$= -\dot{p}q - L(q,\dot{q})$$

$$\rightarrow \frac{\partial H(p,q)}{\partial \dot{q}} = 0 - \frac{\partial L}{\partial \dot{q}}$$

$$0 = -\frac{\partial L}{\partial \dot{q}}$$

which is wrong unless L doesn't depend on $\dot{q}$ at all, in other words you have to integrate by parts all $\dot{q}$'s such that there is none left in L for the above to hold. One thing you must *not* do is to integrate by parts on certain terms and not integrate by parts on others. Thus we have to always be consistent.

Let's see how all this works out by going back to our example of $L = a\dot{b} - mab$. What hamiltonian should we assign to this lagrangian? we have a few choices since we can do

integration by parts, they are

$$H_1 = p_b\dot{b} - a\dot{b} + mab, \qquad\qquad\qquad p_a = 0, \ p_b = a$$

$$H_2 = p_a\dot{a} + \dot{a}b + mab, \qquad\qquad\qquad p_a = -b, \ p_b = 0$$

$$H_3 = p_a\dot{a} + p_b\dot{b} - K_1 a\dot{b} + K_2\dot{a}b + mab, \quad p_a = -K_2 b, \ p_b = K_1 a, \ K_1 + K_2 = 1$$

$$H_4 = p_a\dot{a} + p_b\dot{b} - a\dot{b} + mab, \qquad\qquad p_a = -b, \ p_b = a$$

where we have shown the corresponding conjugate momenta for each construction of the hamiltonians. H_3 is where we choose to distribute the time derivative among a and b. H_4 is interesting because we do the integration by parts as and when we need it, *i.e.* we move the time derivative to a when we calculate p_a and let the time derivative stay on b when we calculate p_b (as well as when we do legendre transform to get H_4 from L).

Integration by parts

Before we proceed let's see a pitfall of doing integration by parts w.r.t time. Say we want to calculate $\dot{p}_b = -\frac{\partial H_1}{\partial b}$ and we do an integration by parts on $a\dot{b} \to -\dot{a}b$

$$\dot{p}_b = -\frac{\partial H_1}{\partial b}$$

$$\overset{?}{=} -p_b\frac{\partial\dot{b}}{\partial b} - \dot{a}\frac{\partial b}{\partial b} - ma\frac{\partial b}{\partial b}$$

$$\dot{p}_b \overset{?}{=} 0 - \dot{a} - ma, \quad p_b = a$$

$$\dot{a} \overset{?}{=} -\dot{a} - ma$$

$$0 \overset{?}{=} 2\dot{a} + ma$$

which is the wrong e.o.m for a. The inconsistency was because we did integration by parts on $a\dot{b}$ but did not do it for $p_b\dot{b}$, to be consistent we have to

$$\dot{p}_b = -\frac{\partial H_1}{\partial b}$$

$$\overset{?}{=} \dot{p}_b\frac{\partial b}{\partial b} - \dot{a}\frac{\partial b}{\partial b} - ma\frac{\partial b}{\partial b}$$

$$\dot{p}_b \overset{?}{=} \dot{p}_b - \dot{a} - ma, \quad p_b = a$$

$$\dot{a} \overset{?}{=} \dot{a} - \dot{a} - ma$$

$$0 \overset{?}{=} \dot{a} + ma$$

which is the correct e.o.m for a, so as long as we are consistent we "should" be ok. The reason for this discussion about integration by parts is because we need to do it to put the time derivative on both a and b and if we are not careful we will get all sorts of wrong results.

Dealing with constraints

Going back to our hamiltonians, we see that H_1 and H_2 don't work because

$$\dot{a} = \frac{\partial H_1}{\partial p_a} = 0$$
$$\dot{b} = \frac{\partial H_2}{\partial p_b} = 0$$

which is inconsistent with the e.o.m for a and b obtained from the lagrangian, $i.e.$ $\dot{a} + ma = 0$, $-\dot{b} + mb = 0$. The rest of the e.o.m's are okay and can be obtained by the usual tricks $\dot{q} = \partial H/\partial p$, $\dot{p} = -\partial H/\partial q$ although there are complications because we have constraints from the fact that we cannot invert p to get $\dot{q}$ in terms of p.

H_4 is a difficult case and should be avoided because it's inconsistent in terms of doing integration by parts. Say we want to calculate $\dot{p}_b = -\frac{\partial H_4}{\partial b}$ and we do integration by parts $-a\dot{b} \to \dot{a}b$ in H_4 and so we get

$$\dot{p}_b = -\frac{\partial H_4}{\partial b}$$
$$= -\dot{a} - ma, \quad p_b = a$$
$$\to 0 = 2\dot{a} + ma$$

where we have substituted the constraint $p_b = a$. We see that we get the wrong answer, to get the correct answer we need to leave $-a\dot{b}$ as they are

$$\dot{p}_b = -\frac{\partial H_4}{\partial b}$$
$$= -ma, \quad p_b = a$$
$$\to 0 = \dot{a} + ma$$

and we get the correct answer, so you might ask does that mean that we should not do

20

any integration by parts at all? Not really because

$$\dot{p}_a = -\frac{\partial H_4}{\partial a}, \quad p_a = -b$$
$$-\dot{b} = \dot{b} - mb$$
$$\to 0 = 2\dot{b} - mb$$

again we get the wrong e.o.m for b but this time we didn't do any integration by parts, to get the right result we should do integration by parts $-a\dot{b} \to \dot{a}b$

$$\dot{p}_a = -\frac{\partial H_4}{\partial a}, \quad p_a = -b$$
$$-\dot{b} = -mb$$
$$\to 0 = \dot{b} - mb$$

Why is this? This is because when we calculated p_a we did an integration by parts $a\dot{b} \to -\dot{a}b$ in the lagrangian and when we calculated p_b we didn't do integration by parts thus we have to be consistent when we calculate the hamiltonian e.o.m's that utilize them. Somehow the hamiltonian keeps track "what you did last summer".

Going back to constraints, in the above calculations involving H_4 we utilized the constraints to get the correct e.o.m's (this is also true for $H_{1,2,3}$), we should not do that.

Ambiguities, do they matter?

We need to apply Dirac-Bergmann algorithm to deal with constraints. However, there is an indefiniteness in applying this algorithm, for example take H_1, since p_a is zero should we include them in the hamiltonian?

$$H'_{11} = p_b\dot{b} - a\dot{b} + mab + \lambda_1 C_1 + \lambda_2 C_2 \quad \text{OR} \quad H'_{12} = p_b\dot{b} + p_a\dot{a} - a\dot{b} + mab + \lambda_1 C_1 + \lambda_2 C_2$$

where the C_m's are the constraints given by $C_1 = p_b - a$, $C_2 = p_a$. You can argue that by not including p_a in H_1 we have used the constraint of $p_a = 0$. Using the notation in Dirac's book we denote the above hamiltonians (with the unsolved lagrange multipliers (λ_m)) H'_{11} and H'_{12}.

We'll begin by seeing what happens if we include p_a, i.e. H'_{12}. First we take the various

Poisson brackets, the Poisson brackets between the constraints $C_1 = p_b - a$, $C_2 = p_a$ is

$$\{C_1, C_2\} = \frac{\partial C_1}{\partial b}\frac{\partial C_2}{\partial p_b} - \frac{\partial C_1}{\partial p_b}\frac{\partial C_2}{\partial b} + \frac{\partial C_1}{\partial a}\frac{\partial C_2}{\partial p_a} - \frac{\partial C_1}{\partial p_a}\frac{\partial C_2}{\partial a}$$

$$= 0 - 0 - 1 - 0$$

$$= -1$$

This Poisson bracket relation is true for both H'_{11} and H'_{12}. This step is required in Dirac-Bergmann algorithm to see if the constraints are first or second class.

Next we do the consistency checks by taking the Poisson bracket between the constraints and the hamiltonian, $i.e.$ we want to make sure that the constraint is constant w.r.t time and for convenience I reproduce the hamiltonian below

$$H'_{12} = p_b\dot{b} + p_a\dot{a} - a\dot{b} + mab + \lambda_1 C_1 + \lambda_2 C_2, \quad C_1 = p_b - a, \ C_2 = p_a$$

$$\{C_1, H'_{12}\} = \frac{\partial C_1}{\partial b}\frac{\partial H'_{12}}{\partial p_b} - \frac{\partial C_1}{\partial p_b}\frac{\partial H'_{12}}{\partial b} + \frac{\partial C_1}{\partial a}\frac{\partial H'_{12}}{\partial p_a} - \frac{\partial C_1}{\partial p_a}\frac{\partial H'_{12}}{\partial a}$$

$$= 0 - (ma) + (-1)(\dot{a} + \lambda_2) - 0$$

$$= -ma - \dot{a} - \lambda_2$$

$$\lambda_2 = -\dot{a} - ma$$

$$\{C_2, H'_{12}\} = \frac{\partial C_2}{\partial b}\frac{\partial H'_{12}}{\partial p_b} - \frac{\partial C_2}{\partial p_b}\frac{\partial H'_{12}}{\partial b} + \frac{\partial C_2}{\partial a}\frac{\partial H'_{12}}{\partial p_a} - \frac{\partial C_2}{\partial p_a}\frac{\partial H'_{12}}{\partial a}$$

$$= 0 - 0 + 0 - (-\dot{b} + mb - \lambda_1)$$

$$\lambda_1 = -\dot{b} + mb$$

Let's see what happens for H'_{11}

$$H'_{11} = p_b\dot{b} - a\dot{b} + mab + \lambda_1(p_b - a) + \lambda_2 p_a$$

$$\{C_1, H'_{11}\} = \frac{\partial C_1}{\partial b}\frac{\partial H'_{11}}{\partial p_b} - \frac{\partial C_1}{\partial p_b}\frac{\partial H'_{11}}{\partial b} + \frac{\partial C_1}{\partial a}\frac{\partial H'_{11}}{\partial p_a} - \frac{\partial C_1}{\partial p_a}\frac{\partial H'_{11}}{\partial a}$$

$$= 0 - (ma) - (\lambda_2) - 0$$

$$= -ma - \lambda_2$$

$$\lambda_2 = -ma$$

$$\{C_2, H'_{11}\} = \frac{\partial C_2}{\partial b}\frac{\partial H'_{11}}{\partial p_b} - \frac{\partial C_2}{\partial p_b}\frac{\partial H'_{11}}{\partial b} + \frac{\partial C_2}{\partial a}\frac{\partial H'_{11}}{\partial p_a} - \frac{\partial C_2}{\partial p_a}\frac{\partial H'_{11}}{\partial a}$$

$$= 0 - 0 + 0 - (-\dot{b} + mb - \lambda_1)$$

$$\lambda_1 = -\dot{b} + mb$$

First class contributions

So we see that we get different solutions for λ_2. Before moving on, we see that the consistency checks did not produce more constraints. But do we get first class contributions to λ's, the ones called $V_{am}\{C_n, C_m\} = 0$ by Dirac?

In other words we just solved $\lambda_m\{C_n, C_m\} = -\{C_n, H\}$ since

$$\{C_n, H'\} = \{C_n, H\} + \lambda_m\{C_n, C_m\}, \quad H' = H + \lambda_m C_m$$

$$0 \approx \{C_n, H\} + \lambda_m\{C_n, C_m\}$$

$$-\{C_n, H\} \approx \lambda_m\{C_n, C_m\}$$

which is equivalent to solving a set of linear equations

$$\lambda_1 a_{11} + \lambda_2 a_{12} = b_1$$

$$\lambda_1 a_{21} + \lambda_2 a_{22} = b_2$$

where the matrix $a_{nm} = \{C_n, C_m\}$ and the vector $b_n = -\{C_n, H'\}$. What we want to know is if we can add the homogeneous solution V_{am} to λ_m

$$V_{a1} a_{11} + V_{a2} a_{12} = 0$$

$$V_{a1} a_{21} + V_{a2} a_{22} = 0$$

the only way for V_{am} to be *non* zero is for the matrix a_{mn} to be non invertible, $\det(a) = 0$ (the index $_a$ in V_{am} is to indicate that there might be more than one *set* of homogenous solution) but we see above that $a_{nm} = \{C_n, C_m\}$ is antisymmetric with nonzero components $a_{12} = -a_{21} = -1$. Hence a is invertible and $V_{am} = 0$, we don't have any first class contributions to λ_m! Note that we can always take a linear combination of homogeneous solutions $v_a V_{am}$ to get another one.

Equations of motion

We have seen above that we only have second class constraints $\{C_1, C_2\} \neq 0$ and there's no first class contributions to λ_m. We now want to check if the choices of the two hamiltonians H'_{11} and H'_{12} give different e.o.m's. A word about notation, in Dirac's book once we solve for the λ's we denote the hamiltonian with superscript ^{T}, *i.e.* $H^T_{11,12}$.

Before we dwelve into the e.o.m's there's another inconclusiveness we need to tackle. The e.o.m's can be calculated three ways, the most traditional way is by varying the action

$$L = p\dot{q} - H_T = p\dot{q} - H - \lambda_m C_m$$

$$\delta S \to \delta L = \delta p \dot{q} - \dot{p} \delta q - \frac{\partial H}{\partial q} \delta q - \frac{\partial H}{\partial p} \delta p - \frac{\partial(\lambda_m C_m)}{\partial q} \delta q - \lambda_m \frac{\partial(\lambda_m C_m)}{\partial p} \delta p$$

$$= \delta p \left(\dot{q} - \frac{\partial H}{\partial p} - \frac{\partial(\lambda_m C_m)}{\partial p} \right) + \delta q \left(-\dot{p} - \frac{\partial H}{\partial q} - \frac{\partial(\lambda_m C_m)}{\partial q} \right)$$

secondly, by taking the Poisson brackets which generate contact transformations

$$\dot{p} = \{p, H\} + \{p, \lambda_m C_m\}$$

$$\dot{q} = \{q, H\} + \{q, \lambda_m C_m\}$$

$$C_m = 0$$

The puzzle is in whether we vary λ_m w.r.t p and q, *i.e.* in the first and second methods

$$\frac{\partial(\lambda_m C_m)}{\partial q} \stackrel{?}{=} \lambda_m \frac{\partial C_m}{\partial q} \qquad\qquad \{q, \lambda_m C_m\} \stackrel{?}{=} \lambda_m \{q, C_m\}$$

The thing is that we solved λ_m in terms of p and q through the consistency checks $\{C_n, H'\} \approx 0$ and thus $\lambda_m = f_m(p, q)$. To eliminate this incertitude let's go back to the beginning, in the beginning ... there were lagrange multipliers

$$L(q, \dot{q}) = p\dot{q} - H(p, q) + \lambda_m(t) C_m$$

$$\delta L(q, \dot{q}) = \delta(p\dot{q} - H(p, q)) + (\delta \lambda_m(t)) C_m + \lambda_m(t)(\delta C_m)$$

and it is the term $(\delta \lambda_m) C_m = 0$ that forces $C_m = 0$. Our initial assumption is that λ is just an independent parameter that we can vary together with L and they are solved only after we obtain the e.o.m's which is also the standard procedure in lagrangian mechanics.

The same can be said with λ_m's because we only solve them only after we calculate all Poisson brackets during our consistency checks, thus

$$\frac{\partial(\lambda_m C_m)}{\partial q} = \lambda_m \frac{\partial C_m}{\partial q} \qquad\qquad \{q, \lambda_m C_m\} = \lambda_m \{q, C_m\}$$

and the e.o.m's are given through the following three equivalent ways

$$0 = \delta p \left(\dot{q} - \frac{\partial H}{\partial p} - \lambda_m \frac{\partial C_m}{\partial p} \right) \qquad\qquad 0 = \delta q \left(-\dot{p} - \frac{\partial H}{\partial q} - \lambda_m \frac{\partial C_m}{\partial q} \right)$$

$$\dot{p} = \{p, H\} + \lambda_m\{p, C_m\}$$

$$\dot{q} = \{q, H\} + \lambda_m\{q, C_m\}$$

$$C_m = 0$$

$$\dot{p} = \{p, H_T\}^* = \{p, H_T\} - \{p, C_m\}\{C_m, C_n\}^{-1}\{C_n, H_T\}$$

$$\dot{q} = \{q, H_T\}^* = \{q, H_T\} - \{q, C_m\}\{C_m, C_n\}^{-1}\{C_n, H_T\}$$

$$\{C_m, C_n\}^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Now, let's see if both H'_{11} and H'_{12} give us the same e.o.m's through all three formulations. Let's just compute $\dot{a}$ for now since it is the prioblematic one due to the absence of p_a. First, from $H_{12}^T = p_b\dot{b} + p_a\dot{a} - a\dot{b} + mab + \lambda_1(p_b - a) + \lambda_2(p_a)$ we get

$$\dot{a} = \{a, H_{12}^T\} = \{a, H_{12}\} + \lambda_m\{a, C_m\}$$

$$= (\dot{a}) + (\lambda_2)$$

$$0 = -\dot{a} - ma$$

$$\rightarrow 0 = \dot{a} + ma$$

and from $H_{11}^T = p_b\dot{b} - a\dot{b} + mab + \lambda_1(p_b - a) + \lambda_2(p_a)$ we get

$$\dot{a} = \{a, H_{11}^T\} = \{a, H_{11}\} + \lambda_m\{a, C_m\}$$

$$\dot{a} = 0 + (\lambda_2)$$

$$\rightarrow 0 = \dot{a} + ma$$

So we see that they both give the same e.o.m for a, let's see if the traditional way gives the correct result, first up H_{12}^T

$$\dot{a} = \frac{\partial H_{12}^T}{\partial p_a} = \frac{\partial H_{12}}{\partial p_a} + \lambda_m\frac{\partial C_m}{\partial p_a}$$

$$\dot{a} = (\dot{a}) + (\lambda_2)$$

$$\rightarrow 0 = \dot{a} + ma$$

Now H_{11}^T's turn

$$\dot{a} = \frac{\partial H_{11}^T}{\partial p_a} = \frac{\partial H_{11}}{\partial p_a} + \lambda_m\frac{\partial C_m}{\partial p_a}$$

$$\dot{a} = (0) + (\lambda_2)$$

$$\rightarrow 0 = \dot{a} + ma$$

So the two methods give consistent results. Let's see if Dirac brackets also gives the right results. According to Dirac's recipe we should set second class constraints to zero, but of course we only set them to zero in the hamiltonian *not* in the Dirac brackets themselves, otherwise they'll just be ordinary Poisson brackets. As usual first up H_{12}^T

$$
\begin{aligned}
\dot{a} = \{a, H_{12}^T\}^* &= \{a, H_{12}^T\} - \{a, C_m\}\{C_m, C_n\}^{-1}\{C_n, H_{12}^T\} \\
&= (\dot{a} + \lambda_2) - \{a, C_1\}\{C_1, C_2\}^{-1}\{C_2, H_{12}^T\} - \{a, C_2\}\{C_2, C_1\}^{-1}\{C_1, H_{12}^T\} \\
&= (\dot{a} + \lambda_2) - 0 - (1)(-1)(\{C_1, H_{12}\} + \{C_1, \lambda_1 C_1\} + \{C_1, \lambda_2 C_2\}) \\
&= (\dot{a} + \lambda_2) + (-\dot{a} - ma + \{C_1, \lambda_1 C_1\} + \{C_1, \lambda_2 C_2\})
\end{aligned}
$$

We need to take a pause here, H_T is defined as H' with all λ's solved, thus the Poisson bracket $\{C_m, \lambda_n\} \neq 0$

$$
\begin{aligned}
\{C_1, \lambda_1 C_1\} &= \lambda_1 \{C_1, C_1\} + \{C_1, \lambda_1\} C_1 \\
&= 0 + \{C_1, -\dot{b} + mb\} C_1 \\
&= \{p_b - a, -\dot{b} + mb\}(p_b - a) \\
&= -m(p_b - a)
\end{aligned}
$$

while for the last term we get

$$
\begin{aligned}
\{C_1, \lambda_2 C_2\} &= \lambda_2 \{C_1, C_2\} + \{C_1, \lambda_2\} C_2 \\
&= \lambda_2 \{p_b - a, p_a\} + \{p_b - a, -\dot{a} - ma\}(p_b - a) \\
&= -\lambda_2 \\
&= \dot{a} + ma
\end{aligned}
$$

collecting everything gives us

$$
\begin{aligned}
\dot{a} &= (\dot{a} + \lambda_2) + (-\dot{a} - ma + \{C_1, \lambda_1 C_1\} + \{C_1, \lambda_2 C_2\}) \\
0 &= \lambda_2 + (-\dot{a} - ma - m(p_b - a) + \dot{a} + ma) \\
0 &= \dot{a} + ma + m(p_b - a)
\end{aligned}
$$

While the e.o.m of H_{11}^T is

$$\dot{a} = \{a, H_{11}^T\}^* = \{a, H_{11}^T\} - \{a, C_m\}\{C_m, C_n\}^{-1}\{C_n, H_{11}^T\}$$

$$= (\lambda_2) - \{a, C_1\}\{C_1, C_2\}^{-1}\{C_2, H_{11}^T\} - \{a, C_2\}\{C_2, C_1\}^{-1}\{C_1, H_{11}^T\}$$

$$= (\lambda_2) - 0 - (1)(-1)(\{C_1, H_{11}\} + \{C_1, \lambda_1 C_1\} + \{C_1, \lambda_2 C_2\})$$

$$= (\lambda_2) + (-ma + \{C_1, \lambda_1 C_1\} + \{C_1, \lambda_2 C_2\})$$

Again, the second to last term gives us

$$\{C_1, \lambda_1 C_1\} = \lambda_1\{C_1, C_1\} + \{C_1, \lambda_1\}C_1$$

$$= 0 + \{C_1, -\dot{b} + mb\}C_1$$

$$= \{p_b - a, -\dot{b} + mb\}(p_b - a)$$

$$= -m(p_b - a)$$

while the last term is

$$\{C_1, \lambda_2 C_2\} = \lambda_2\{C_1, C_2\} + \{C_1, \lambda_2\}C_2$$

$$= \lambda_2\{p_b - a, p_a\} + \{p_b - a, -ma\}(p_b - a)$$

$$= -\lambda_2$$

$$= ma$$

substituting all of them back into the Dirac bracket gives

$$\dot{a} = (\lambda_2) + (-ma + \{C_1, \lambda_1 C_1\} + \{C_1, \lambda_2 C_2\})$$

$$\dot{a} = \lambda_2 + (-ma - m(p_b - a) + ma)$$

$$0 = \dot{a} + ma - m(p_b - a)$$

So we get similar results from H_{11}^T and H_{12}^T. We see that the extra term $m(p_b - a)$ are just the second class constraint and should be set to zero because the Dirac bracket of any function of q and p with any of the second class constraint is zero

$$\{f(q, p), C_m\}^* = \{f(q, p), C_m\} - \{f(q, p), C_n\}\{C_n, C_k\}^{-1}\{C_k, C_m\}$$

$$= \{f(q, p), C_m\} - \{f(q, p), C_n\}\delta_{nm}$$

$$= \{f(q, p), C_m\} - \{f(q, p), C_m\}$$

$$= 0$$

27

which means that second *class* constraints do not generate any transformation.

Let's see what we'll get if we set all second class constraints to zero in the hamiltonian before taking the Dirac brackets, *i.e.* $H = H_T$, first up H_{12}

$$\dot{a} = \{a, H_{12}^T\}^* = \{a, H_{12}\} - \{a, C_m\}\{C_m, C_n\}^{-1}\{C_n, H_{12}\}$$
$$= (\dot{a}) - \{a, C_1\}\{C_1, C_2\}^{-1}\{C_2, H_{12}\} - \{a, C_2\}\{C_2, C_1\}^{-1}\{C_1, H_{12}\}$$
$$= (\dot{a}) - 0 - (1)(-1)(\{C_1, H_{12}\})$$
$$\dot{a} = (\dot{a}) + (-\dot{a} - ma)$$
$$0 = -\dot{a} - ma$$

and next H_{11}

$$\dot{a} = \{a, H_{11}^T\}^* = \{a, H_{11}\} - \{a, C_m\}\{C_m, C_n\}^{-1}\{C_n, H_{11}\}$$
$$= -\{a, C_1\}\{C_1, C_2\}^{-1}\{C_2, H_{11}\} - \{a, C_2\}\{C_2, C_1\}^{-1}\{C_1, H_{11}\}$$
$$= -0 - (1)(-1)(\{C_1, H_{11}\})$$
$$\dot{a} = -ma$$
$$0 = \dot{a} + ma$$

In this case we immediately get the e.o.m for a without the extra term (C_1) since we set it to zero in the hamiltonian.

A note of caution, in deriving e.o.m's from H_{11}' we have used the constraint implicitly, *i.e.* we exclude $p_a\dot{a}$ from $\sum p_k\dot{q}_k - L$ during the legendre transformation and by doing this we have inadvertently enforced $p_a = 0$, which is the constraint. There are two things to note here, first, we only apply the constraint of $p_a = 0$ but not $p_b = a$ when doing the legendre transformation, this is random. Next, even though we got the same e.o.m for a from H_{11}' and H_{12}' (H_{12}' keep the $p_a\dot{a}$ term), we might have been just lucky. For example consider the following scenario where we substituted all p's using their corresponding constraints. To simplify thing let's just assume that we can solve p's explicitly, *i.e.* $C_i \equiv p_i - f^{(i)}(\overline{q})$ where $\overline{q} = \{q_1, ..., q_n\}$.

First we see how the consistency check for the constraint $C_i = p_i - f^{(i)}(\bar{q})$ sets $\lambda_{m \neq i}$

$$\{C_i, H'\} = \{p_i - f^{(i)}(\bar{q}), \sum_k p_k \dot{q}_k\} + \{C_i, -L + \sum_m \lambda_m C_m\}$$

$$\rightarrow -\lambda_{m \neq i} \propto \{p_i, \sum_k p_k \dot{q}_k\} - \{f^{(i)}(\bar{q}), \sum_k p_k \dot{q}_k\} + ...$$

$$= 0 - \sum_m \sum_k (\partial_{q_m} f^{(i)}(\bar{q}))(\partial_{p_m} p_k) \dot{q}_k + ...$$

$$= -\sum_m \sum_k (\partial_{q_m} f^{(i)}(\bar{q}))(\delta_{mk}) \dot{q}_k + ...$$

$$\lambda_{m \neq i} \propto \sum_k (\partial_{q_k} f^{(i)}(\bar{q})) \dot{q}_k + ...$$

The minus sign in line 2 in front of $\lambda_{m \neq i}$ is because $\lambda_{m \neq i}$ were initially on the RHS. The ... means other terms that we are not interested in. Let's see how $\lambda_{m \neq i}$ is modified when we substituted $f^{(k)}$ for each p_k

$$\{p_i - f^{(i)}(\bar{q}), \sum_k p_k \dot{q}_k\} = \{p_i, \sum_k p_k \dot{q}_k\} - \{f^{(i)}(\bar{q}), \sum_k p_k \dot{q}_k\}$$

$$= 0 - \sum_m \sum_k (\partial_{q_m} f^{(i)}(\bar{q}))(\partial_{p_m} p_k) \dot{q}_k$$

$$= -\sum_m \sum_k (\partial_{q_m} f^{(i)}(\bar{q}))(\delta_{mk}) \dot{q}_k$$

$$\lambda_{m \neq i} \propto \sum_k (\partial_{q_k} f^{(i)}(\bar{q})) \dot{q}_k + ...$$

which is not the same as $\lambda_{m \neq i}$ obtained previously, it is not immediately clear if they will give the same e.o.m's for q_k.

Trivial first class constraint

Before moving on to the actual Dirac lagrangian, there's one more thing we need to discuss. Going back to $H_3 = p_a \dot{a} + p_b \dot{b} - K_1 a \dot{b} + K_2 \dot{a} b + mab$, $p_a = -K_2 b$, $p_b = K_1 a$, $K_1 + K_2 = 1$, there's actually a first class constraint hiding inside, albeit a trivial one, $i.e.$ $C_3 = K_1 + K_2 - 1$, since $K_{1,2}$ do not involve any d.o.f.

Still, it is a redundancy in the definition of the lagrangian/hamiltonian. Thus to be complete, H_3' should be written as

$$H_3' = p_a \dot{a} + p_b \dot{b} - K_1 a \dot{b} + K_2 \dot{a} b + mab +$$

$$\lambda_1(p_b - K_1 a) + \lambda_2(p_a + K_2 b) + \lambda_3(K_1 + K_2 - 1)$$

Since $C_3 \equiv K_1 + K_2 - 1$ is first class, we can *not* solve for λ_3 because the consistency check doesn't even give equations involving λ_3 because $\{C_3, f(p,q)\} = 0$ for any function $f(p,q)$.

I mention this here (even though it is trivial) because it is still a polysemousness in our formulation and a potential source of confusion as per my discussion on integration by parts above.

Other equations of motion

In the above discussion I have chosen only $\dot{a}$ because this is the one that has a problem with the traditional hamiltonian approach (due to the absence of p_a). Other e.o.m's for $\dot{b}, \dot{p}_a, \dot{p}_b$ can be calculated in a straighforward manner from H_{12}^T and H_{11}^T. Below we show them through the three ways we have used to calculate $\dot{a}$ above.

<table>
<tr><td colspan="2" align="center">$\dot{b}$</td></tr>
<tr>
<td>

$\begin{aligned} \{b, H_{12}^T\} &= \{b, H_{12}\} + \lambda_m \{b, C_m\} \\ &= (\dot{b}) + (\lambda_1) \\ &= mb \end{aligned}$

</td>
<td>

$\begin{aligned} \{b, H_{11}^T\} &= \{b, H_{11}\} + \lambda_m \{b, C_m\} \\ &= (\dot{b}) + (\lambda_1) \\ &= mb \end{aligned}$

</td>
</tr>
<tr>
<td>

$\begin{aligned} \frac{\partial H_{12}^T}{\partial p_b} &= \frac{\partial H_{12}}{\partial p_b} + \lambda_m \frac{\partial C_m}{\partial p_b} \\ &= (\dot{b}) + (\lambda_1) \\ &= mb \end{aligned}$

</td>
<td>

$\begin{aligned} \frac{\partial H_{11}^T}{\partial p_b} &= \frac{\partial H_{11}}{\partial p_b} + \lambda_m \frac{\partial C_m}{\partial p_b} \\ &= (\dot{b}) + (\lambda_1) \\ &= mb \end{aligned}$

</td>
</tr>
<tr>
<td>

$\begin{aligned} \{b, H_{12}^T\}^* &= \{b, H_{12}\} - [b, H_{12}]_{mn} \\ &= (\dot{b}) - [b, H_{12}]_{12} - [b, H_{12}]_{21} \\ &= (\dot{b}) - (1)(1)\{p_a, H_{12}\} - 0 \\ &= mb \end{aligned}$

</td>
<td>

$\begin{aligned} \{b, H_{11}^T\}^* &= \{b, H_{11}\} - [b, H_{11}]_{mn} \\ &= (\dot{b}) - [b, H_{11}]_{12} - [b, H_{11}]_{21} \\ &= (\dot{b}) - (1)(1)\{p_a, H_{11}\} - 0 \\ &= mb \end{aligned}$

</td>
</tr>
</table>

where I have introduced a shorthand notation

$$[f, g]_{mn} \equiv \{f, C_m\}\{C_m, C_n\}^{-1}\{C_n, g\}$$

$\dot{p}_a$	
$\begin{aligned}\{p_a, H_{12}^T\} &= \{p_a, H_{12}\} + \lambda_m\{p_a, C_m\} \\ &= -(-\dot{b} + mb) + (\lambda_1) \\ &= 0\end{aligned}$	$\begin{aligned}\{p_a, H_{11}^T\} &= \{p_a, H_{11}\} + \lambda_m\{p_a, C_m\} \\ &= -(-\dot{b} + mb) + (\lambda_1) \\ &= 0\end{aligned}$
$\begin{aligned}\frac{\partial H_{12}^T}{\partial a} &= \frac{\partial H_{12}}{\partial a} + \lambda_m \frac{\partial C_m}{\partial a} \\ &= (-\dot{b} + mb) - (\lambda_1) \\ &= 0\end{aligned}$	$\begin{aligned}\frac{\partial H_{11}^T}{\partial a} &= \frac{\partial H_{11}}{\partial a} + \lambda_m \frac{\partial C_m}{\partial a} \\ &= (-\dot{b} + mb) - (\lambda_1) \\ &= 0\end{aligned}$
$\begin{aligned}\{p_a, &H_{12}^T\}^* \\ &= \{p_a, H_{12}\} - [p_a, H_{12}]_{mn} \\ &= \{p_a, H_{12}\} - [p_a, H_{12}]_{12} - [p_a, H_{12}]_{21} \\ &= \{p_a, H_{12}\} - (1)(1)\{p_a, H_{12}\} - 0 \\ &= 0\end{aligned}$	$\begin{aligned}\{p_a, &H_{11}^T\}^* \\ &= \{p_a, H_{11}\} - [p_a, H_{11}]_{mn} \\ &= \{p_a, H_{11}\} - [p_a, H_{11}]_{12} - [p_a, H_{11}]_{21} \\ &= \{p_a, H_{11}\} - (1)(1)\{p_a, H_{11}\} - 0 \\ &= 0\end{aligned}$

$\dot{p}_b$	
$\begin{aligned}\{p_b, H_{12}^T\} &= \{p_b, H_{12}\} + \lambda_m\{p_b, C_m\} \\ &= -ma\end{aligned}$	$\begin{aligned}\{p_a, H_{11}^T\} &= \{p_a, H_{11}\} + \lambda_m\{p_a, C_m\} \\ &= -ma\end{aligned}$
$\begin{aligned}\frac{\partial H_{12}^T}{\partial b} &= \frac{\partial H_{12}}{\partial b} + \lambda_m \frac{\partial C_m}{\partial b} \\ &= ma\end{aligned}$	$\begin{aligned}\frac{\partial H_{11}^T}{\partial b} &= \frac{\partial H_{11}}{\partial b} + \lambda_m \frac{\partial C_m}{\partial b} \\ &= ma\end{aligned}$
$\begin{aligned}\{p_b, &H_{12}^T\}^* \\ &= \{p_b, H_{12}\} - [p_b, H_{12}]_{mn} \\ &= \{p_b, H_{12}\} - [p_b, H_{12}]_{12} - [p_b, H_{12}]_{21} \\ &= -ma - 0 - 0 \\ &= -ma\end{aligned}$	$\begin{aligned}\{p_b, &H_{11}^T\}^* \\ &= \{p_b, H_{11}\} - [p_b, H_{11}]_{mn} \\ &= \{p_b, H_{11}\} - [p_b, H_{11}]_{12} - [p_b, H_{11}]_{21} \\ &= -ma - 0 - 0 \\ &= -ma\end{aligned}$

Note that $\dot{p} = -\partial H/\partial q$ so there's an extra minus sign on the second row that was omitted to avoid clutter.

It might seem redundant to list the above e.o.m's. After all, they are just algebraic exercise. The take home message is that the above e.o.m's work because of the role of λ_m's. If you look carefully, without these multipliers we do not get the correct e.o.m's. But there's an even more important lesson to note. We only used H_{12}^T and H_{11}^T where the time derivative was not distributed among a and b. Inspite of this fact we still recovver

all e.om's thanks to the multipliers λ_m's. This only shows how powerful and crucial the consistency checks, $\{C_m, H'\}$, are. Only through this process do we solve each λ_m which in turn gives us the correct e.o.m's. This is why we should not set $C_m \to 0$ before taking Poisson brackets, otherwise we get the wrong e.o.m's (this point was mentioned in Dirac's book only in passing, but now you see it more vividly).

Moral of the story

We have discussed a few important thing when dealing with the Dirac hamiltonian and other hamiltonians in general, they are:

- Doing integration by parts is tricky, consistency must be maintained at every step and since the hamiltonian keeps memory of integration by parts, it might come back to bite later on.

- Applying a constraint directly to p might still work for simple cases, application to other general cases might yield the wrong result.

- We need *not* distribute the time derivative among a and b and this thanks to the consistency checks.

- The *importance* and *efficacy* of the consistency checks, $\{C_m, H'\} \approx 0$, can not be stressed enough, only through this process can we get the correct equations of motion.

Application to Dirac hamiltonian

Let us now apply what we learned above to the Dirac hamiltonian, for convenience I reproduce the lagrangian and the corresponding hamiltonian below and since we know that we do not need to distribute the time derivative, we will keep the time derivative only on ψ and *not* on $\overline{\psi}$. Also, to keep things simple we will set $p_{\overline{\psi}} = 0$ directly in the

hamiltonian.

$$\mathcal{L} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - \overline{\psi}\left(e\gamma^\mu A_\mu + m\right)\psi$$
$$= i\overline{\psi}\gamma^\mu\partial_\mu\psi - \overline{\psi}M\psi, \qquad M = e\gamma^\mu A_\mu + m$$
$$p_\psi = i\overline{\psi}\gamma^0 \;\rightarrow\; C_1 = p_\psi - i\overline{\psi}\gamma^0$$
$$p_{\overline{\psi}} = 0 \qquad \rightarrow\; C_2 = p_{\overline{\psi}}$$
$$\mathcal{H}' = p_\psi\partial_0\psi - i\overline{\psi}\gamma^\mu\partial_\mu\psi + \overline{\psi}M\psi + \lambda_1(p_\psi - i\overline{\psi}\gamma^0) + \lambda_2(p_{\overline{\psi}})$$

and the corresponding equations of motion derived from the lagrangian are

$$\delta S = \int d^4x \;\delta\overline{\psi}(i\gamma^\mu\partial_\mu\psi - M\psi) + (-i\partial_\mu(\overline{\psi}\gamma^\mu)\psi - \overline{\psi}M)\delta\psi$$
$$\rightarrow 0 = i\gamma^\mu\partial_\mu\psi - M\psi$$
$$\rightarrow 0 = -i\partial_\mu(\overline{\psi}\gamma^\mu) - \overline{\psi}M$$

again we see the strangeness of getting the e.o.m of ψ from varying $\overline{\psi}$ and vice versa

Before we take possion brackets among different quantities, we need to straighten out a few things.

- First, we are taking Poisson brackets w.r.t H' and *not* $\mathcal{H}'$. At first this looks strange since $H' = \int_x \mathcal{H}'$ and so it is *not* a function of $\vec{x}$ and its functional derivative doesn't make sense since H' is no longer a function of $\vec{x}$. But this is what we need to get the correct e.o.m's otherwise we will have $\delta(\vec{x} - \vec{w})$ all over the place and the way we do it is

$$\{f(\vec{x}), H'\} = \int d^3y\{f(\vec{x}), \mathcal{H}'(y)\}$$

- Second, for field theory, we take the possion bracket at two different spatial locations but at the *same* moment in time, *i.e.* $\{f(\vec{x},t), g(\vec{y},t)\}$.

- Third, since we are dealing with fields, we are going to use functional derivative, and functional derivative are defined only for the spatial variables *not* the time variable, *i.e.*

$$\frac{\delta f(\vec{x},t)}{\delta f(\vec{y},t)} = \delta(\vec{x} - \vec{y})$$

33

the delta function is only spatial. Why is this? because a field is nothing more than just a collection of degrees of freedom, one at every point in space, $\phi(\vec{x}, t) = \sum_i \phi_i(t)$ and thus we don't touch the time variable.

- Fourth and more importantly we have new terms that have to be handled correctly, *i.e.* terms that contain spatial derivative, for example

$$\frac{\delta(-i\overline{\psi}\gamma^\mu\partial_\mu\psi)}{\delta\psi} = \frac{\delta(-i\overline{\psi}\gamma^0\partial_0\psi)}{\delta\psi} + \frac{\delta(-i\overline{\psi}\gamma^i\partial_i\psi)}{\delta\psi}$$

Let's tackle them one term at a time, it is extremely helpful to explicitly specify the spacetime coordinates of ψ and also on γ and the derivative since we will be doing plenty of changes of coordinates. It is also helpful to supress the time variable since it has no effect on the functional derivatives

$$\frac{\delta(-i\overline{\psi}\gamma^{i_x}\partial_{i_x}\psi)}{\delta\psi} = -i\overline{\psi}\gamma^{i_x}\frac{\delta(\partial_{i_x}\psi)}{\delta\psi}$$

$$\rightarrow \frac{\delta(\partial_{i_x}\psi(\vec{x}))}{\delta\psi(\vec{w})} = \frac{\delta}{\delta\psi(\vec{w})}\int d^3y\ \partial_{i_y}\psi(\vec{y})\delta(\vec{y}-\vec{x})$$

$$= -\frac{\delta}{\delta\psi(\vec{w})}\int d^3y\ \psi(\vec{y})\partial_{i_y}\delta(\vec{y}-\vec{x})$$

$$= -\int d^3y\ \frac{\delta\psi(\vec{y})}{\delta\psi(\vec{w})}\partial_{i_y}\delta(\vec{y}-\vec{x})$$

$$= -\int d^3y\ \delta(\vec{y}-\vec{w})\partial_{i_y}\delta(\vec{y}-\vec{x})$$

$$= \int d^3y\ \partial_{i_y}\delta(\vec{y}-\vec{w})\delta(\vec{y}-\vec{x})$$

$$= \partial_{i_x}\delta(\vec{x}-\vec{w})$$

$$\rightarrow \frac{\delta(-i\overline{\psi}\gamma^i\partial_i\psi)}{\delta\psi(\vec{w})} = -i\overline{\psi}\gamma^{i_x}\partial_{i_x}\delta(\vec{x}-\vec{w})$$

where ∂_{i_x} means that partial derivative w.r.t to the coordinate $\vec{x}$, ditto with γ^{i_x}. The integration by parts on $\delta(\vec{y}-\vec{w})\partial_{i_y}\delta(\vec{y}-\vec{x}) \rightarrow -\partial_{i_y}\delta(\vec{y}-\vec{w})\delta(\vec{y}-\vec{x})$ is needed because the whole term is multiplied by γ^{i_x}, otherwise we'd get $\partial_{i_w}\delta(\vec{w}-\vec{x})$ in the last line and it wouldn't contract with γ^{i_x}. Another (better) reason is because our initial delta function was $\delta(\vec{y}-\vec{x})$ sowhatever multiplies this was the function to be integrated and in this case this is $\partial_{i_y}\delta(\vec{y}-\vec{w})$.

Now there are two integrals, one associated to $H' = \int_x \mathcal{H}'$ and another from the Poisson bracket itself $\{,\} \to \int d^3w \; \frac{\delta}{\delta\psi(w)}\frac{\delta}{\delta p_\psi(w)} - \frac{\delta}{\delta p_\psi(w)}\frac{\delta}{\delta\psi(w)}$

$$\int d^3x \int d^3w \; \frac{\delta(-i\overline{\psi}\gamma^{i_x}\partial_{i_x}\psi)}{\delta\psi(\vec{w})} = -\int d^3w \int d^3x \; i\overline{\psi}\gamma^{i_x}\partial_{i_x}\delta(\vec{x}-\vec{w})$$

$$= \int d^3w \int d^3x \; i\partial_{i_x}(\overline{\psi}\gamma^{i_x})\delta(\vec{x}-\vec{w})$$

$$= i\int d^3w \; \partial_{i_w}(\overline{\psi}(\vec{w})\gamma^{i_w})$$

This is no different from doing integration by parts before taking the functional derivative. My initial attempt was to avoid doing spatial integration by parts because it is somewhat random and introduce yet another trivial first class constraint like the one discussed before but in terms of the spatial derivative instead. Looks like there's no escaping spatial integration by parts.

Now, it is quite a different story with $\delta(-i\overline{\psi}\gamma^0\partial_0\psi)/\delta\psi$

$$\frac{\delta(-i\overline{\psi}\gamma^0\partial_0\psi)}{\delta\psi} = -i\overline{\psi}\gamma^0\frac{\delta\partial_0\psi}{\delta\psi}$$

$$\frac{\delta\partial_t\psi(\vec{x},t)}{\delta\psi(\vec{w},t)} = \frac{\delta}{\delta\psi(\vec{w},t)}\int dT \; \partial_T\psi(\vec{x},T)\delta(T-t)$$

$$= -\frac{\delta}{\delta\psi(\vec{w},t)}\int dT \; \psi(\vec{x},T)\partial_T\delta(T-t)$$

$$= -\int dT \; \frac{\delta\psi(\vec{x},T)}{\delta\psi(\vec{w},T)}\partial_T\delta(T-t)$$

$$= -\int dT \; \delta(\vec{x}-\vec{w})\partial_T\delta(T-t)$$

$$= -\delta(\vec{x}-\vec{w})\int dT \; (1)\partial_T\delta(T-t)$$

$$= \delta(\vec{x}-\vec{w})\int dT \; \partial_T(1)\delta(T-t)$$

$$= 0$$

Combining these two results we get

$$\int d^3w \frac{\delta(-i\overline{\psi}\gamma^\mu\partial_\mu\psi)}{\delta\psi} = \int d^3w \frac{\delta(-i\overline{\psi}\gamma^0\partial_0\psi)}{\delta\psi} + \frac{\delta(-i\overline{\psi}\gamma^i\partial_i\psi)}{\delta\psi}$$

$$= \int d^3w \; i\partial_{i_w}(\overline{\psi}(\vec{w})\gamma^{i_w})$$

35

We can now proceed with our various Poisson brackets, let's first calculate $\{C_m, C_n\}$, with $C_1 = (p_\psi - i\overline{\psi}\gamma^0)$ and $C_2 = (p_{\overline{\psi}})$

$$\{C_1(\vec{x}), C_2(\vec{y})\} = \int d^3w \left(\frac{\delta C_1}{\delta \psi(\vec{w})}\frac{\delta C_2}{\delta p_\psi(\vec{w})} - \frac{\delta C_1}{\delta p_\psi(\vec{w})}\frac{\delta C_2}{\delta \psi(\vec{w})} + \frac{\delta C_1}{\delta \overline{\psi}(\vec{w})}\frac{\delta C_2}{\delta p_{\overline{\psi}}(\vec{w})} - \frac{\delta C_1}{\delta p_{\overline{\psi}}(\vec{w})}\frac{\delta C_2}{\delta \overline{\psi}(\vec{w})} \right)$$

$$= \int d^3w \ \left((0)(0) - \delta_{wx}(0) + (-i\gamma^0 \delta_{wx})(\delta_{wy}) - (0)(0) \right)$$

$$= -\int d^3w \ i\gamma^0 \delta_{wx}\delta_{wy}$$

$$= -i\gamma^0 \delta(\vec{x} - \vec{y})$$

where we have introduced the shorthand notation $\delta_{wx} \equiv \delta(\vec{w} - \vec{x})$. Let's now do the consistency checks for $C_1 = (p_\psi - i\overline{\psi}\gamma^0)$ and $C_2 = (p_{\overline{\psi}})$ by taking their possion brackets with H'.

$$\int_y \{C_1(\vec{x}), \mathcal{H}'(\vec{y})\} = \int_{yw} \left(\frac{\delta C_1}{\delta \psi(\vec{w})}\frac{\delta \mathcal{H}'}{\delta p_\psi(\vec{w})} - \frac{\delta C_1}{\delta p_\psi(\vec{w})}\frac{\delta \mathcal{H}'}{\delta \psi(\vec{w})} + \frac{\delta C_1}{\delta \overline{\psi}(\vec{w})}\frac{\delta \mathcal{H}'}{\delta p_{\overline{\psi}}(\vec{w})} - \frac{\delta C_1}{\delta p_{\overline{\psi}}(\vec{w})}\frac{\delta \mathcal{H}'}{\delta \overline{\psi}(\vec{w})} \right)$$

$$= \int_{yw} \ \left(0 - \delta_{xw}\left\{ i\partial_i(\overline{\psi}\gamma^i) + \overline{\psi}M \right\}\delta_{wy} - \lambda_2 \delta_{yw}\delta_{wx}i\gamma^0 - 0 \right)$$

$$= -i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M - i\lambda_2\gamma^0$$

where we have introduced a shorthand notation $\int_x \equiv \int d^3x$. This means that

$$i\lambda_2\gamma^0 = -i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M$$

$$(-i)(i)\lambda_2(\gamma^0\gamma^0) = -\partial_i(\overline{\psi}\gamma^i)\gamma^0 + i\overline{\psi}M\gamma^0$$

$$\lambda_2 = -\partial_i(\overline{\psi}\gamma^i)\gamma^0 + i\overline{\psi}M\gamma^0$$

where $\gamma^0\gamma^0 = 1_{4\times4}$, a 4×4 identity matrix. Note also that without the integral $\int_y$ we will have leftover delta functions. The consistency check for $C_1 = (p_{\overline{\psi}})$ is

$$\int_y \{C_2(\vec{x}), \mathcal{H}'(\vec{y})\} = \int_{yw} \left(\frac{\delta C_2}{\delta \psi(\vec{w})}\frac{\delta \mathcal{H}'}{\delta p_\psi(\vec{w})} - \frac{\delta C_2}{\delta p_\psi(\vec{w})}\frac{\delta \mathcal{H}'}{\delta \psi(\vec{w})} + \frac{\delta C_2}{\delta \overline{\psi}(\vec{w})}\frac{\delta \mathcal{H}'}{\delta p_{\overline{\psi}}(\vec{w})} - \frac{\delta C_2}{\delta p_{\overline{\psi}}(\vec{w})}\frac{\delta \mathcal{H}'}{\delta \overline{\psi}(\vec{w})} \right)$$

$$= \int_{yw} \ \left(0 - 0 + 0 - \delta_{xw}\left\{ -i\gamma^\mu\partial_\mu\psi + M\psi - \lambda_1 i\gamma^0 \right\}\delta_{wy} \right)$$

$$= i\gamma^\mu\partial_\mu\psi - M\psi + i\gamma^0\lambda_1$$

which means that

$$\lambda_1 = -\gamma^0\gamma^\mu\partial_\mu\psi - i\gamma^0 M\psi$$

We are now ready to calculate the e.o.m's three way just like before. First up, the most problematic of all

$$\dot{\overline{\psi}}(\vec{x}) = \int d^3y \{\overline{\psi}(\vec{x}), \mathcal{H}^T(\vec{y})\} = \int d^3y \{\overline{\psi}(\vec{x}), \mathcal{H}(\vec{y})\} + \lambda_m \{\overline{\psi}(\vec{x}), C_m(\vec{y})\}$$

$$= 0 + \lambda_2 \int d^3y \int d^3w \; \frac{\delta \overline{\psi}(\vec{x})}{\delta \overline{\psi}(\vec{w})} \frac{\delta C_2(\vec{y})}{\delta p_{\overline{\psi}}(\vec{w})}$$

$$= \lambda_2 \int d^3y \int d^3w \; \delta(\vec{x} - \vec{w})\delta(\vec{y} - \vec{w})$$

$$\partial_0 \overline{\psi} = -\partial_i(\overline{\psi}\gamma^i)\gamma^0 + i\overline{\psi}M\gamma^0$$

$$\rightarrow 0 = -i\partial_\mu(\overline{\psi}\gamma^\mu) - \overline{\psi}M$$

which is the correct adjoint Dirac equation. Now you see that we need to take the Poisson bracket w.r.t H' instead of $\mathcal{H}'$ because we need $\int_y$ to get rid of the delta function. Our second method yields

$$\dot{\overline{\psi}}(\vec{x}) = \frac{\delta H^T}{\delta p_{\overline{\psi}}(\vec{x})} = \frac{\delta H}{\delta p_{\overline{\psi}}} + \int d^3y \; \lambda_m \frac{\delta C_m(\vec{y})}{\delta p_{\overline{\psi}}(\vec{x})}$$

$$= 0 + \int d^3y \; \lambda_2(\vec{y})\delta_{xy}$$

$$= \lambda_2(\vec{x})$$

which is consistent with our first method. To figure out what coordinate to use in the functional derivative we just need to math LHS and RHS, since LHS is a function of $\vec{x}$ we need to take the functional derivative w.r.t $\delta/\delta p_{\overline{\psi}}(\vec{x})$.

Lastly our dear friend the Dirac bracket but first we need to calculate the inverse matrix $\{C_m, C_n\}^{-1}$. In the case of field theory, this is a bit tricky because beside the indices $_{mn}$, the Poisson bracket also carries spatial coordinates and we will denote it

$$\Delta^{\mathbf{xy}}_{mn} \equiv \{C_m(\vec{x}), C_n(\vec{y})\}$$

so in a sense we have extra indices to take care of and the identity matrix now carries the same extra indices as well which we will denote $\mathbf{1}^{\mathbf{xy}}_{mn} \equiv \delta_{xy}1_{m\times n}$. One thing to note is that the indices $\mathbf{xy}$ are continuous and their contractions require integration. The inverse $(\Delta^{\mathbf{xy}}_{mn})^{-1}$ we want is thus

$$\int_y \Delta^{\mathbf{xy}}_{mn}(\Delta^{\mathbf{yz}}_{nk})^{-1} = \mathbf{1}^{\mathbf{xz}}_{mk}$$

From our calculation $\{C_1(\vec{x}), C_2(\vec{y})\}$ above we get

$$\int_y \{C_m(\vec{x}), C_n(\vec{y})\}\{C_n(\vec{y}), C_k(\vec{z})\}^{-1} = \delta(\vec{x} - \vec{y})1_{m \times n}$$

$$\int_y \begin{pmatrix} 0 & -i\gamma^0 \delta_{xy} \\ i\gamma^0 \delta_{xy} & 0 \end{pmatrix} \begin{pmatrix} 0 & -i\gamma^0 \delta_{yz} \\ i\gamma^0 \delta_{yz} & 0 \end{pmatrix} = \begin{pmatrix} \delta_{xz} & 0 \\ 0 & \delta_{xz} \end{pmatrix}$$

Which means that in this case $(\Delta_{mn}^{\mathbf{xy}})^{-1} = \Delta_{mn}^{\mathbf{xy}}$ *i.e.*

$$\{C_m(\vec{x}), C_n(\vec{y})\}^{-1} = \begin{pmatrix} 0 & -i\gamma^0 \delta_{xy} \\ i\gamma^0 \delta_{xy} & 0 \end{pmatrix}$$

More heuristically, since the left hand side is a function of $\vec{x}$ and $\vec{y}$, the right hand side must be as well and the only reasonable candidate is the delta function. Yet another way of looking at it is that a field is just a collection of oscillators, one at each point in space and thus the coordinate variable is just an index

$$\sum_y \{C_x^{(m)}, C_y^{(n)}\}\{C_y^{(n)}, C_k^{(z)}\} = \int_y \{C_m(\vec{x}), C_n(\vec{y})\}\{C_n(\vec{y}), C_k(\vec{z})\}^{-1}$$

and using the usual dictionary of passing from discrete to continuum we need to change $\sum_y \to \int_y$.

Going back to the e.o.m of $\overline{\psi}$, we take its Dirac bracket with H^T but following Dirac, since all constraints are second class, we'll set them to zero *inside* $H^T \to H^T = H$

$$\dot{\overline{\psi}}(\vec{x}) = \int_y \{\overline{\psi}(\vec{x}), \mathcal{H}^T(\vec{y})\}^*$$

$$= \int_y \{\overline{\psi}(\vec{x}), \mathcal{H}(\vec{y})\} - \int_{zuy} \{\overline{\psi}(\vec{x}), C_2(\vec{z})\}\{C_2(\vec{z}), C_1(\vec{u})\}^{-1}\{C_1(\vec{u}), \mathcal{H}(\vec{y})\}$$

$$= 0 - \int_{zuy} \{\overline{\psi}(\vec{x}), C_2(\vec{z})\}\{C_2(\vec{z}), C_1(\vec{u})\}^{-1}\{C_1(\vec{u}), \mathcal{H}(\vec{y})\}$$

$$= -\int_{zuy} \delta_{xz}(i\gamma^0 \delta_{zu})\left(-\delta_{uy}\left\{i\partial_i(\overline{\psi}\gamma^i) + \overline{\psi}M\right\}\right)$$

$$\partial_0 \overline{\psi} = i\left\{i\partial_i(\overline{\psi}\gamma^i) + \overline{\psi}M\right\}\gamma^0$$

$$\to 0 = -i\partial_\mu(\overline{\psi}\gamma^\mu) - \overline{\psi}M$$

which is consistent with our previous calculation.

The rest of the work is straightforward, we just need to apply the same steps to get $\dot{\psi}, \dot{p}_{\overline{\psi}}, \dot{p}_\psi$ and they are given in the following tables

38

$\dot{\psi}$
$\{\psi, H^T\} = \int_y \{\psi, \mathcal{H}(\vec{y})\} + \int_y \lambda_m \{\psi, C_m(\vec{y})\}$
$\quad = (\dot{\psi}) + (\lambda_1)$
$\quad \cancel{\dot{\psi}} = \cancel{\dot{\psi}} - \gamma^0\gamma^\mu\partial_\mu\psi - i\gamma^0 M\psi$
$\quad 0 = i\gamma^\mu\partial_\mu\psi - M\psi$
$\frac{\delta H^T}{\delta p_\psi(\vec{x})} = \int_y \frac{\delta\mathcal{H}(\vec{y})}{\delta p_\psi(\vec{x})} + \int_y \lambda_m \frac{\delta C_m(\vec{y})}{\delta p_\psi(\vec{x})}$
$\quad = (\dot{\psi}) + (\lambda_1)$
$\quad \cancel{\dot{\psi}} = \cancel{\dot{\psi}} - \gamma^0\gamma^\mu\partial_\mu\psi - i\gamma^0 M\psi$
$\quad 0 = i\gamma^\mu\partial_\mu\psi - M\psi$
$\{\psi, H^T\}^* = \int_y \{\psi, \mathcal{H}\} - \int_y [\psi, \mathcal{H}]_{mn}$
$\quad = (\dot{\psi}) - \int_y [\psi, \mathcal{H}]_{12} - \int_y [\psi, \mathcal{H}]_{21}$
$\quad = \dot{\psi} - \int_y (1)(-i\gamma^0)\{p_{\overline{\psi}}, \mathcal{H}(\vec{y})\} - 0$
$\quad \cancel{\dot{\psi}} = \cancel{\dot{\psi}} + i\gamma^0(i\gamma^\mu\partial_\mu\psi - M\psi)$
$\quad 0 = i\gamma^\mu\partial_\mu\psi - M\psi$

I have introduced a shorthand notation

$$[f(\vec{x}), g(\vec{y})]_{mn} \equiv \int_{zu} \{f(\vec{x}), C_m(\vec{z})\}\{C_m(\vec{z}), C_n(\vec{u})\}^{-1}\{C_n(\vec{u}), g(\vec{y})\}$$

$\dot{p}_{\overline{\psi}}$
$\{p_{\overline{\psi}}, H^T\} = \int_y \{p_{\overline{\psi}}, \mathcal{H}(\vec{y})\} + \int_y \lambda_m \{p_{\overline{\psi}}, C_m(\vec{y})\}$
$\quad = \int_y \delta_{xy}(i\gamma^\mu\partial_\mu\psi - M\psi + (\lambda_1 i\gamma^0))$
$\quad = 0$
$\frac{\delta H^T}{\delta\overline{\psi}(\vec{x})} = \int_y \frac{\delta\mathcal{H}(\vec{y})}{\delta\overline{\psi}(\vec{x})} + \int_y \lambda_m \frac{\delta C_m(\vec{y})}{\partial\overline{\psi}(\vec{x})}$
$\quad = \int_y \delta_{xy}(i\gamma^\mu\partial_\mu\psi - M\psi + (\lambda_1 i\gamma^0))$
$\quad = 0$
$\{p_{\overline{\psi}}, H^T\}^* = \int_y \{p_{\overline{\psi}}, \mathcal{H}\} - \int_y [p_{\overline{\psi}}, \mathcal{H}]_{mn}$
$\quad = \int_y \{p_{\overline{\psi}}, \mathcal{H}\} - \int_y [p_{\overline{\psi}}, \mathcal{H}]_{12} - \int_y [p_{\overline{\psi}}, \mathcal{H}]_{21}$
$\quad = \int_y (\{p_{\overline{\psi}}, H\} - (i\gamma^0)(-i\gamma^0)\{p_{\overline{\psi}}, H\} - 0)$
$\quad = \int_y (\{p_{\overline{\psi}}, H\} - \{p_{\overline{\psi}}, H\})$
$\quad = 0$

note that there should be an extra minus sign in front of the terms in the second row $\dot{p} = -\partial H/\partial q$ but it doesn't affect the conclusion in this case.

$\dot{p}_\psi$
$\begin{aligned} \{p_\psi, H^T\} &= \int_y \{p_\psi, \mathcal{H}(\vec{y})\} + \int_y \lambda_m \{p_\psi, C_m(\vec{y})\} \\ \dot{p}_\psi &= \int_y \delta_{yx}((-i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M) + (0)) \\ i\partial_t\overline{\psi}\gamma^0 &= -i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M \\ 0 &= -i\partial_\mu(\overline{\psi}\gamma^\mu) - \overline{\psi}M \end{aligned}$
$\begin{aligned} -\frac{\delta H^T}{\delta\psi(\vec{x})} &= -\int_y \frac{\delta\mathcal{H}(\vec{y})}{\delta\psi(\vec{x})} - \int_y \lambda_m \frac{\delta C_m(\vec{y})}{\delta\psi(\vec{x})} \\ \dot{p}_\psi &= \int_y \delta_{yx}((-i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M) + (0)) \\ i\partial_t\overline{\psi}\gamma^0 &= -i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M \\ 0 &= -i\partial_\mu(\overline{\psi}\gamma^\mu) - \overline{\psi}M \end{aligned}$
$\begin{aligned} \{p_\psi, H^T\}^* &= \int_y \{p_\psi, \mathcal{H}\} - \int_y [p_\psi, \mathcal{H}]_{mn} \\ &= \int_y \{p_\psi, H_{12}\} - \int_y [p_\psi, \mathcal{H}]_{11} - \int_y [p_\psi, \mathcal{H}]_{21} \\ \dot{p}_\psi &= \int_y \delta_{yx}((-i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M) - (0) - (0)) \\ i\partial_t\overline{\psi}\gamma^0 &= -i\partial_i(\overline{\psi}\gamma^i) - \overline{\psi}M \\ 0 &= -i\partial_\mu(\overline{\psi}\gamma^\mu) - \overline{\psi}M \end{aligned}$

where I have applied the constraint $p_\psi = i\overline{\psi}\gamma^0$ which is permitted since $C_m = 0$ are also part of the equations of motion.

We have now recovered all equations of motion of the Dirac hamiltonian using the Dirac-Bergmann algorithm which is needed since we have second class constraints. Our next task it to apply Dirac-Bergmann algorithm to the photons.

C. Photonic constraints

The constraint situation here is different from the fermionic case as we will get a secondary constraint. For this discussion we will need to revisit our coastal cousins once again and it does make a big difference which metric you use. We will add a mass term for the photon. Massive photons are pedagogical although sadly they are almost always ignored in standard textbooks. We immediately have a difficulty here, what should the

mass term be? It depends where you get it from

$$m^2 A^\mu A_\mu = m^2(-A_0 A_0 + A_i A_i) \quad \text{east coast}$$
$$= m^2(+A_0 A_0 - A_i A_i) \quad \text{west coast}$$

Again our choice matters but in this case it has severe consequences. Well, let's see what the big deal is, the massive photon lagrangian is

$$\mathcal{L}_{\text{mph}} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu$$
$$= \frac{1}{2} \left(\vec{E}^2 - \vec{B}^2 \right) + \frac{1}{2} m^2 A_\mu A^\mu$$

We would like to include the sources for our massive A_μ as well

$$\mathcal{L}_{\text{mph}} = \frac{1}{2} \left(\vec{E}^2 - \vec{B}^2 \right) + \frac{1}{2} m^2 A_\mu A^\mu - J^\mu A_\mu$$

And so in our peaceful west coast I have chosen the mass term as $+\frac{1}{2} m^2 A_\mu A^\mu$ instead of its negative cousin, this might of course backfire but we will see.

Following the usual Dirac-Bergmann recipe we include the primary (second class) constraint which in this case is $p_A^0 = 0 \to C_1 = p_A^0$ in the hamiltonian (of course at this point we don't know whether this constraint is first or second class)

$$\mathcal{H}'_{\text{mph}} = \frac{1}{2} \left(\vec{E}^2 + \vec{B}^2 \right) - A_0 (\partial_i E^i) - \frac{1}{2} m^2 A_\mu A^\mu + J^\mu A_\mu + \lambda_1 p_A^0$$

I have excluded the term $p_A^0 \partial_0 A_0$ in the hamiltonian as per our discussion earlier that this term makes no difference in the end.

Next, we do a consistency check to make sure that the constraints remain fixed in time

$$\dot{p}_A^0(\vec{x}) = \{ p_A^0(\vec{x}), H' \}$$
$$= \int d^3 y \, \{ p_A^0(\vec{x}), \mathcal{H}'(\vec{y}) \}$$
$$= -\int d^3 y \int d^3 w \, \frac{\delta p_A^0(\vec{x})}{\delta p_A^0(\vec{w})} \frac{\delta \mathcal{H}'(\vec{y})}{\delta A_0(\vec{w})}$$
$$= -\int d^3 y \int d^3 w \, \delta(\vec{x} - \vec{w})(-\partial_i E^i(\vec{y}) - m^2 A^0(\vec{y}) + J^0(\vec{y})) \delta(\vec{y} - \vec{w})$$
$$= \partial_i E^i(\vec{x}) + m^2 A^0(\vec{x}) - J^0(\vec{x})$$

In the above calculation we have been careful by including all integrals and spatial coordinates, in our subsequent calculations we might bypass a few things since the result

of the functional derivatives is no different than the usual partial derivatives. I have also dropped the subscript $_{\text{mph}}$ to avoid clutter (and will do so when there's no potential confusion). This is now our secondary constraint, as a consistency check, we see if its time derivative, $\{\partial_i E^i + m^2 A^0 - J^0, H'\}$, is zero, if not we'll get a tertiary second class constraint. First we calculate $\{\partial_i E^i, H'\}$. To calculate it, we can use the trick that a function is an integral of itself with the delta function

$$\frac{\delta \partial_i E^i(\vec{x})}{\delta E^j(\vec{y})} = \frac{\delta}{\delta E^j(\vec{y})} \int d^3z \; \partial_i E^i(\vec{z}) \delta(\vec{z} - \vec{x})$$

$$= -\frac{\delta}{\delta E^j(\vec{y})} \int d^3z \; E^i(\vec{z}) \partial_i \delta(\vec{z} - \vec{x})$$

$$= -\int d^3z \; \frac{\delta E^i(\vec{z})}{\delta E^j(\vec{y})} \partial_i \delta(\vec{z} - \vec{x})$$

$$= -\int d^3z \; \delta^i_j \delta(\vec{z} - \vec{y}) \partial_i \delta(\vec{z} - \vec{x})$$

$$\frac{\delta \partial_i E^i(\vec{x})}{\delta E^j(\vec{y})} = -\partial_j \delta(\vec{y} - \vec{x})$$

Integration by parts revisited

Note that there's a potential uncertainty here. We could have done integration by parts in the second to last line to get

$$= -\int d^3z \; \delta^i_j \delta(\vec{z} - \vec{y}) \partial_i \delta(\vec{z} - \vec{x})$$

$$= \int d^3z \; \delta^i_j \partial_i \delta(\vec{z} - \vec{y}) \delta(\vec{z} - \vec{x})$$

$$\frac{\delta \partial_i E^i(\vec{x})}{\delta E^j(\vec{y})} = \partial_j \delta(\vec{x} - \vec{y})$$

and since the delta function is symmetric $\delta(\vec{x} - \vec{y}) = \delta(\vec{y} - \vec{x})$, it seems like we have lost a minus sign. At this point it is prudent to be more careful. If we write the derivative explicitly, in the first derivation we get

$$\frac{\delta \partial_i E^i(\vec{x})}{\delta E^j(\vec{y})} = -\frac{\partial}{\partial y^j} \delta(\vec{y} - \vec{x})$$

while after doing integration by parts we get

$$\frac{\delta \partial_i E^i(\vec{x})}{\delta E^j(\vec{y})} = \frac{\partial}{\partial x^j} \delta(\vec{x} - \vec{y})$$

and if we do a substitution of $\vec{z} = \vec{x} - \vec{y}$

$$\frac{\partial}{\partial x^j}\delta(\vec{x} - \vec{y}) = \frac{\partial}{\partial z^k}\delta(\vec{z})\frac{\partial z^k}{\partial x^j}$$
$$= \frac{\partial}{\partial z^k}\delta(\vec{z})\delta_j^k$$
$$= \frac{\partial}{\partial z^j}\delta(\vec{z})$$
$$-\frac{\partial}{\partial y^j}\delta(\vec{y} - \vec{x}) = -\frac{\partial}{\partial y^j}\delta(\vec{x} - \vec{y})$$
$$= -\frac{\partial}{\partial z^k}\delta(\vec{z})\frac{\partial z^k}{\partial y^j}$$
$$= -\frac{\partial}{\partial z^k}\delta(\vec{z})(-\delta_j^k)$$
$$= \frac{\partial}{\partial z^j}\delta(\vec{z})$$
$$\rightarrow \frac{\partial}{\partial x^j}\delta(\vec{x} - \vec{y}) = -\frac{\partial}{\partial y^j}\delta(\vec{y} - \vec{x})$$

which shows that the two results are consistent, there is a virtue of being careful after all.

Going back to our consistency check

$$\{\partial_i E^i, H'\} = -\int d^3y \, \frac{\delta \partial_i E^i(\vec{x})}{\delta E^j(\vec{y})}\frac{\delta H'}{\delta A_j(\vec{y})}$$
$$= \int d^3y \, \partial_j\delta(\vec{y} - \vec{x})\frac{\delta H'}{\delta A_j(\vec{y})}$$

as a reminder, the integral $\int d^3y$ is the counterpart of $\sum_j$ in the Poison bracket for multiple d.o.f's and this integral is crucial, otherwise things won't work. The second term in the above integral is

$$\rightarrow \frac{\delta H'}{\delta A_j(\vec{y})} = \int d^3x \, \frac{\delta \mathcal{H}'(\vec{x})}{\delta A_j(\vec{y})} = \int d^3x \, \frac{\delta \frac{1}{2}\vec{B}^2}{\delta A_j(\vec{y})} - m^2 A^j + J^j$$
$$= \int d^3x \, \frac{\delta \frac{1}{4}F^{ik}F_{ik}}{\delta A_j(\vec{y})} - m^2 A^j + J^j$$
$$= \int d^3x \, \frac{1}{4}\delta(\vec{x} - \vec{y})\left\{\left(-\partial_i F^{ij} + \partial_k F^{jk} - \partial^i F_i{}^j + \partial^k F^j{}_k\right) - m^2 A^j + J^j\right\}$$
$$= \frac{1}{4}\left(-\partial_i F^{ij} + \partial_k F^{jk} - \partial_i F^{ij} + \partial_k F^{jk}\right) - m^2 A^j + J^j$$
$$\frac{\delta H'}{\delta A_j(\vec{y})} = \frac{1}{4}\left(-2\partial_i F^{ij} + 2\partial_k F^{jk}\right) - m^2 A^j + J^j$$
$$= \partial_i F^{ji} - m^2 A^j + J^j$$

43

and so

$$\{\partial_i E^i, H'\} = \int d^3y \ \partial_j \delta(\vec{y} - \vec{x}) \ (\partial_i F^{ji} - m^2 A^j + J^j)$$

$$= -\int d^3y \ \delta(\vec{y} - \vec{x}) \ (\partial_j \partial_i F^{ji}(\vec{y}) - m^2 \partial_j A^j + \partial_j J^j)$$

$$\{\partial_i E^i, H'\} = \partial_i \partial_j F^{ij} + m^2 \partial_j A^j - \partial_j J^j$$

$$= m^2 \partial_j A^j - \partial_j J^j$$

The first term is zero because we contract symmetric and antisymmetric objects. All this seems cumbersome, there's a faster way to calculate $\{\partial_i E^i, H\}$

$$\{\partial_i E^i, H'\} = \partial_t(\partial_i E^i) = \partial_i(\partial_t E^i)$$

$$= \partial_i \{E^i, H'\}$$

$$\rightarrow \{E^i, H'\} = -\int d^3y \ \delta_j^i \delta(\vec{x} - \vec{y})(\partial_i F^{ji} - m^2 A^j + J^j)$$

$$= \partial_j F^{ij} + m^2 A^i - J^i$$

and we immediately see that

$$\partial_i \{E^i, H'\} = \partial_i \partial_j F^{ij} + m^2 \partial_i A^i - \partial_i J^i$$

$$\rightarrow \{\partial_i E^i, H'\} = m^2 \partial_i A^i - \partial_i J^i$$

Now we need to take care the second term in the Poisson bracket $\{\partial_i E^i + m^2 A^0 - J^0, H'\}$

$$\{A^0, H'\} = \int d^3y \ \frac{\delta A^0}{\delta A_0(\vec{y})} \frac{\delta H'}{\delta p^0(\vec{y})}$$

$$= \int d^3y \ \delta(\vec{x} - \vec{y})\lambda_1$$

$$\rightarrow m^2\{A^0, H'\} = m^2 \lambda_1$$

Collecting all of our partial results together

$$\{\partial_i E^i + m^2 A^0 - J^0, H'\} = \{\partial_i E^i, H'\} + \{m^2 A^0, H'\} + \{-J^0, H'\}$$

$$= (m^2 \partial_i A^i - \partial_i J^i) + (m^2 \lambda_1) - \langle \partial_0 J^0 \rangle^{\text{fm}}$$

$$\rightarrow \lambda_1 = \frac{1}{m^2}(\langle \partial_0 J^0 \rangle^{\text{fm}} + \partial_i J^i - \partial_i A^i)$$

Some explanation is in order here. I have introduced a strange notation $\langle \ \rangle^{\text{fm}}$ surrounding $\partial_0 J^0$ because since J is not a function of A or p_A, thus H'_{mph} does not generate contact

transformations for J but H'_{fm} does and the complete QED hamiltonian includes both the photonic and fermionic parts, we will encounter this subtlety again later on.

In the above derivation we can solve for λ_1 since all consistency checks are weakly equal to zero. And since this consistency check solves λ_1 it doesn't produce anymore constraints. So we have a total of two constraints $C_1 = p_A^0$ and $C_2 = \partial_i E^i + m^2 A^0 - J^0$. We need to update H' by adding C_2

$$\mathcal{H}' = \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) - A_0(\partial_i E^i) - \frac{1}{2}m^2 A_\mu A^\mu + J^\mu A_\mu + \lambda_1 p_A^0 + \lambda_2(\partial_i E^i + m^2 A^0 - J^0)$$

To solve λ_2 we calculate $\{p_A^0, H'\}$ with the updated H' above (we can't solve λ_2 by taking $\{C_2, H'\}$ because $\{C_2, C_2\} = 0$ and there will be no term with λ_2 left)

$$\{p^0, H'\} = -(-\partial_i E^i - m^2 A^0 + J^0 + m^2 \lambda_2)$$
$$\rightarrow \lambda_2 = \frac{1}{m^2}(\partial_i E^i + m^2 A^0 - J^0)$$

From the above calculation we see that both constraints are second class because their Poisson brackets solve the lagrange multipliers but of course explicitly the Poisson brackets among themselves shouldn't be zero

$$\{p^0(\vec{x}), \partial_i E^i(\vec{y}) + m^2 A^0(\vec{y}) - J^0(\vec{y})\} = -\int d^3w \, \frac{\delta p^0(\vec{x})}{\delta p^0(\vec{w})} \frac{\delta(m^2 A^0(\vec{y}))}{\delta A_0(\vec{w})}$$
$$= -\int d^3w \, \delta(\vec{x} - \vec{w}) \, (m^2 \delta(\vec{y} - \vec{w}))$$
$$= -m^2 \delta(\vec{x} - \vec{y})$$

Now all this has been derived for $m \neq 0$, for the case of massless photons $m = 0$, both constraints would become first class while the number of constraints is the same since if we take $m \rightarrow 0$ we get

$$\{p^0(\vec{x}), H'\} = \partial_i E^i(\vec{x}) - J^0(\vec{x})$$
$$\{\partial_i E^i(\vec{x}) - J^0(\vec{x}), H'\} = 0$$
$$\{p^0(\vec{x}), \partial_i E^i(\vec{y}) - J^0(\vec{y})\} = 0$$

Massive settlement

Going back to our discussion of what the sign of the mass term should be in the lagrangian, we can see a tremendous difference between $+$ and $-$. Since both constraints

are second class (in the massive case) we can set them to zero in the hamiltonian. However, in doing so we unwittingly solve for A_0 as well since

$$0 = \partial_i E^i + m^2 A^0 - J^0$$

$$\rightarrow A_0 = \frac{1}{m^2}(J^0 - \partial_i E^i)$$

substituting this result back into our hamiltonian we get

$$
\begin{aligned}
\mathcal{H}^T &= \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) - A_0(\partial_i E^i) - \frac{1}{2}m^2(A_0 A_0 - A_i A_i) + J^0 A_0 + J^i A_i \\
&= \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) - \frac{(J^0 - \partial_i E^i)}{m^2}(\partial_i E^i) - \frac{1}{2}m^2\left(\frac{(J^0 - \partial_i E^i)^2}{m^4} - A_i A_i\right) \\
&\quad + J^0\frac{(J^0 - \partial_i E^i)}{m^2} + J^i A_i \\
&= \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) + \frac{(J^0 - \partial_i E^i)^2}{m^2} - \frac{1}{2}m^2\left(\frac{(J^0 - \partial_i E^i)^2}{m^4} - A_i A_i\right) + J^i A_i \\
&= \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) + \frac{(J^0 - \partial_i E^i)^2}{2m^2} + \frac{m^2}{2}A_i A_i + J^i A_i
\end{aligned}
$$

If we set $J^\mu = 0$ for a moment, we see that the above hamiltonian is positive definite, which means that the ground state is stable. Well great, but wait, what would have happened had we chosen $-m^2 A^\mu A_\mu$ in the lagrangian to start with? We would have gotten

$$
\begin{aligned}
\mathcal{H}^T &= \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) - \frac{(\partial_i E^i - J^0)}{m^2}(\partial_i E^i) + \frac{1}{2}m^2\frac{(\partial_i E^i - J^0)^2}{m^4} + \frac{1}{2}m^2 A_i A^i \\
&\quad + J^0\frac{(\partial_i E^i - J^0)}{m^2} + J^i A_i \\
&= \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) - \frac{(\partial_i E^i - J^0)^2}{2m^2} - \frac{m^2}{2}A_i A_i + J^i A_i
\end{aligned}
$$

which is not positive definite, which in turn means that the ground state might *not* be stable. This is a massive difference (pun intended). As far is the east from the west, so far has one minus sign ruins one's day.

Jokes aside, this should cast some shadow on the prevailing mantra "you just need to include all invariant terms in the lagrangian" in building a model. True that we should include all invariant terms, but what are the signs? because we have seen that a plus or minus error will render a system stable or unstable. Symmetry alone is not enough.

The only thing left to do is to calculate the equations of motion for both massive and massless photons. Their hamiltonians are

$$\mathcal{H}^T_{m\neq 0} = \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) + \frac{(J^0 - \partial_i E^i)^2}{2m^2} + \frac{m^2}{2}A_i A_i + J^i A_i$$

$$\mathcal{H}^T_{m=0} = \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) - A_0(\partial_i E^i) + J^\mu A_\mu + \lambda_1 p^0_A + \lambda_2(\partial_i E^i - J^0)$$

Note that a massive photon has three degrees of freedom and thus $\mathcal{H}^T_{m\neq 0}$ has no unphysical d.o.f left while $\mathcal{H}^T_{m=0}$ still has two extra unphysical d.o.f's plus the unsolvable λ's. This is interesting because we have eliminated A_0 without the use of any gauge fixing. A question someone might ask is as to why we want to solve for A_0 instead of say J^0?, *i.e.* substituting J^0 for A_0 in the hamiltonian. The reason is that a massive photon has three degrees of freedom so one of them must be unphysical while the current J^0 is something external that we cannot control so it is more appropriate to eliminate A_0.

For completeness we will calculate all e.o.m's three way, just like before. We will first tackle the massive case.

Massive motion

To make sure we get the correct e.o.m's we'll first derive them from the lagrangian. The fastest way to do this is to morph the lagrangian into

$$\mathcal{L} = \frac{1}{2}A_\nu\left(g^{\nu\mu}(\partial^2 + m^2) - \partial^\mu\partial^\nu\right)A_\mu - J^\mu A_\mu$$

such that the e.o.m's are

$$\left(g^{\nu\mu}(\partial^2 + m^2) - \partial^\mu\partial^\nu\right)A_\nu = J^\mu$$

$$\partial_\nu F^{\nu\mu} + m^2 A^\mu = J^\mu$$

Following the same device as before

$$\partial_\nu F^{\nu\mu} + m^2 A^\mu = J^\mu$$

$$(\partial_i F^{i0} + m^2 A^0) + (\partial_0 F^{0k} + \partial_i F^{ik} + m^2 A^k) = J^0 + J^k$$

$$\rightarrow \partial_i F^{i0} + m^2 A^0 = J^0$$

$$\rightarrow \partial_0 F^{0k} + \partial_i F^{ik} + m^2 A^k = J^k$$

This means that

$$\partial_i F^{i0} + m^2 A^0 = J^0$$

$$m^2 A_0 = -\partial_i E^i + J^0$$

$$\rightarrow A_0 = \frac{1}{m^2}(J^0 - \partial_i E^i)$$

which is consistent with the result we get from setting $C_2 = 0$. This is conspicuously different from the massless case, the rest of the e.o.m's are

$$\partial_0 F^{0k} + \partial_i F^{ik} + m^2 A^k = J^k$$

$$\partial_0(-E^k) + \partial_i(-\varepsilon^{ikm} B_m) + m^2 A^k = J^k$$

$$\rightarrow (\nabla \times \vec{B}) + m^2 \vec{A} = \partial_0 \vec{E} + \vec{J}$$

so it's basically the massive version of Maxwell's equations. Massive photons are not as strange as you might think. For example, inside a superconductor photons become massive making them unable to penetrate through the superconductor.

The conserved current usually found in Maxwell's equations are now given by

$$\partial_\nu F^{\nu\mu} + m^2 A^\mu = J^\mu$$

$$\partial_\mu \partial_\nu F^{\nu\mu} + m^2 \partial_\mu A^\mu = \partial_\mu J^\mu$$

$$m^2 \partial_\mu A^\mu = \partial_\mu J^\mu$$

$$m^2 \partial_0 A^0 + m^2 \partial_i A^i = \langle \partial_0 J^0 \rangle^{\text{fm}} + \partial_i J^i$$

$$\partial_0 A^0 = \frac{1}{m^2}(\langle \partial_0 J^0 \rangle^{\text{fm}} + \partial_i J^i - m^2 \partial_i A^i)$$

Again I have used the notation $\langle \ \rangle^{\text{fm}}$ introduced earlier to accentuate the point that J's time evolution is generated by the fermionic part of the hamiltonian not the photonic part we are dealing with in this section (massive or otherwise). My point in introducing this funny notation is to underscore how easy it is to miss things that are coincidental, *i.e.* ∂_0 applied to A is different from ∂_0 applied to J inspite of the same notational manifestation.

The calculation for the e.o.m's from the hamiltonian is straightforward and is given in the following tables. To employ the Dirac bracket we need the matrix

$$\{C_m(\vec{x}), C_n(\vec{y})\} = \begin{pmatrix} 0 & -m^2 \delta_{xy} \\ m^2 \delta_{xy} & 0 \end{pmatrix} \qquad \{C_m(\vec{x}), C_n(\vec{y})\}^{-1} = \begin{pmatrix} 0 & \frac{1}{m^2}\delta_{xy} \\ -\frac{1}{m^2}\delta_{xy} & 0 \end{pmatrix}$$

$$\dot{A}_0$$

$\dot{A}_0 = \frac{\delta H}{\delta p_A^0} + \lambda_m \frac{\delta C_m}{\delta p_A^0}$ $= 0 + \lambda_1$ $= \frac{1}{m^2}(\langle \partial_0 J^0 \rangle^{\mathrm{fm}} + \partial_i J^i - m^2 \partial_i A^i)$	$\dot{A}_0 = \{A_0, H\} + \lambda_m \{A_0, C_m\}$ $= 0 + \lambda_1$ $= \frac{1}{m^2}(\langle \partial_0 J^0 \rangle^{\mathrm{fm}} + \partial_i J^i - m^2 \partial_i A^i)$

$$\{A_0, H^T\}^* = \{A_0, H\} - \{A_0, C_1\}\{C_1, C_2\}^{-1}\{C_2, H\} - \{A_0, C_2\}\{C_2, C_1\}^{-1}\{C_1, H\}$$
$$= 0 - \frac{1}{m^2}(m^2 \partial_i A^i - \partial_i J^i - \langle \partial_0 J^0 \rangle^{\mathrm{fm}}) - 0$$
$$\to m^2 \partial_0 A_0 = \langle \partial_0 J^0 \rangle^{\mathrm{fm}} + \partial_i J^i - m^2 \partial_i A^i$$

$$\dot{A}_i$$

$\dot{A}_i = \frac{\delta H}{\delta E^i} + \lambda_m \frac{\delta C_m}{\delta E^i}$ $= \left(E^i + \frac{\partial_i(J^0 - \partial_k E^k)}{m^2}\right) + 0$ $\partial_0 A_i = E^i + \partial_i A_0$ $\to E_i = \partial_i A_0 - \partial_0 A_i$	$\dot{A}_i = \{A_i, H\} + \lambda_m \{A_i, C_m\}$ $= \left(E^i + \frac{\partial_i(J^0 - \partial_k E^k)}{m^2}\right) + 0$ $\partial_0 A_i = E^i + \partial_i A_0$ $\to E_i = \partial_i A_0 - \partial_0 A_i$

$$\{A_i, H^T\}^* = \{A_i, H\} - \{A_i, C_1\}\{C_1, C_2\}^{-1}\{C_2, H\} - \{A_i, C_2\}\{C_2, C_1\}^{-1}\{C_1, H\}$$
$$= \left(E^i + \frac{\partial_i(J^0 - \partial_k E^k)}{m^2}\right) - 0 - 0$$
$$\partial_0 A_i = E^i + \partial_i A_0$$
$$\to E_i = \partial_i A_0 - \partial_0 A_i$$

$\dot{A}_i$ gives nothing but the definition of $E_i = F_{i0}$. The last term in line one of $\{A_i, H^T\}^*$ is zero because $\{C_1, H\} = 0$. We have also used integration by parts on $\partial_i E^i$ when we calculated the Poisson bracket $\{A_i, (J^0 - \partial_i E^i)^2/2m^2\}$.

$$\dot{p}_A^0$$

$\dot{p}_A^0 = -\frac{\delta H}{\delta A_0} - \lambda_m \frac{\delta C_m}{\delta A_0}$ $= -0 - \lambda_2(m^2)$ $= -0 - \frac{C_2}{m^2}(m^2)$ $= 0$	$\dot{p}_A^0 = \{p_A^0, H\} + \lambda_m \{p_A^0, C_m\}$ $= -0 - \lambda_2(m^2)$ $= -0 - \frac{C_2}{m^2}(m^2)$ $= 0$

$$\{p_A^0, H^T\}^* = \{p_A^0, H\} - \{p_A^0, C_1\}\{C_1, C_2\}\{C_2, H\} - \{p_A^0, C_2\}\{C_2, C_1\}\{C_1, H\}$$
$$= 0 - 0 - (-m^2)\left(\frac{1}{m^2}\right)(0)$$
$$\partial_0 p_A^0 = 0$$

Here we can see a difference from the fermionic case. Even though we have second class constraints there as well we didn't see the constraint popping up in the e.o.m's. The result

for $\dot{p}^0_A$ the traditional way is $\dot{p}^0_A = -C_2$ while using the Dirac bracket $\dot{p}^0_A = 0$, because we set all second class constraints to zero before calculating the Dirac brackets.

$\dot{E}^i$	
$\begin{aligned} \dot{E}^i &= -\frac{\delta H}{\delta A_i} - \lambda_m \frac{C_m}{\delta A_i} \\ &= -(-\partial_k F^{ki} + m^2 A_i + J^i) + 0 \\ \partial_0 E^i &= (\nabla \times \vec{B})^i + m^2 A^i - J^i \end{aligned}$	$\begin{aligned} \dot{E}^i &= \{E^i, H\} + \lambda_m \{E^i, C_m\} \\ &= -(-\partial_k F^{ki} + m^2 A_i + J^i) + 0 \\ \partial_0 E^i &= (\nabla \times \vec{B})^i + m^2 A^i - J^i \end{aligned}$
$\begin{aligned} \{E^i, H^T\}^* &= \{E^i, H\} - \{E^i, C_1\}\{C_1, C_2\}\{C_2, H\} - \{E^i, C_2\}\{C_2, C_1\}\{C_1, H\} \\ \partial_0 E^i &= -(-\partial_k F^{ki} + m^2 A_i + J^i) - 0 - 0 \\ \partial_0 E^i &= (\nabla \times \vec{B})^i + m^2 A^i - J^i \end{aligned}$	

From the above tables we see that half of the e.o.m's were ineffectual, they tell us something we already knew which affirms our earlier observation that Maxwell's equations are equations about the conjugate momenta primarily.

Massless motion

Things are more interesting for the massless case since we still have more d.o.f's than necessary. Let's look at A_0

$$\partial_0 A_0 = \frac{\delta H^T_{m=0}}{\delta p^0_A}$$
$$= \lambda_1$$
$$\to A_0 = \int dt \, \lambda_1$$

which means that A_0 is completely random since λ_1 is random. Some people lump this together with λ_2 since it also multiplies $\partial_i E^i$. Our hamiltonian is then

$$\mathcal{H}^T_{m=0} = \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) + J^\mu A_\mu + \lambda_1 p^0_A + \lambda_2 (K \partial_i E^i - J^0)$$

where $K \equiv 1 - A_0/\lambda_2$ is an arbitrary function.

Note that we can *not* use Dirac bracket here since $\{C_m, C_n\} = 0$ and is thus non

50

invertible. For the next e.o.m's, $\dot{A}_i$, we need to be careful, let's see why

$$\dot{A}_i(\vec{x}) = \int_y \frac{\delta\mathcal{H}(\vec{y})}{\delta E^i(\vec{x})} + \int_y \lambda_m \frac{\delta C_m(\vec{y})}{\delta E^i(\vec{x})}$$

$$= \int_y \delta_{xy} E^i - (\lambda_2 K)(\partial_i \delta_{xy})$$

$$= \int_y \delta_{xy} E^i + \partial_i(\lambda_2 - A_0)\delta_{xy}$$

$$= \int_y \delta_{xy} E^i - \partial_i A_0 \delta_{xy}, \qquad \lambda_2 \to 0$$

$$\to E_i \overset{?}{=} -\partial_i A_0 - \partial_0 A_i$$

which is wrong! Note that we have set $\lambda_2 = 0$ to simplify things, we can do this because $\lambda_2(t)$ is arbitrary, we also used $\lambda_2 K = \lambda_2 - A_0$. Before discussing what was wrong, the above example shows how important it is to keep track of what is what. Had we lumped up $\lambda_2 - A_0 \to \tilde{\lambda}_2$ (some random coefficient), we will just get $E_i = -\partial_0 A_i + \partial_i \tilde{\lambda}_2$ and since $\tilde{\lambda}_2$ and A_0 are both random we might as well set $\tilde{\lambda}_2 = A_0$. But this is sloppy, two random entities might not be equal.

What is wrong with the above calculation is that we were not careful in calculating $\delta C_m/\delta E^i$. If we spelled out all variables explicitly we would get

$$\dot{A}_i(\vec{x}) = \int_y \frac{\delta\mathcal{H}(\vec{y})}{\delta E^i(\vec{x})} + \int_y \lambda_m \frac{\delta C_m(\vec{y})}{\delta E^i(\vec{x})}$$

$$= \int_y \delta_{xy} E^i(\vec{y}) - (\lambda_2 K)(\partial_{x^i} \delta_{xy})$$

$$= \int_y \delta_{xy} E^i(\vec{y}) + \partial_{x^i}(\lambda_2 - A_0(\vec{y}))\delta_{xy}$$

$$= \int_y \delta_{xy} E^i(\vec{y}) - (\partial_{x^i} A_0(\vec{y}))\delta_{xy}$$

but $A_0(\vec{y})$ is *not* a function of $\vec{x}$! We need to use the identity we got earlier $\partial_{x^i}\delta_{xy} = -\partial_{y^i}\delta_{xy}$

such that

$$= \int_y \delta_{xy} E^i(\vec{y}) - (\lambda_2 K)(\partial_{x^i} \delta_{xy})$$

$$= \int_y \delta_{xy} E^i(\vec{y}) + (\lambda_2 K)(\partial_{y^i} \delta_{xy})$$

$$= \int_y \delta_{xy} E^i(\vec{y}) - \partial_{y^i}(\lambda_2 - A_0(\vec{y}))\delta_{xy}$$

$$= \int_y \delta_{xy} E^i(\vec{y}) + (\partial_{y^i} A_0(\vec{y}))\delta_{xy}$$

$$\rightarrow E_i = \partial_i A_0 - \partial_0 A_i$$

which is correct. This is one of the pitfalls of working with tensors, the numerous amount of indices can often obscure what the indices represent, prudence is imperative. As always $\dot{A}_i = \{A_i, H\} + \lambda_m \{A_i, C_m\}$ gives the same result.

If we reintroduce λ_2 back into the equation we get

$$E_i = \partial_i A_0 - \partial_0 A_i - \partial_i \lambda_2$$

but this is not quite right because E_i should not be affected by gauge transformation

$$\tilde{A}_\mu = A_\mu + \partial_\mu \Lambda$$

$$E_i = \partial_i \tilde{A}_0 - \partial_0 \tilde{A}_i$$

$$= \partial_i (A_0 + \partial_0 \Lambda) - \partial_0 (A_i + \partial_i \Lambda)$$

$$= \partial_i A_0 - \partial_0 A_i + (\partial_i \partial_0 - \partial_0 \partial_i)\Lambda, \qquad \partial_0 \partial_i = \partial_i \partial_0$$

$$E_i = \partial_i A_0 - \partial_0 A_i$$

which is always true no matter what Λ is because even if $\partial_i \Lambda = 0$, $\partial_0 \Lambda$ might not and since we can swap the order of differentiation the above consistently holds. The right way of seeing this is then

$$E_i = \partial_i(A_0 - \lambda_2) - \partial_0 A_i$$

$$= \partial_i \tilde{A}_0 - \partial_0 \tilde{A}_i$$

where λ_2 is a gauge transformation such that

$$\tilde{A}_\mu = A_\mu + \partial_\mu \Lambda$$

$$\partial_0 \Lambda = -\lambda_2$$

$$\partial_i \Lambda = 0$$

so this is *not* any gauge transformation! Some people prefer expressing it the following way

$$\partial_0 A_i = -E_i - \partial_i(\lambda_2 - A_0)$$
$$A_i = -\int dt \ (E_i + \partial_i(\lambda_2 - A_0))$$

which makes it seem like A_i is arbitrary but this is *not* quite right as shown above, it is A_0 that is actually arbitrary and since E_i is fixed (by Gauss's, Ampere's, and Faraday's laws) it is λ_2 that has to adjust *not* A_i. The next e.o.m is

$\dot{p}_A^0$	
$\dot{p}_A^0 = -\delta H/\delta A_0 - \lambda_m(\delta C_m/\delta A_0)$	$\dot{p}_A^0 = \{p_A^0, H\} + \lambda_m\{p_A^0, C_m\}$
$\qquad = -(J^0 - \partial_i E^i) - 0$	$\qquad = -(J^0 - \partial_i E^i) - 0$
$\qquad = 0$	$\qquad = 0$

Now this is somewhat interesting. $\delta C_m/\delta A_0 = -\{p_A^0, C_m\} = 0$ but we still get $J^0 - \partial_i E^i$ which is nothing but C_2 from $\delta C_m/\delta A_0 = -\{p_A^0, H\}$ and since $C_2 = 0 \to \dot{p}_A^0 = 0$. The last e.o.m is

$\dot{E}^i$	
$\dot{E}^i = -\frac{\delta H}{\delta A_i} - \lambda_m \frac{\delta C_m}{\delta A_i}$	$\dot{E}^i = \{E^i, H\} + \lambda_m\{E^i, C_m\}$
$\quad = -(-\partial_k F^{ki} + J^i) + 0$	$\quad = -(-\partial_k F^{ki} + J^i) + 0$
$\partial_0 E^i = (\nabla \times \vec{B})^i - J^i$	$\partial_0 E^i = (\nabla \times \vec{B})^i - J^i$

Hence the Dirac-Bergmann algorithm can be summarized as

- Add all primary constraints (be it second or first *class*) to $H \to H' = H + \lambda_m C_m$

- Calculate the naive Poisson bracket of the primary constraints with H' as a consistency check, this check is to generate secondary constraints and also to solve for the lagrange multipliers λ_m (if at all possible)

$$\{C_n, H_{\text{ext}}\} = \{C_n, H\} + \lambda_m\{C_n, C_m\} = 0$$
$$0 = \{C_n, H\} + \lambda_m C_{nm}$$
$$\to \lambda_m = -C_{nm}^{-1}\{C_n, H\}$$

which can only be solved for second class constraints, *i.e.* C_{mn}^{-1} exists, $C_{nm} \neq 0$. For first class constraints their multipliers remain as a arbitrary factor in the equations of motion.

- Find first class contributions for λ's if available, they are obtained from the homogenous portion of the set of linear equations pertaining the constraints $V_{am}\{C_n, C_m\} = 0$.

- Substitute solution for λ's (including the first class contributions) back into $H' \rightarrow H^T = H + \lambda_m(p, q)C_m + v_a V_{am}C_m$ where v_a is a random quantity, *i.e.* we can always take a linear combination of homogenous solutions to get another homogeneous solution.

- Calculate equations of motion using the Dirac bracket if all constraints are second class or using the usual Poisson bracket $\dot{f}(p, q) = \{f, H\} + (\lambda_m + v_a V_{am})\{f, C_m\}$.

V. POSTSCRIPT

This is a lightning speed rundown on first and second class constraints and their relationship to pre/symplectic mechanics. The usual Euler-Lagrange e.o.m.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0$$

$$\left(\frac{\partial L}{\partial \dot{q}^i \partial q^j}\dot{q}^j + \frac{\partial L}{\partial \dot{q}^i \partial \dot{q}^j}\ddot{q}^j\right) - \frac{\partial L}{\partial q^i} = 0$$

$$\ddot{q}^j\left(\frac{\partial L}{\partial \dot{q}^i \partial \dot{q}^j}\right) = -\frac{\partial L}{\partial q^i \partial \dot{q}^j}\dot{q}^j + \frac{\partial L}{\partial q^i}$$

We define the matrix

$$W_{ij} \equiv \left(\frac{\partial L}{\partial \dot{q}^i \partial \dot{q}^j}\right)$$

$$\rightarrow \ddot{q}^j = W_{ij}^{-1}\left(-\frac{\partial L}{\partial q^i \partial \dot{q}^j}\dot{q}^j + \frac{\partial L}{\partial q^i}\right)$$

If W_{ij} doesn't have an inverse then p_j can not be inverted to get $\dot{q}^j$ as a function of p_j, why is that? because

$$W_{ij} \equiv \left(\frac{\partial L}{\partial \dot{q}^i \partial \dot{q}^j} \right) = \frac{\partial \left(\frac{\partial L}{\partial \dot{q}^i} \right)}{\partial \dot{q}^j}, \quad \frac{\partial L}{\partial \dot{q}^i} = p_i$$

$$= \frac{\partial p_i}{\partial \dot{q}^j}$$

and by chain rule

$$W_{ij} = \frac{\partial p_i}{\partial \dot{q}^j} \rightarrow W_{ij}^{-1} = \frac{\partial \dot{q}^j}{\partial p_i}$$

$$\text{as} \quad \frac{\partial p_i}{\partial \dot{q}^j} \frac{\partial \dot{q}^j}{\partial p_k} = \frac{\partial p_i}{\partial p_k} = \delta_k^i$$

Thus W_{ij} not having an inverse means that $\partial \dot{q}^j / \partial p_k$ is tenuous and that in turn means that $\dot{q}^j$ cannot be defined as a function of p_k. As an example consider an action and its conjugate momenta

$$S = \int dt \; \theta_\mu(z)\dot{z}^\mu - H(z)$$

$$p_\alpha = \frac{\partial L}{\partial \dot{z}^\alpha}$$

$$= \theta_\alpha(z)$$

We can*not* invert the momenta to get $\dot{z}^\alpha = f(p_\alpha)$. The eom is

$$\delta S = \int dt \; \left(\frac{\partial \theta_\mu(z)}{\partial z^\nu} \delta z^\nu \right) \dot{z}^\mu + \theta_\mu(z)\delta\dot{z}^\mu - \frac{\partial H(z)}{\partial z^\nu} \delta z^\nu$$

$$= \int dt \; \left(\partial_\nu \theta_\mu(z)\dot{z}^\mu - \dot{\theta}_\mu(z)\delta_\nu^\mu - \partial_\nu H(z) \right) \delta z^\nu$$

where we have used integration by parts to move the time derivative on the second term to $\theta(z)$. We now have to use chain rule again to get

$$\dot{\theta}_\mu(z)\delta_\nu^\mu = \dot{\theta}_\nu(z) = \frac{\partial \theta(z)}{\partial z^\mu} \frac{\partial z^\mu}{\partial t}$$

$$\rightarrow \dot{\theta}_\mu(z)\delta_\nu^\mu = \partial_\mu \theta(z)\dot{z}^\mu$$

Thus the eom is

$$\delta S = \int dt \ \left(\partial_\nu \theta_\mu \dot{z}^\mu - \dot{\theta}_\mu \delta^\mu_\nu - \partial_\nu H \right) \delta z^\nu$$

$$= \int dt \ \left(\partial_\nu \theta_\mu \dot{z}^\mu - \partial_\mu \theta_\nu \dot{z}^\mu - \partial_\nu H \right) \delta z^\nu$$

$$= \int dt \ \left(\left(\partial_\nu \theta_\mu - \partial_\mu \theta_\nu \right) \dot{z}^\mu - \partial_\nu H \right) \delta z^\nu$$

$$= \int dt \ \left(\omega_{\nu\mu} \dot{z}^\mu - \partial_\nu H \right) \delta z^\nu$$

with $\omega_{\nu\mu} \equiv \partial_\nu \theta_\mu - \partial_\mu \theta_\nu$. This e.o.m can be rewritten as

$$\dot{z}^\mu = \omega^{\mu\nu} \partial_\nu H(z) = \{ z^\mu, H(z) \}$$

Here we have defined a new way of calculating the Poisson bracket

$$\{ f(z), g(z) \} = \frac{\partial f(z)}{\partial z^\mu} \omega^{\mu\nu} \frac{\partial g(z)}{\partial z^\nu}$$

which means that $\{ z^\mu, z^\nu \} = \omega^{\mu\nu}$. To make contact with the usual hamiltonian mechanics

$$z^\mu = \begin{pmatrix} q^i \\ p_i \end{pmatrix} \qquad \omega^{\mu\nu} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

to see why this is consider only one q and one p, i.e. $z^1 = q$ and $z^2 = p$.

$$\{ f(p,q), g(p,q) \} = \frac{\partial f}{\partial q} \frac{\partial g}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial g}{\partial q}$$

$$= \frac{\partial f}{\partial z^1} \frac{\partial g}{\partial z^2} - \frac{\partial f}{\partial z^2} \frac{\partial g}{\partial z^1}$$

$$= \frac{\partial f}{\partial z^1} \omega^{12} \frac{\partial g}{\partial z^2} - \frac{\partial f}{\partial z^2} \omega^{21} \frac{\partial g}{\partial z^1}$$

$$= \frac{\partial f(z)}{\partial z^\mu} \omega^{\mu\nu} \frac{\partial g(z)}{\partial z^\nu}$$

which is also true for any number of d.o.f's, for example for two degrees of freedom q^1, q^2, p_1, p_2

$$z^\mu = \begin{pmatrix} q^1 \\ q^2 \\ p_1 \\ p_2 \end{pmatrix} \qquad \omega^{\mu\nu} = \begin{pmatrix} 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

and so on for more degrees of freedom. This is what is usually called symplectic manifold (symjplectic because $\omega_{\mu\nu}$ has an inverse otherwise it is called presymplectic). This is analogous to Rimannian manifold with the role of the metric replaced by $\omega_{\mu\nu}$ (the symplectic two-form) and the "length" of a vector is given by the Poisson bracket which acts as an inner product. In differential form notation $\omega = \mathrm{d}q \wedge \mathrm{d}p$.

What it means is that in phase space there are subspaces which are physical, *i.e.* where the constraints is zero. When this happens $\omega_{\mu\nu}$ will no longer be constant, just like in Riemannian manifold when the space is no longer flat the metric cannot be constant because only on this subspace the motion physical. This is the case for second class constraints, these constraints specify the subspace of phase space where things are physical and this is what Dirac bracket does. It makes sure you don't leave this physical subspace as you move along in time by making sure that the Dirac bracket with the (second class) constraints are always zero, *i.e.* the contact transformation generated by second class constraints do not generate movement away from the physical subspace.

If $\omega_{\mu\nu}$ has no inverse we have gauge symmetry because now $\dot{z}^\mu = \omega^{\mu\nu}\partial_\nu H$ is undefined and so we cannot generate physical motion for z^μ. This is coherent with the definition of gauge transformations, they do *not* generate physical transformations. This is first class constraint. The subspace of phase space defined by $C_m = 0$ do not indicate different physical configurations, movement in the subspace has no physical meaning while movement leaving the subspace is physical.

VI.   EPILOGUE

One thing that is never mentioned in standard textbooks is that there is some sort of gauge redundancy in our humble quantum mechanics. Physical states are represented by rays in Hilbert space. Movement along the ray doesn't produce any physical consequences. We have an operation that changes a state and yet not alter its physical meaning, this is gauge. However, no one talks about gauging quantum mechanics.

Nevertheless, we do treat things like we are dealing with gauges. We always choose a gauge no matter what, either by explicitly setting the phase to zero $e^{i0}|\psi\rangle = |\psi\rangle$ or

by enforcing the "gauge condition" that only $\|\psi\|^2 = \langle\psi|\psi\rangle$ is observable. The difference about this particular "gauge" is that the generator is trivial and it only applies to the *overall* phase.

This might be one aspect that differentiates quantum mechanics from its classical cousin, there is no redundancy in classical mechanics. A lot of people attribute the quirkiness of quantum mechanics to its utilization of complex numbers but I think this is quite mistaken. The main impetus is redundancy and it can only be achieved through a phase which in turn requires complex numbers.