PEETER JOOT

CLASSICAL MECHANICS

CLASSICAL MECHANICS

PEETER JOOT

Independent study and phy354 notes and problems

August 2023 – version V0.1.17-1

Peeter Joot: *Classical Mechanics,* Independent study and phy354 notes and problems, © August 2023

COPYRIGHT

Copyright ©2023 Peeter Joot All Rights Reserved

This book may be reproduced and distributed in whole or in part, without fee, subject to the following conditions:

- The copyright notice above and this permission notice must be preserved complete on all complete or partial copies.
- Any translation or derived work must be approved by the author in writing before distribution.
- If you distribute this work in part, instructions for obtaining the complete version of this document must be included, and a means for obtaining a complete version provided.
- Small portions may be reproduced as illustrations for reviews or quotes in other works without this permission notice if proper citation is given.

Exceptions to these rules may be granted for academic purposes: Write to the author and ask.

Disclaimer: I confess to violating somebody's copyright when I copied this copyright statement.

DOCUMENT VERSION

Version V0.1.17-1

Sources for this notes compilation can be found in the github repository https://github.com/peeterjoot/classicalmechanics

The last commit (Aug/24/2023), associated with this pdf was 097af4ffcae6bd7e4869a2d94ba4af829aa6126f

Should you wish to actively contribute typo fixes (or even more significant changes) to this book, you can do so by contacting me, or by forking your own copy of the associated git repositories and building the book pdf from source, and submitting a subsequent merge request.

```
#!/bin/bash

git clone git@github.com:peeterjoot/latex-notes-
    compilations.git peeterjoot
cd peeterjoot

submods="figures/classicalmechanics
    classicalmechanics mathematica latex"
for i in $submods ; do
   git submodule update --init $i
   (cd $i && git checkout master)
done

export PATH=`pwd`/latex/bin:$PATH

cd classicalmechanics
make
```

I reserve the right to impose dictatorial control over any editing and content decisions, and may not accept merge requests as-is, or at all. That said, I will probably not refuse reasonable suggestions or merge requests.

Dedicated to:
Aurora and Lance, my awesome kids, and
Sofia, who not only tolerates and encourages my studies, but is also
awesome enough to think that math is sexy.

PREFACE

This book contains a large quantity of solved problems and study notes for classical mechanics.

- A collection of miscellaneous notes and problems for my personal (independent) classical mechanics studies. A fair amount of those notes were originally in my collection of Geometric (Clifford) Algebra related material so may assume some knowledge of that subject.

- My notes for some of the PHY354 lectures I attended. That class was taught by Prof. Erich Poppitz. I audited some of the Wednesday lectures since the timing was convenient. I took occasional notes, did the first problem set, and a subset of problem set 2.

  These notes, when I took them, likely track along with the Professor's hand written notes very closely, since his lectures follow his notes very closely. The text for PHY354 is [16]. I'd done my independent study from [4], also a great little book.

- Some assigned problems from the PHY354 course, ungraded (not submitted since I did not actually take the course). I ended up only doing the first problem set and two problems from the second problem set.

- Miscellaneous worked problems from other sources.

Peeter Joot peeterjoot@pm.me

CONTENTS

LIST OF FIGURES

1 POTENTIAL AND KINETIC ENERGY.

1.1 POTENTIAL AND KINETIC ENERGY.

Attempting some Lagrangian calculation problems I found I got all the signs of my potential energy terms wrong. Here is a quick step back to basics to clarify for myself what the definition of potential energy is, and thus implicitly determine the correct signs.

Starting with kinetic energy, expressed in vector form:

$$K = \frac{1}{2} m \mathbf{r}' \cdot \mathbf{r}' = \frac{1}{2} \mathbf{p} \cdot \mathbf{r}', \tag{1.1}$$

one can calculate the rate of change of that energy:

$$\begin{aligned} \frac{dK}{dt} &= \frac{1}{2} \left(\mathbf{p}' \cdot \mathbf{r}' + \mathbf{p} \cdot \mathbf{r}'' \right) \\ &= \frac{1}{2} \left(\mathbf{p}' \cdot \mathbf{r}' + \mathbf{r}' \cdot \mathbf{p}' \right) \\ &= \mathbf{p}' \cdot \mathbf{r}'. \end{aligned} \tag{1.2}$$

Note that the mass has been assumed constant above.

Integrating this time rate of change of kinetic energy produces a force line integral:

$$\begin{aligned} K_2 - K_1 &= \int_{t1}^{t2} \frac{dK}{dt} dt \\ &= \int_{t1}^{t2} \mathbf{p}' \cdot \mathbf{r}' dt \\ &= \int_{t1}^{t2} \mathbf{p}' \cdot \frac{d\mathbf{r}'}{dt} dt \\ &= \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F} \cdot d\mathbf{r}. \end{aligned} \tag{1.3}$$

For the path integral to depend on only the end points or the corresponding end times requires a conservative force that can be expressed as a gradient. Let us say that $\mathbf{F} = \nabla f$, then integrating:

$$\begin{aligned} K_2 - K_1 &= \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F} \cdot d\mathbf{r} \\ &= \int_{\mathbf{r}_1}^{\mathbf{r}_2} \nabla f \cdot d\mathbf{r} \\ &= \text{limit}_{\epsilon \to 0} \int_{\mathbf{r}_1}^{\mathbf{r}_1 + \epsilon \hat{\mathbf{r}}} \left(\hat{\mathbf{r}} \frac{f(\mathbf{r} + \epsilon \hat{\mathbf{r}})}{\epsilon} \right) \cdot d\mathbf{r} \\ &= \text{handwaving} \\ &= f(\mathbf{r}_2) - f(\mathbf{r}_1). \end{aligned} \tag{1.4}$$

Assembling the quantities for times 1, and 2, we have

$$K_2 - f(\mathbf{r}_2) = K_1 - f(\mathbf{r}_1) = \text{constant}. \tag{1.5}$$

This constant is what we give the name Energy. The quantities $-f(\mathbf{r}_i)$ we label potential energy V_i, and finally write the total energy as the sum of the kinetic and potential energies for a particle at a point in time and space:

$$K_2 + V_2 = K_1 + V_1 = E, \tag{1.6}$$

$$\mathbf{F} = -\nabla V. \tag{1.7}$$

1.1.1 *Work with a specific example. Newtonian gravitational force.*

Take the gravitational force:

$$F = -\frac{GmM}{r^2} \hat{\mathbf{r}}. \tag{1.8}$$

The rate of change of kinetic energy with respect to such a force (FIXME: think though signs ... with or against?), is:

$$\begin{aligned} \frac{dK}{dt} &= \mathbf{p}' \cdot \mathbf{r}' \\ &= -\frac{GmM}{r^2} \hat{\mathbf{r}} \cdot \frac{d\mathbf{r}}{dt} \\ &= -\frac{GmM}{r^3} \mathbf{r} \cdot \frac{d\mathbf{r}}{dt}. \end{aligned} \tag{1.9}$$

The vector dot products above can be eliminated with the standard trick:

$$\begin{aligned}\frac{dr^2}{dt} &= \frac{\mathbf{r}\cdot\mathbf{r}}{dt} \\ &= 2\frac{d\mathbf{r}}{dt}\cdot\mathbf{r}.\end{aligned} \tag{1.10}$$

Thus,

$$\begin{aligned}\frac{dK}{dt} &= -\frac{GmM}{2r^3}\frac{dr^2}{dt} \\ &= -\frac{GmM}{r^2}\frac{dr}{dt} \\ &= \frac{d}{dt}\left(\frac{GmM}{r}\right).\end{aligned} \tag{1.11}$$

This can be integrated to find the kinetic energy difference associated with a change of position in a gravitational field:

$$\begin{aligned}K_2 - K_1 &= \int_{t_1}^{t_2}\frac{d}{dt}\left(\frac{GmM}{r}\right)dt \\ &= GmM\left(\frac{1}{r_2}-\frac{1}{r_1}\right).\end{aligned} \tag{1.12}$$

Rearranging

$$K_2 - \frac{GmM}{r_2} = K_1 - \frac{GmM}{r_1} = E. \tag{1.13}$$

Taking gradients of this negative term:

$$\begin{aligned}\nabla\left(-\frac{GmM}{r}\right) &= \hat{\mathbf{r}}\frac{\partial}{\partial r}\left(-\frac{GmM}{r}\right) \\ &= \hat{\mathbf{r}}\frac{GmM}{r^2},\end{aligned} \tag{1.14}$$

returns the negation of the original force, so if we write $V = -GmM/r$, it implies the force is:

$$\mathbf{F} = -\nabla V. \tag{1.15}$$

By this example we see how one arrives at the negative sign convention for the potential energy. Our Lagrangian in a gravitational field is thus:

$$L = \frac{1}{2}m\mathbf{v}^2 + \frac{GmM}{r}. \tag{1.16}$$

Now, we have seen strictly positive terms mgh in the Lagrangian in the Tong and Goldstein examples. We can account for this by Taylor expanding this potential in the vicinity of the surface R of the Earth:

$$\begin{aligned}\frac{GmM}{r} &= \frac{GmM}{R+h} \\ &= \frac{GmM}{R(1+h/R)} \\ &\approx \frac{GmM}{R}(1-h/R).\end{aligned} \tag{1.17}$$

The Lagrangian is thus:

$$L \approx \frac{1}{2}m\mathbf{v}^2 + \frac{GmM}{R} - \frac{GmM}{R^2}h. \tag{1.18}$$

but the constant term will not change the EOM, so can be dropped from the Lagrangian, and with $g = \frac{GM}{R^2}$ we have:

$$L' = \frac{1}{2}m\mathbf{v}^2 - gmh. \tag{1.19}$$

Here the potential term of the Lagrangian is negative, but in the Goldstein and Tong examples the reference point is up, and the height is measured down from that point. Put another way, if the total energy is

$$E = V_0. \tag{1.20}$$

when the mass is unmoving in the air, and then drops gaining Kinetic energy, an unchanged total energy means that potential energy must be counted as lost, in proportion to the distance fallen:

$$E = V_0 = K_1 + V_1 = \frac{1}{2}m\mathbf{v}^2 - mgh. \tag{1.21}$$

So, one can write

$$V = -mgh, \tag{1.22}$$

and

$$L' = \frac{1}{2}m\mathbf{v}^2 + gmh. \tag{1.23}$$

BUT. Here the height h is the distance fallen from the reference point, compared to eq. (1.19), where h was the distance measured up from the surface of the Earth (or other convenient local point where the gravitational field can be linearly approximated)!

Care must be taken here because it is all too easy to get the signs wrong blindly plugging into the equations without considering where they come from and how exactly they are defined.

2 CALCULUS OF VARIATIONS.

Exercise 2.1 Shortest curve variational problem. (*[4] 2.1*)

Prove that the shortest length curve between two points in space is a straight line.

Exercise 2.2 Geodesics on sphere. (*[4] 2.2*)

Prove that the geodesics (shortest length paths) on a spherical surface are great circles.

Exercise 2.3 Euler-Lagrange equations for second order systems. (*[4] 2.4*)

For $f = f(y, \dot{y}, \ddot{y}, x)$, find the equations for extreme values of

$$I = \int_a^b f dx.$$

2.1 SOLUTIONS.

Answer for Exercise 2.1

In a first attempt of this I used:

$$ds = \sqrt{1 + (dy/dx)^2 + (dz/dx)^2} dx. \tag{2.1}$$

Application of the Euler-Lagrange equations does show that one ends up with a linear relation between the y and z coordinates, but no mention of x. Rather than write that up, consider instead a parametrization of the coordinates:

$$\begin{aligned} x &= x_1(\lambda) \\ y &= x_2(\lambda) \\ z &= x_3(\lambda), \end{aligned} \tag{2.2}$$

in terms of this arbitrary parametrization we have a segment length of:

$$ds = \sqrt{\sum \left(\frac{dx_i}{d\lambda}\right)^2} d\lambda = f\left(x_i\right) d\lambda. \tag{2.3}$$

Application of the Euler-Lagrange equation to f we have:

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= 0 \\ &= \frac{d}{d\lambda} \frac{\partial}{\partial \dot{x}_i} \sqrt{\sum \dot{x}_j^2} \\ &= \frac{d}{d\lambda} \frac{\dot{x}_i}{\sqrt{\sum \dot{x}_j^2}}. \end{aligned} \tag{2.4}$$

Therefore each of these quotients can be equated to a constant:

$$\begin{aligned} \frac{\dot{x}_i}{\sqrt{\sum \dot{x}_j^2}} &= {c_i}^{-2} \\ {c_i}^2 \dot{x}_i^2 &= \sum \dot{x}_j^2 \\ ({c_i}^2 - 1)\dot{x}_i^2 &= \sum_{j \neq i} \dot{x}_j^2 \\ (1 - {c_i}^2)\dot{x}_i^2 + \sum_{j \neq i} \dot{x}_j^2 &= 0. \end{aligned} \tag{2.5}$$

This last form shows explicitly that not all of these squared derivative terms can be linearly independent. In particular, we have a zero determinant:

$$0 = \begin{vmatrix} 1 - c_1^2 & 1 & 1 & 1 & \dots \\ 1 & 1 - c_2^2 & 1 & 1 & \vdots \\ 1 & 1 & 1 - c_3^2 & 1 & \\ & & & \ddots & \\ & & & & 1 - {c_n}^2 \end{vmatrix}. \tag{2.6}$$

Now, expanding this for a couple specific cases is not too hard. For $n = 2$ we have:

$$\begin{aligned} 0 &= (1 - c_1^2)(1 - c_2^2) - 1 \\ c_1^2 + c_2^2 &= c_1^2 c_2^2 \\ c_1^2 &= \frac{c_2^2}{c_2^2 - 1} \\ c_2^2 - 1 &= \frac{c_2^2}{c_1^2}. \end{aligned} \tag{2.7}$$

This can be substituted back into one our c_2^2 equation:

$$\begin{aligned} ({c_2}^2 - 1)\dot{x}_2^2 &= \dot{x}_1^2 \\ \frac{c_2^2}{c_1^2}\dot{x}_2^2 &= \dot{x}_1^2 \\ \pm\frac{c_2}{c_1}\dot{x}_2 &= \dot{x}_1 \\ \pm\frac{c_2}{c_1}x_2 &= x_1 + \kappa. \end{aligned} \tag{2.8}$$

This is precisely the straight line that was desired, but we have setup for proving that consideration of all path variations from two points in $\mathbb{R}^N$ space has the shortest distance when that path is a straight line. Despite the general setup, I am going to chicken out and show this only for the $\mathbb{R}^3$ case. In that case our determinant expands to:

$$c_1^2 + c_2^2 + c_3^2 = c_1^2 c_2^2 c_3^2. \tag{2.9}$$

Since not all of the $\dot{x}_i^2$ can be linearly independent, one can be eliminated:

$$\begin{aligned} (1 - c_1^2)\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2 &= 0 \\ (1 - c_2^2)\dot{x}_2^2 + \dot{x}_3^2 + \dot{x}_1^2 &= 0 \\ (1 - c_3^2)\dot{x}_3^2 + \dot{x}_1^2 + \dot{x}_2^2 &= 0. \end{aligned} \tag{2.10}$$

Let us pick $\dot{x}_1^2$ to eliminate, and subst 2 into 3:

$$\begin{aligned} (1 - c_3^2)\dot{x}_3^2 + (-(1 - c_2^2)\dot{x}_2^2 - \dot{x}_3^2) + \dot{x}_2^2 = 0 \implies \\ -c_3^2\dot{x}_3^2 + c_2^2\dot{x}_2 = 0 \\ \pm c_3\dot{x}_3 = c_2\dot{x}_2. \end{aligned} \tag{2.11}$$

Since these equations are symmetric, we can do this for all, with the result:

$$\begin{aligned} \pm c_3\dot{x}_3 &= c_2\dot{x}_2 \\ \pm c_3\dot{x}_3 &= c_1\dot{x}_1 \\ \pm c_2\dot{x}_2 &= c_1\dot{x}_1. \end{aligned} \tag{2.12}$$

Since the c_i constants are arbitrary, then we can for example pick the negative sign for $\pm c_2$, and the positive for the rest, then add all of these, and scale by two:

$$c_3\dot{x}_3 - c_2\dot{x}_2 = c_1\dot{x}_1, \tag{2.13}$$

and integrating:

$$c_3x_3 - c_2x_2 = c_1x_1 + \kappa. \tag{2.14}$$

Again, we have the general equation of a line, subject to the desired constraints on the end points. In the end we did not need to evaluate the determinant after all, as done in the $\mathbb{R}^2$ case.

Answer for Exercise 2.2

As a variational problem, the first step is to formulate an element of length on the surface. If we write our vector in spherical coordinates (ϕ on the equator, and θ measuring from the north pole) we have:

$$\mathbf{r} = (x, y, z) = R(\sin\theta cos\phi, \sin\theta\sin\phi, \cos\theta). \tag{2.15}$$

A differential vector element on the surface is (set $R = 1$ without loss of generality) :

$$\begin{aligned} d\mathbf{r} &= \frac{d\mathbf{r}}{d\theta}\frac{d\theta}{d\lambda}d\lambda + \frac{d\mathbf{r}}{d\phi}\frac{d\phi}{d\lambda}d\lambda \\ &= (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta)\dot{\theta}d\lambda + (-\sin\theta\sin\phi, \sin\theta\cos\phi, 0)\dot{\phi}d\lambda \\ &= (\cos\theta\cos\phi\dot{\theta} - \sin\theta\sin\phi\dot{\phi}, \cos\theta\sin\phi\dot{\theta} + \sin\theta\cos\phi\dot{\phi}, -\sin\theta\dot{\theta})d\lambda. \end{aligned} \tag{2.16}$$

Calculation of the length ds of this vector yields:

$$ds = |d\mathbf{r}| = \sqrt{\dot{\theta}^2 + (\sin\theta)^2\dot{\phi}^2}d\lambda. \tag{2.17}$$

This completes the setup for the minimization problem, and we want to minimize:

$$s = \int \sqrt{\dot{\theta}^2 + (\dot{\phi}\sin\theta)^2} d\lambda, \tag{2.18}$$

and can therefore apply the Euler-Lagrange equations to the function

$$f(\theta, \phi, \dot{\theta}, \dot{\phi}, \lambda) = \sqrt{\dot{\theta}^2 + (\dot{\phi}\sin\theta)^2}. \tag{2.19}$$

The ϕ is cyclic, and we have:

$$\frac{\partial f}{\partial \phi} = 0 = \frac{d}{d\lambda}\frac{\dot{\phi}\sin^2\theta}{f}. \tag{2.20}$$

Thus we have:

$$\begin{aligned}
\dot{\phi}^2 \sin^4\theta &= K^2\left(\dot{\theta}^2 + (\dot{\phi}\sin\theta)^2\right) \\
\dot{\phi}^2 \sin^2\theta\left(\sin^2\theta - K^2\right) &= K^2\dot{\theta}^2 \\
\dot{\phi}^2 &= \frac{K^2\dot{\theta}^2}{\sin^2\theta\left(\sin^2\theta - K^2\right)} \\
\dot{\phi} &= \frac{K\dot{\theta}}{\sin\theta\sqrt{\sin^2\theta - K^2}}.
\end{aligned} \tag{2.21}$$

This is in a nicely separated form, but it is not obvious that this describes paths that are great circles. Let us have a look at the second equation.

$$\begin{aligned}
\frac{\partial f}{\partial \theta} &= \frac{d}{d\lambda}\frac{\partial f}{\partial \dot{\theta}} \\
\frac{\sin\theta\cos\theta\dot{\phi}^2}{f} &= \frac{d}{d\lambda}\frac{\dot{\theta}}{f} \\
&= \frac{\ddot{\theta}}{f} - \frac{1}{2}\frac{\left(\dot{\theta}^2 + (\dot{\phi}\sin\theta)^2\right)'}{f^3} \\
&= \frac{\ddot{\theta}}{f} - \frac{\dot{\theta}\ddot{\theta} + \dot{\phi}\sin\theta\left(\ddot{\phi}\sin\theta + \dot{\phi}\cos\theta\dot{\theta}\right)}{f^3}.
\end{aligned} \tag{2.22}$$

This implies

$$\begin{aligned}
&-\sin\theta\cos\theta\dot{\phi}^2\left(\dot{\theta}^2 + (\dot{\phi}\sin\theta)^2\right) \\
&= -\ddot{\theta}\left(\dot{\theta}^2 + (\dot{\phi}\sin\theta)^2\right) + \dot{\theta}\ddot{\theta} + \dot{\phi}\sin\theta\left(\ddot{\phi}\sin\theta + \dot{\phi}\cos\theta\dot{\theta}\right),
\end{aligned} \tag{2.23}$$

or,

$$\begin{aligned} 0 &= -\ddot{\theta}\dot{\theta}^2 - \ddot{\theta}\dot{\phi}^2 \sin^2\theta + \dot{\theta}\ddot{\theta} \\ &\quad + \dot{\phi}\ddot{\phi}\sin^2\theta + \dot{\phi}^2\dot{\theta}\sin\theta\cos\theta + \dot{\phi}^2\dot{\theta}^2\sin\theta\cos\theta + \dot{\phi}^4\sin^3\theta\cos\theta. \end{aligned} \tag{2.24}$$

What a mess! I do not feel inclined to try to reduce this at the moment. I will come back to this problem later. Perhaps there is a better parametrization? Did come back to this later, in [12], but still did not get the problem fully solved. Maybe the third time, some time later, will be the charm.

Answer for Exercise 2.3

Here we want y and $\dot{y}$ fixed at the end points. Following the first derivative derivation write the functions in terms of the desired extremum functions plus a set of arbitrary functions:

$$\begin{aligned} y(x,\alpha) &= y(x,0) + \alpha n(x) \\ \dot{y}(x,\alpha) &= \dot{y}(x,0) + \alpha m(x). \end{aligned} \tag{2.25}$$

Here we specify that these arbitrary variational functions vanish at the endpoints:

$$n(a) = n(b) = m(a) = m(b) = 0. \tag{2.26}$$

The functions $y(x,0)$, and $\dot{y}(x,0)$ are the functions we are looking for as solutions to the min/max problem. Calculating derivatives we have:

$$\frac{dI}{d\alpha} = \int \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha} + \frac{\partial f}{\partial \ddot{y}}\frac{\partial \ddot{y}}{\partial \alpha} \right) dx. \tag{2.27}$$

Assuming sufficient continuity look at the second term where we have:

$$\begin{aligned} \frac{\partial \dot{y}}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\frac{\partial y}{\partial x} \\ &= \frac{\partial}{\partial x}\frac{\partial y}{\partial \alpha} \\ &= \frac{\partial}{\partial x} n(x) \\ &= \frac{d}{dx} n(x) \\ &= \frac{d}{dx}\frac{\partial y}{\partial \alpha}. \end{aligned} \tag{2.28}$$

Similarly for the third term we have:

$$\frac{\partial \dot{y}}{\partial \alpha} = \frac{d}{dx}\frac{\partial \dot{y}}{\partial \alpha}, \tag{2.29}$$

$$uv' = (uv)' - u'v$$

$$\frac{dI}{d\alpha} = \int \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha}dx + \boxed{\frac{\partial f}{\partial \dot{y}}\frac{d}{dx}\frac{\partial y}{\partial \alpha}}dx + \frac{\partial f}{\partial \ddot{y}}\frac{d}{dx}\frac{\partial \dot{y}}{\partial \alpha}dx. \tag{2.30}$$

Now integrating by parts:

$$\frac{dI}{d\alpha} = \int \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha}dx + \int \frac{\partial f}{\partial \dot{y}}\frac{d}{dx}\frac{\partial y}{\partial \alpha}dx + \int \frac{\partial f}{\partial \ddot{y}}\frac{d}{dx}\frac{\partial \dot{y}}{\partial \alpha}dx$$

$$n(x)$$

$$\frac{dI}{d\alpha} = \int \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha}dx + \frac{\partial f}{\partial \dot{y}}\boxed{\frac{\partial y}{\partial \alpha}\bigg|_a^b} - \int \frac{\partial y}{\partial \alpha}\frac{d}{dx}\frac{\partial f}{\partial \dot{y}}dx \tag{2.31}$$

$$m(x)$$

$$+ \frac{\partial f}{\partial \ddot{y}}\boxed{\frac{\partial \dot{y}}{\partial \alpha}\bigg|_a^b} - \int \frac{\partial \dot{y}}{\partial \alpha}\frac{d}{dx}\frac{\partial f}{\partial \ddot{y}}dx.$$

Because $m(a) = m(b) = n(a) = n(b)$ the non-integral terms are all zero, leaving:

$$\frac{dI}{d\alpha} = \int \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha}dx - \int \frac{\partial y}{\partial \alpha}\frac{d}{dx}\frac{\partial f}{\partial \dot{y}}dx - \int \frac{\partial \dot{y}}{\partial \alpha}\frac{d}{dx}\frac{\partial f}{\partial \ddot{y}}dx. \tag{2.32}$$

Now consider just this last integral, which we can again integrate by parts:

$$u' \qquad v$$

$$\int \frac{\partial \dot{y}}{\partial \alpha}\frac{d}{dx}\frac{\partial f}{\partial \ddot{y}}dx = \int \boxed{\frac{d}{dx}\frac{\partial y}{\partial \alpha}}\boxed{\frac{d}{dx}\frac{\partial f}{\partial \ddot{y}}}dx$$

$$n(x) \tag{2.33}$$

$$= \boxed{\frac{\partial y}{\partial \alpha}}\frac{d}{dx}\frac{\partial f}{\partial \ddot{y}}\bigg|_a^b - \int \frac{\partial y}{\partial \alpha}\frac{d}{dx}\frac{d}{dx}\frac{\partial f}{\partial \ddot{y}}dx$$

$$= -\int \frac{\partial y}{\partial \alpha}\frac{d^2}{dx^2}\frac{\partial f}{\partial \ddot{y}}dx.$$

This gives:

$$\begin{aligned}
\frac{dI}{d\alpha} &= \int \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha}dx - \int \frac{\partial y}{\partial \alpha}\frac{d}{dx}\frac{\partial f}{\partial \dot{y}}dx + \int \frac{\partial y}{\partial \alpha}\frac{d^2}{dx^2}\frac{\partial f}{\partial \ddot{y}}dx \\
\frac{dI}{d\alpha} &= \int dx \frac{\partial y}{\partial \alpha}\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} + \frac{d^2}{dx^2}\frac{\partial f}{\partial \ddot{y}}\right) \\
&= \int dx n(x)\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} + \frac{d^2}{dx^2}\frac{\partial f}{\partial \ddot{y}}\right).
\end{aligned} \tag{2.34}$$

So, if we want this derivative to equal zero for any $n(x)$ we require the inner quantity to by zero:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} + \frac{d^2}{dx^2}\frac{\partial f}{\partial \ddot{y}} = 0. \tag{2.35}$$

Question. Goldstein writes this in total differential form instead of a derivative:

$$\begin{aligned}
dI &= \frac{dI}{d\alpha}d\alpha \\
&= \int dx \left(\frac{\partial y}{\partial \alpha}d\alpha\right)\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} + \frac{d^2}{dx^2}\frac{\partial f}{\partial \ddot{y}}\right).
\end{aligned} \tag{2.36}$$

and then calls this quantity $\frac{\partial y}{\partial \alpha}d\alpha = \delta y$, the variation of y. There must be a mathematical subtlety which motivates this but it is not clear to me what that is. Since the variational calculus texts go a different route, with norms on functional spaces and so forth, perhaps understanding that motivation is not worthwhile. In the end, the conclusion is the same, namely that the inner expression must equal zero for the extremum condition.

3 SPECIAL RELATIVITY.

Placeholder.

4 ACTION AND EULER-LANGRANGE EQUATIONS.

4.1 SCALAR FORM OF EULER-LAGRANGE EQUATIONS.

[17] presents a multivector Lagrangian treatment. To preparation for understanding that I have gone back and derived the scalar case myself. As in my recent field Lagrangian derivations Feynman's [3] simple action procedure will be used. Write

$$\begin{aligned} L &= L(q^i, \dot{q}^i, \lambda) \\ q^i &= \bar{q}^i + n^i \\ S &= \int_{\partial\lambda} L d\lambda. \end{aligned} \tag{4.1}$$

Here $\bar{q}^i$ are the desired optimal solutions, and the functions n^i are all zero at the end points of the integration range $\partial\lambda$.

A multivariable function $f(a^i) = f(a^1, a^2, \cdots, a^n)$ may be expanded, to first order, in Taylor series

$$f(a^i + x^i) \approx f(a^i) + \sum_i (a^i + x^i) \left.\frac{\partial f}{\partial x^i}\right|_{x^i = a^i}. \tag{4.2}$$

In this case the x^i take the values q^i, and $\dot{q}^i$, so the first order Lagrangian approximation requires summation over differential contributions for both sets of terms

$$L(q^i, \dot{q}^i, \lambda) \approx L(\bar{q}^i, \dot{\bar{q}}^i, \lambda) + \sum_i (\bar{q}^i + n^i) \left.\frac{\partial L}{\partial q^i}\right|_{q^i = \bar{q}^i} + \sum_i (\dot{\bar{q}}^i + \dot{n}^i) \left.\frac{\partial L}{\partial \dot{q}^i}\right|_{q^i = \bar{q}^i}. \tag{4.3}$$

Now form the action, and group the terms in fixed and variable sets

$$
\begin{aligned}
S &= \int L d\lambda \\
&\approx \int d\lambda \left(L(\bar{q}^i, \dot{\bar{q}}^i, \lambda) + \sum_i \bar{q}^i \left.\frac{\partial L}{\partial q^i}\right|_{q^i=\bar{q}^i} + \sum_i \dot{\bar{q}}^i \left.\frac{\partial L}{\partial \dot{q}^i}\right|_{q^i=\bar{q}^i} \right) \\
&+ \overbrace{\boxed{\sum_i \int d\lambda \left(n^i \left.\frac{\partial L}{\partial q^i}\right|_{q^i=\bar{q}^i} + \dot{n}^i \left.\frac{\partial L}{\partial \dot{q}^i}\right|_{q^i=\bar{q}^i} \right)}}^{\delta S}.
\end{aligned}
\tag{4.4}
$$

For the optimal solution we want $\delta S = 0$ for all possible paths n^i. Now do the integration by parts writing $u' = \dot{n}^i$, and $v = \partial L / \partial \dot{q}^i$

$$
\int u' v = uv - \int u v'. \tag{4.5}
$$

The action variation is then

$$
\delta S = + \sum_i \left. \left(n^i \frac{\partial L}{\partial \dot{q}^i} \right) \right|_{\partial \lambda} + \sum_i \int d\lambda n^i \left(\left.\frac{\partial L}{\partial q^i}\right|_{q^i=\bar{q}^i} - \frac{d}{d\lambda} \left.\frac{\partial L}{\partial \dot{q}^i}\right|_{q^i=\bar{q}^i} \right). \tag{4.6}
$$

The non-integral term is zero since by definition $n^i = 0$ on the boundary of the desired integration region, so for the total variation to equal zero for all possible paths n^i one must have

$$
\frac{\partial L}{\partial q^i} - \frac{d}{d\lambda} \frac{\partial L}{\partial \dot{q}^i} = 0. \tag{4.7}
$$

Evaluation of these derivatives at the optimal desired paths has been suppressed since these equations now define that path.

4.1.1 *Some comparison to the Goldstein approach.*

[4] calls the quantity eq. (4.7) the functional derivative

$$
\frac{\delta S}{\delta q^i} = \frac{\partial L}{\partial q^i} - \frac{d}{d\lambda} \frac{\partial L}{\partial \dot{q}^i}. \tag{4.8}
$$

(with higher order derivatives if the Lagrangian has dependencies on more than generalized position and velocity terms). Goldstein's approach is

also harder to follow than Feynman's (Goldstein introduces a parameter ϵ, writing

$$q^i = \bar{q}^i + \epsilon n^i. \tag{4.9}$$

He then takes derivatives under the integral sign for the end result.

While his approach is a bit harder to follow initially, that additional ϵ parametrization of the variation path also fits nicely with this linearization procedure. After the integration by parts and subsequent differentiation under integral sign nicely does the job of discarding all the "fixed" $\bar{q}^i$ contributions to the action leaving:

$$\frac{dS}{d\epsilon} = \int d\lambda \sum_i n^i \left.\frac{\delta S}{\delta q^i}\right|_{q^i=\bar{q}^i}. \tag{4.10}$$

Introducing this idea does firm things up, eliminating some hand waving. To obtain the extremal solution it does make sense to set the derivative of the action equal to zero, and introducing an additional scalar variational control in the paths from the optimal solution provides that something to take derivatives with respect to.

Goldstein also writes that this action derivative is then evaluated at $\epsilon = 0$. This really says the same thing as Feynman... toss all the higher order terms, since factors of epsilon will be left associated with of these. With my initial read of Goldstein this was not the least bit clear... it was really yet another example of the classic physics approach of solving something with a first order linear approximation.

4.2 PROBLEMS.

Exercise 4.1 Lorentz force Lagrangian. (*'12 phy356 ps1.1*)

1. For the non-covariant electrodynamic Lorentz force Lagrangian.

$$L = \frac{1}{2}m\mathbf{v}^2 + q\mathbf{v}\cdot\mathbf{A} - q\phi, \tag{4.11}$$

derive the Lorentz force equation

$$\begin{aligned}
\mathbf{F} &= q(\mathbf{E} + \mathbf{v}\times\mathbf{B}) \\
\mathbf{E} &= -\boldsymbol{\nabla}\phi - \frac{\partial \mathbf{A}}{\partial t} \\
\mathbf{B} &= \boldsymbol{\nabla}\times\mathbf{A}.
\end{aligned} \tag{4.12}$$

2. With a gauge transformation of the form:

$$\begin{aligned} \phi &\to \phi + \frac{\partial \chi}{\partial t} \\ \mathbf{A} &\to \mathbf{A} - \boldsymbol{\nabla}\chi, \end{aligned} \tag{4.13}$$

show that the Lagrangian is invariant.

Exercise 4.2 Find trajectory using action. (*'12 phy356 ps1.2.*)

For a ball thrown upward, guess a solution for the height y of the form $y(t) = a_2 t^2 + a_1 t + a_0$. Assuming that $y(0) = y(T) = 0$, this quickly becomes $y(t) = a_2(t^2 - Tt)$. Calculate the action (to do that, you need to first write the Lagrangian, of course) between $t = 0$ and $t = T$, and show that it is minimized when $a_2 = -g/2$.

Exercise 4.3 Change of coordinates. (*'12 phy356 ps1.3*)

Consider a Lagrangian $L(q, \dot{q}) \equiv L(q_1, \cdots, q_N, \dot{q}_1, \cdots \dot{q}_N)$. Now change the coordinates to some new ones, e.g. let $q_i = q_i(x_1, \cdots, x_N), i = 1 \cdots N$, or in short $q_i = q_i(x)$. This defines a new Lagrangian:

$$\tilde{L}(x, \dot{x}) = L(q_1(x), \cdots q_N(x), \frac{d}{dt} q_1(x), \cdots \frac{d}{dt} q_N(x)), \tag{4.32}$$

which is now a function of x_i and $\dot{x}_i$. Show that the Euler-Lagrange equations for $L(q, \dot{q})$:

$$\frac{\partial L(q, \dot{q})}{\partial q_i} = \frac{d}{dt} \frac{\partial L(q, \dot{q})}{\partial \dot{q}_i}, \tag{4.33}$$

imply that the Euler-Lagrange equations for $\tilde{L}(x, \dot{x})$ hold (provided the change of variables $q \to x$ is nonsingular):

$$\frac{\partial \tilde{L}(x, \dot{x})}{\partial x_i} = \frac{d}{dt} \frac{\partial \tilde{L}(x, \dot{x})}{\partial \dot{x}_i}. \tag{4.34}$$

The moral is that the action formalism is very convenient: one can write the Lagrangian in any set of coordinates; the Euler-Lagrange equations for the new coordinates can then be obtained by using the Lagrangian expressed in these coordinates.

Hint: Solving this problem only requires repeated use of the chain rule.

Exercise 4.4 Symmetries and conservation (E.M.) (*'12 phy356 ps2.1*)

Let us continue studying the Lagrangian of Problem 1 of Homework 1, namely, its symmetries, and the relevant conserved quantities. To this end, we will consider various cases of external scalar and vector potentials.

1. Consider first the case of time-independent $\mathbf{A}$ and ϕ. Find the expression for the conserved energy, $\mathcal{E}$, of the particle.
2. For external $\mathbf{A}$ and ϕ dependent on time, find $d\mathcal{E}/dt$.
3. Let now $\mathbf{A}$ and ϕ be spatially homogeneous, i.e. $\mathbf{x}$-independent. Find the conserved momentum. Is it equal to the usual $m\mathbf{v}$?
4. Consider motion in the field of an electrostatic source (creating an external static $\phi(\mathbf{x})$). Find the angular momentum of the particle. Is it conserved for all $\phi(\mathbf{x})$?

Exercise 4.5 Angular momentum, non-rectangular. (*'12 phy356 ps2.5*)

1. Find $M_x, M_y, M_z, \mathbf{M}^2$ in spherical coordinates (r, θ, ϕ).
2. Find $M_x, M_y, M_z, \mathbf{M}^2$ in cylindrical coordinates (r, ϕ, z).

Exercise 4.6 Angular momentum, three particle system. (*[4] 1.8*)

A system is composed of three particles of equal mass m. Between any two of them there are forces derivable from a potential

$$V = -ge^{-\mu r}, \tag{4.105}$$

where r is the distance between the two particles. In addition, two of the particles each exert a force on the third which can be obtained from a generalized potential of the form

$$U = -f\mathbf{v}\cdot\mathbf{r}, \tag{4.106}$$

$\mathbf{v}$ being the relative velocity of the interacting particles and f a constant. Set up the Lagrangian for the system, using as coordinates the radius vector $\mathbf{R}$ of the center of mass and the two vectors

$$\begin{aligned} \boldsymbol{\rho}_1 &= \mathbf{r}_1 - \mathbf{r}_3 \\ \boldsymbol{\rho}_2 &= \mathbf{r}_2 - \mathbf{r}_3. \end{aligned} \tag{4.107}$$

Is the total angular momentum of the system conserved?

Exercise 4.7 Kinetic energy for barbell shaped object. (*[4] 1.6*)

Two points of mass m are joined by a rigid weightless rod of length l, the center of which is constrained to move on a circle of radius a. Set up the kinetic energy in generalized coordinates.

Exercise 4.8 Purely kinetic system.) (*[24] p1*)

Derive the Euler-Lagrange equations for

$$L = \frac{1}{2}\sum g_{ab}(q_c)\dot{q}^a\dot{q}^b. \tag{4.145}$$

Exercise 4.9 Alternate Lagrangian. (*[24] p2.*)

$$L = \frac{1}{12}m^2\dot{x}^4 + m\dot{x}^2V - V^2. \tag{4.155}$$

Exercise 4.10 Relativistic EOM. (*[24] p3.*)

Derive the relativistic equations of motion for a point particle in a position dependent potential:

$$L = -mc^2\sqrt{1 - \mathbf{v}^2/c^2} - V(\mathbf{r}). \tag{4.162}$$

Exercise 4.11 Double pendulum. (*[24] p4.*)

Derive the Lagrangian for a double pendulum.

Exercise 4.12 Lorentz force Lagrangian. (*[24] p6.*)

Various non-orthogonal coordinate treatments of the Lorentz force Lagrangian

$$L = \frac{1}{2}m\mathbf{v}^2 - q\phi + q\mathbf{v}/c\cdot\mathbf{A},$$

Exercise 4.13

In [7], the non-covariant Lagrangian for the Lorentz equation is given as

$$L = -mc^2\sqrt{1 - \mathbf{u}^2/c^2} + \frac{e}{c}\mathbf{u}\cdot\mathbf{A} - e\phi. \tag{4.224}$$

Evaluate this to show that this produces the Lorentz force law.

Exercise 4.14 Dipole Moment induced by a constant electric field.

In [6] it is stated that the force per unit angle on a dipole system as illustrated in fig. 4.1 is

$$F_\theta = -p\mathcal{E}\sin\theta, \tag{4.233}$$

where $\mathbf{p} = q\mathbf{r}$. The text was also referring to torques, and it wasn't clear to me if the result was the torque or the force. Derive the result to resolve any doubt (in retrospect dimensional analysis would also have worked).

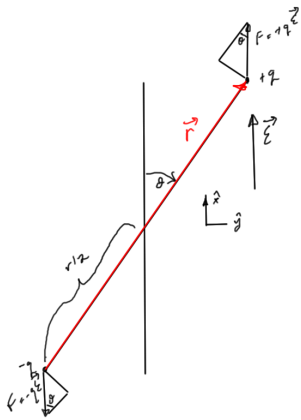


Figure 4.1: Dipole moment coordinates.

4.3 SOLUTIONS.

Answer for Exercise 4.1

Solution Part 1. *Evaluate the Euler-Lagrange equations.* In coordinates, employing summation convention, this Lagrangian is

$$L = \frac{1}{2}m\dot{x}_j\dot{x}_j + q\dot{x}_jA_j - q\phi. \tag{4.14}$$

Taking derivatives

$$\frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + qA_i, \tag{4.15}$$

$$\begin{aligned}\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} &= m\ddot{x}_i + q\frac{\partial A_i}{\partial t} + q\frac{\partial A_i}{\partial x_j}\frac{dx_j}{dt} \\ &= m\ddot{x}_i + q\frac{\partial A_i}{\partial t} + q\frac{\partial A_i}{\partial x_j}\dot{x}_j.\end{aligned} \tag{4.16}$$

This must equal

$$\frac{\partial L}{\partial x_i} = q\dot{x}_j\frac{\partial A_j}{\partial x_i} - q\frac{\partial \phi}{\partial x_i}, \tag{4.17}$$

so we have

$$\begin{aligned}m\ddot{x}_i &= -q\frac{\partial A_i}{\partial t} - q\frac{\partial A_i}{\partial x_j}\dot{x}_j + q\dot{x}_j\frac{\partial A_j}{\partial x_i} - q\frac{\partial \phi}{\partial x_i} \\ &= -q\left(\frac{\partial A_i}{\partial t} - \frac{\partial \phi}{\partial x_i}\right) + qv_j\left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\right).\end{aligned} \tag{4.18}$$

The first term is just E_i. If we expand out $(\mathbf{v}\times\mathbf{B})_i$ we see that matches

$$\begin{aligned}(\mathbf{v}\times\mathbf{B})_i &= v_a B_b \epsilon_{abi} \\ &= v_a \partial_r A_s \epsilon_{rsb}\epsilon_{abi} \\ &= v_a \partial_r A_s \delta_{rs}^{[ia]} \\ &= v_a(\partial_i A_a - \partial_a A_i).\end{aligned} \tag{4.19}$$

A $a \to j$ substitution, and comparison of this with the Euler-Lagrange result above completes the exercise.

Solution Part 2. *Gauge invariance.* We really only have to show that

$$\mathbf{v}\cdot\mathbf{A} - \phi, \tag{4.20}$$

is invariant. Making the transformation we have

$$\begin{aligned}\mathbf{v}\cdot\mathbf{A} - \phi &\to v_j\left(A_j - \partial_j\chi\right) - \left(\phi + \frac{\partial \chi}{\partial t}\right) \\ &= v_j A_j - \phi - v_j\partial_j\chi - \frac{\partial \chi}{\partial t} \\ &= \mathbf{v}\cdot\mathbf{A} - \phi - \left(\frac{dx_j}{dt}\frac{\partial \chi}{\partial x_j} + \frac{\partial \chi}{\partial t}\right) \\ &= \mathbf{v}\cdot\mathbf{A} - \phi - \frac{d\chi(\mathbf{x},t)}{dt}.\end{aligned} \tag{4.21}$$

We see then that the Lagrangian transforms as

$$L \to L + \frac{d}{dt}\left(-q\chi\right), \tag{4.22}$$

and differs only by a total derivative. With the lemma from the lecture, we see that this gauge transformation does not have any effect on the end result of applying the Euler-Lagrange equations.

Answer for Exercise 4.2

We are told to guess at a solution

$$y = a_2 t^2 + a_1 t + a_0, \tag{4.23}$$

for the height of a particle thrown up into the air. With initial condition $y(0) = 0$ we have

$$a_0 = 0, \tag{4.24}$$

and with a final condition of $y(T) = 0$ we also have

$$\begin{aligned} 0 &= a_2 T^2 + a_1 T \\ &= T(a_2 T + a_1), \end{aligned} \tag{4.25}$$

so have

$$\begin{aligned} y(t) &= a_2 t^2 - a_2 T t = a_2\left(t^2 - Tt\right) \\ \dot{y}(t) &= a_2(2t - T). \end{aligned} \tag{4.26}$$

So our Lagrangian is

$$L = \frac{1}{2} m a_2^2 \left(2t - T\right)^2 - m g a_2 \left(t^2 - Tt\right), \tag{4.27}$$

and our action is

$$S = \int_0^T dt \left(\frac{1}{2} m a_2^2 \left(2t - T\right)^2 - m g a_2 \left(t^2 - Tt\right)\right). \tag{4.28}$$

To minimize this action with respect to a_2 we take the derivative

$$\frac{\partial S}{\partial a_2} = \int_0^T \left(m a_2 \left(2t - T\right)^2 - m g \left(t^2 - Tt\right)\right). \tag{4.29}$$

Integrating we have

$$
\begin{aligned}
0 &= \frac{\partial S}{\partial a_2} \\
&= \left(\frac{1}{6}ma_2\left(2t-T\right)^3 - mg\left(\frac{1}{3}t^3 - \frac{1}{2}Tt^2\right)\right)\Bigg|_0^T \\
&= \frac{1}{6}ma_2T^3 - mg\left(\frac{1}{3}T^3 - \frac{1}{2}T^3\right) - \frac{1}{6}ma_2(-T)^3 \qquad (4.30)\\
&= mT^3\left(\frac{1}{3}a_2 - g\left(\frac{1}{3}-\frac{1}{2}\right)\right) \\
&= \frac{1}{3}mT^3\left(a_2 - g\left(1-\frac{3}{2}\right)\right).
\end{aligned}
$$

or

$$a_2 + g/2 = 0, \qquad (4.31)$$

which is the result we are required to show.

Answer for Exercise 4.3

Here we want to show that after a change of variables, provided such a transformation is non-singular, the Euler-Lagrange equations are still valid.

Let us write

$$r_i = r_i(q_1, q_2, \cdots q_N). \qquad (4.35)$$

Our "velocity" variables in terms of the original parametrization q_i are

$$\dot{r}_j = \frac{dr_j}{dt} = \frac{\partial r_j}{\partial q_i}\frac{\partial q_i}{\partial t} = \dot{q}_i\frac{\partial r_j}{\partial q_i}, \qquad (4.36)$$

so we have

$$\frac{\partial \dot{r}_j}{\partial \dot{q}_i} = \frac{\partial r_j}{\partial q_i}. \qquad (4.37)$$

Computing the LHS of the Euler Lagrange equation we find

$$\frac{\partial L}{\partial q_i} = \frac{\partial L}{\partial r_j}\frac{\partial r_j}{\partial q_i} + \frac{\partial L}{\partial \dot{r}_j}\frac{\partial \dot{r}_j}{\partial q_i}. \qquad (4.38)$$

For our RHS we start with

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial r_j}\frac{\partial r_j}{\partial \dot{q}_i} + \frac{\partial L}{\partial \dot{r}_j}\frac{\partial \dot{r}_j}{\partial \dot{q}_i} = \frac{\partial L}{\partial r_j}\frac{\partial r_j}{\partial \dot{q}_i} + \frac{\partial L}{\partial \dot{r}_j}\frac{\partial r_j}{\partial q_i}, \tag{4.39}$$

but $\partial r_j/\partial \dot{q}_i = 0$, so this is just

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial r_j}\frac{\partial r_j}{\partial \dot{q}_i} + \frac{\partial L}{\partial \dot{r}_j}\frac{\partial \dot{r}_j}{\partial \dot{q}_i} = \frac{\partial L}{\partial \dot{r}_j}\frac{\partial r_j}{\partial q_i}. \tag{4.40}$$

The Euler-Lagrange equations become

$$\begin{aligned} 0 &= \frac{\partial L}{\partial r_j}\frac{\partial r_j}{\partial q_i} + \frac{\partial L}{\partial \dot{r}_j}\frac{\partial \dot{r}_j}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}_j}\frac{\partial r_j}{\partial q_i}\right) \\ &= \frac{\partial L}{\partial r_j}\frac{\partial r_j}{\partial q_i} + \cancel{\frac{\partial L}{\partial \dot{r}_j}\frac{\partial \dot{r}_j}{\partial q_i}} - \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{r}_j}\right)\frac{\partial r_j}{\partial q_i} - \cancel{\frac{\partial L}{\partial \dot{r}_j}\frac{d}{dt}\frac{\partial r_j}{\partial q_i}} \\ &= \left(\frac{\partial L}{\partial r_j} - \frac{d}{dt}\frac{\partial L}{\partial \dot{r}_j}\right)\frac{\partial r_j}{\partial q_i}. \end{aligned} \tag{4.41}$$

Since we have an assumption that the transformation is non-singular, we have for all j

$$\frac{\partial r_j}{\partial q_i} \neq 0, \tag{4.42}$$

so we have the Euler-Lagrange equations for the new abstract coordinates as well

$$0 = \frac{\partial L}{\partial r_j} - \frac{d}{dt}\frac{\partial L}{\partial \dot{r}_j}. \tag{4.43}$$

Answer for Exercise 4.4

Solution Part 1. *Conserved energy.* Recall the argument for energy conservation, the result of considering time dependence of the Lagrangian. We have

$$\begin{aligned} \frac{d}{dt}L(q_i, \dot{q}_i, t) &= \frac{\partial L}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial t}\frac{\partial L}{\partial t} \\ &= \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\frac{\partial q_i}{\partial t} + \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial t} + \frac{\partial L}{\partial t} \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\frac{\partial q_i}{\partial t}\right) + \frac{\partial L}{\partial t}. \end{aligned} \tag{4.44}$$

Rearranging we have the conservation equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L\right) + \frac{\partial L}{\partial t} = 0. \tag{4.45}$$

We define the energy as

$$\boxed{\mathcal{E} = \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L,} \tag{4.46}$$

so that the when the Lagrangian is independent of time $\mathcal{E}$ is conserved, and in general

$$\frac{d\mathcal{E}}{dt} = -\frac{\partial L}{\partial t}. \tag{4.47}$$

Application to this problem where our Lagrangian is

$$L = \frac{1}{2}m\mathbf{v}^2 + q\mathbf{v}\cdot\mathbf{A} - q\phi, \tag{4.48}$$

we have

$$\frac{\partial L}{\partial \mathbf{v}} = m\mathbf{v} + q\mathbf{A}. \tag{4.49}$$

so the energy is

$$\begin{aligned}\mathcal{E} &= (m\mathbf{v} + \cancel{q\mathbf{A}})\cdot\mathbf{v} - \left(\frac{1}{2}m\mathbf{v}^2 + \cancel{q\mathbf{v}\cdot\mathbf{A}} - q\phi\right)\\ &= \frac{1}{2}m\mathbf{v}^2 + q\phi,\end{aligned} \tag{4.50}$$

with an end result of

$$\boxed{\mathcal{E} = \frac{1}{2}m\mathbf{v}^2 + q\phi.} \tag{4.51}$$

Solution Part 2. *Find $d\mathcal{E}/dt$.*

With direct computation. There are two ways we can try this. One is with direct computation of the derivative from eq. (4.50)

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= \mathbf{v}\cdot(m\mathbf{a}) + q\frac{d\phi}{dt}\\ &= \mathbf{v}\cdot(q\mathbf{E} + q\mathbf{v}\times\mathbf{B}) + q\left(\frac{\partial\phi}{\partial t} + \mathbf{v}\cdot\nabla\phi\right)\\ &= q\mathbf{v}\cdot(\mathbf{E} + \nabla\phi) + \cancel{q\mathbf{v}\cdot(\mathbf{v}\times\mathbf{B})} + q\frac{\partial\phi}{\partial t}\\ &= q\mathbf{v}\cdot\left(-\cancel{\nabla\phi} - \frac{\partial\mathbf{A}}{\partial t} + \cancel{\nabla\phi}\right) + q\frac{\partial\phi}{\partial t}.\end{aligned} \tag{4.52}$$

So our end result is

$$\boxed{\frac{d\mathcal{E}}{dt} = -q\mathbf{v}\cdot\frac{\partial\mathbf{A}}{\partial t} + q\frac{\partial\phi}{\partial t}.} \tag{4.53}$$

Using the Lagrangian time partial. Doing it explicitly as above is the hard way. We can do it from the conservation identity eq. (4.47) instead

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= -\frac{\partial L}{\partial t}\\ &= -\frac{\partial}{\partial t}\left(\frac{1}{2}m\mathbf{v}^2 + q\mathbf{v}\cdot\mathbf{A} - q\phi\right)\\ &= -q\mathbf{v}\cdot\frac{\partial\mathbf{A}}{\partial t} + q\frac{\partial\phi}{\partial t},\end{aligned} \tag{4.54}$$

as before.

Aside: Why not the "expected" $q\mathbf{v}\cdot\mathbf{E}$ result? From the relativistic treatment I expected

$$\frac{d\mathcal{E}}{dt} \stackrel{?}{=} q\mathbf{v}\cdot\mathbf{E}, \tag{4.55}$$

but that's not what we got. With $\mathcal{E} = m\mathbf{v}^2/2 + q\phi$, it appears that we get a similar result considering just the Kinetic portion of the energy

$$\frac{1}{2}m\mathbf{v}^2 = \mathcal{E} - q\phi. \tag{4.56}$$

Computing the derivative from above we have

$$\begin{aligned}\frac{d}{dt}\left(\frac{1}{2}m\mathbf{v}^2\right) &= -q\mathbf{v}\cdot\frac{\partial\mathbf{A}}{\partial t} + q\frac{\partial\phi}{\partial t} - q\frac{d\phi}{dt}\\ &= -q\mathbf{v}\cdot\frac{\partial\mathbf{A}}{\partial t} + \cancel{q\frac{\partial\phi}{\partial t}} - \cancel{q\frac{\partial\phi}{\partial t}} - q\mathbf{v}\cdot\boldsymbol{\nabla}\phi\\ &= q\mathbf{v}\cdot\left(-\boldsymbol{\nabla}\phi - \frac{\partial\mathbf{A}}{\partial t}\right),\end{aligned} \tag{4.57}$$

or

$$\frac{d}{dt}\left(\frac{1}{2}m\mathbf{v}^2\right) = \frac{d}{dt}\left(\mathcal{E} - q\phi\right) = q\mathbf{v}\cdot\mathbf{E}. \tag{4.58}$$

Looking back to what we did in the relativistic treatment, I see that my confusion was due to the fact that we actually computed

$$\frac{d\mathcal{E}_{\text{kin}}}{dt} = q\mathbf{v}\cdot\mathbf{E}, \tag{4.59}$$

where $\mathcal{E}_{\text{kin}} = \gamma mc^2$. To first order, removing an additive constant, we have $\gamma mc^2 \approx m\mathbf{v}^2/2$, so everything checks out.

Solution Part 3. *Conserved momentum.* The conserved momentum followed from a Noether's argument where we compute

$$\begin{aligned}\frac{dL'}{d\epsilon} &= \frac{\partial L'}{\partial q_i}\frac{\partial q_i}{\partial \epsilon} + \frac{\partial L'}{\partial \dot{q}_i'}\frac{\partial \dot{q}_i'}{\partial \epsilon} \\ &= \left(\frac{d}{dt}\frac{\partial L'}{\partial \dot{q}_i}\right)\frac{\partial q_i}{\partial \epsilon} + \frac{\partial L'}{\partial \dot{q}_i'}\frac{\partial \dot{q}_i'}{\partial \epsilon} \\ &= \frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{q}_i'}\frac{\partial q_i'}{\partial \epsilon}\right),\end{aligned} \tag{4.60}$$

where it has been assumed that a perturbed Lagrangian

$$L'(\epsilon) = L(q_i'(\epsilon), \dot{q}_i'(\epsilon), t), \tag{4.61}$$

also satisfies the Euler Lagrange equations using the transformed coordinates. With the coordinates transformed by a shift along some constant direction $\mathbf{a}$ as in

$$\mathbf{x}' = \mathbf{x} + \epsilon\mathbf{a}, \tag{4.62}$$

we have $\partial\mathbf{x}'/\partial\epsilon = \mathbf{a}$, so eq. (4.60) takes the form

$$\frac{dL'}{d\epsilon} = \frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{x}_i}a_i\right). \tag{4.63}$$

Our shifted Lagrangian for spatially homogeneous potentials $\phi' = \phi$ and $\mathbf{A}' = \mathbf{A}$ is

$$L' = \frac{1}{2}m\mathbf{v}'^2 + q\mathbf{v}' \cdot \mathbf{A} - q\phi = L, \tag{4.64}$$

but $\mathbf{v}' = \mathbf{v}$, so we've just got our canonical momentum $\mathbf{M} = \partial L/\partial \dot{x}_i$ within the time derivative, and must have for all $\mathbf{a}$

$$\frac{d\mathbf{M}}{dt} \cdot \mathbf{a} = 0. \tag{4.65}$$

The conserved momentum is then just

$$\mathbf{M} = m\mathbf{v} + q\mathbf{A}. \tag{4.66}$$

Solution Part 4. *Conserved angular momentum.* Does the conserved angular momentum take the same from as $\mathbf{x} \times \mathbf{M}$ as we had in a non-velocity dependent Lagrangian? We can check using the same Noether arguments using the following coordinate transformation

$$\mathbf{x}' = e^{-\epsilon j/2} \mathbf{x} e^{\epsilon j/2}, \tag{4.67}$$

where $j = \hat{\mathbf{u}} \wedge \hat{\mathbf{v}}$ is the geometric product of two perpendicular unit vectors, and ϵ is the magnitude of the rotation. This gives us

$$\begin{aligned} \frac{d\mathbf{x}'}{d\epsilon} &= -\frac{j}{2} e^{-\epsilon j/2} \mathbf{x} e^{\epsilon j/2} + e^{-\epsilon j/2} \mathbf{x} e^{\epsilon j/2} \frac{j}{2} \\ &= \frac{1}{2} \left(\mathbf{x}' j - j \mathbf{x}' \right) \\ &= \mathbf{x}' \cdot j. \end{aligned} \tag{4.68}$$

The Noether conservation statement is then

$$\frac{dL'}{d\epsilon} = \frac{d}{dt} \left(\frac{\partial L'}{\partial \dot{x}_i} \mathbf{e}_i \cdot (\mathbf{x}' \cdot j) \right). \tag{4.69}$$

With a static scalar potential $\phi(\mathbf{x})$ is our Lagrangian rotation invariant? We have

$$\begin{aligned} L' &= \frac{1}{2} \mathbf{v}'^2 - q\phi(\mathbf{x}') \\ &= \frac{1}{2} \mathbf{v}^2 - q\phi(\mathbf{x}'). \end{aligned} \tag{4.70}$$

With zero vector potential, our kinetic term is invariant since the squared velocity is invariant, but we require $\phi(\mathbf{x}') = \phi(\mathbf{x})$ for total Lagrangian invariance. We have that if $\phi(\mathbf{x}) = \phi(|\mathbf{x}|)$. Evaluating the conservation identity eq. (4.69) at $\epsilon = 0$ we have

$$0 = \frac{d}{dt} \left(\mathbf{M} \cdot (\mathbf{x} \cdot j) \right). \tag{4.71}$$

We are used to seeing this in dual form using the cross product

$$\begin{aligned}
\mathbf{M}\cdot(\mathbf{x}\cdot j) &= \langle \mathbf{M}(\mathbf{x}\cdot j)\rangle \\
&= \frac{1}{2}\langle \mathbf{M}\mathbf{x}j - \mathbf{M}j\mathbf{x}\rangle \\
&= \frac{1}{2}\langle \mathbf{M}\mathbf{x}j - \mathbf{x}\mathbf{M}j\rangle \\
&= \frac{1}{2}\langle \mathbf{M}\wedge\mathbf{x} - \mathbf{x}\wedge\mathbf{M}\rangle\cdot j \\
&= \frac{1}{2}\langle (\mathbf{M}\wedge\mathbf{x} - \mathbf{x}\wedge\mathbf{M})\cdot j\rangle \\
&= (\mathbf{M}\wedge\mathbf{x})\cdot j \\
&= I(\mathbf{M}\times\mathbf{x})\cdot j.
\end{aligned} \tag{4.72}$$

We are left with

$$0 = I\frac{d}{dt}(\mathbf{x}\times\mathbf{M})\cdot j, \tag{4.73}$$

but since j can be arbitrarily oriented, we have a requirement that

$$0 = \frac{d}{dt}(\mathbf{x}\times\mathbf{M}). \tag{4.74}$$

This verifies that the our angular momentum is conserved, provided that $\phi(\mathbf{x}) = \phi(|\mathbf{x}|)$, and $\mathbf{A} = 0$. With $\mathbf{A} = 0$, so that $\mathbf{M} = m\mathbf{v} + q\mathbf{A} = m\mathbf{v}$ this is just

$$\mathbf{x}\times\mathbf{M} = m\mathbf{x}\times\mathbf{v}. \tag{4.75}$$

Note that the dependency on geometric algebra in the Noether's argument above can probably be eliminated by utilizing a rotational transformation of the form

$$\mathbf{x}' = \mathbf{x} + \hat{\mathbf{n}}\times\mathbf{x}. \tag{4.76}$$

I'd guess (or perhaps recall if I attended that class), that this was the approach used.

Answer for Exercise 4.5

Solution Part 1. *Spherical coordinates.* In Cartesian coordinates our angular momentum is

$$\begin{aligned}
\mathbf{M} &= \mathbf{r}\times(m\mathbf{v}) \\
&= m(yv_z - zv_y)\hat{\mathbf{x}} + m(zv_x - xv_z)\hat{\mathbf{y}} + m(xv_y - yv_x)\hat{\mathbf{z}}.
\end{aligned} \tag{4.77}$$

Substituting x, y, z is easy since we have

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = r \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}, \tag{4.78}$$

but the $\mathbf{v}$ substitution requires more work. We have

$$\begin{aligned} \mathbf{v} &= \frac{d\mathbf{r}}{dt} \\ &= \frac{d}{dt}(r\hat{\mathbf{r}}) \\ &= \dot{r}\hat{\mathbf{r}} + r\frac{d\hat{\mathbf{r}}}{dt}. \end{aligned} \tag{4.79}$$

$$\begin{aligned} \frac{d\hat{\mathbf{r}}}{dt} &= \frac{d}{dt}\begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta\cos\phi\dot{\theta} - \sin\theta\sin\phi\dot{\phi} \\ \cos\theta\sin\phi\dot{\theta} + \sin\theta\cos\phi\dot{\phi} \\ -\sin\theta\dot{\theta} \end{bmatrix}. \end{aligned} \tag{4.80}$$

So we have

$$\mathbf{v} = \begin{bmatrix} \dot{r}\sin\theta\cos\phi + r\cos\theta\cos\phi\dot{\theta} - r\sin\theta\sin\phi\dot{\phi} \\ \dot{r}\sin\theta\sin\phi + r\cos\theta\sin\phi\dot{\theta} + r\sin\theta\cos\phi\dot{\phi} \\ \dot{r}\cos\theta - r\sin\theta\dot{\theta} \end{bmatrix}, \tag{4.81a}$$

$$\frac{\mathbf{M}}{mr} = \begin{bmatrix} \sin\theta\sin\phi v_z - \cos\theta v_y \\ \cos\theta v_x - \sin\theta\cos\phi v_z \\ \sin\theta\cos\phi v_y - \sin\theta\sin\phi v_x \end{bmatrix}. \tag{4.81b}$$

Expanding this is a bit of a mess, but it eventually simplifies. We start with

$$\begin{bmatrix} S_\theta S_\phi(\dot{r}C_\theta - rS_\theta\dot{\theta}) - C_\theta(\dot{r}S_\theta S_\phi + rC_\theta S_\phi\dot{\theta} + rS_\theta C_\phi\dot{\phi}) \\ C_\theta(\dot{r}S_\theta C_\phi + rC_\theta C_\phi\dot{\theta} - rS_\theta S_\phi\dot{\phi}) - S_\theta C_\phi(\dot{r}C_\theta - rS_\theta\dot{\theta}) \\ S_\theta C_\phi(\dot{r}S_\theta S_\phi + rC_\theta S_\phi\dot{\theta} + rS_\theta C_\phi\dot{\phi}) - S_\theta S_\phi(\dot{r}S_\theta C_\phi + rC_\theta C_\phi\dot{\theta} - rS_\theta S_\phi\dot{\phi}) \end{bmatrix},$$

$$(4.82)$$

then

$$\begin{bmatrix} \cancel{\dot{r}C_\theta S_\theta S_\phi} - r\dot{\theta}S_\theta^2 S_\phi - \cancel{\dot{r}S_\theta C_\theta S_\phi} - r\dot{\theta}C_\theta^2 S_\phi - r\dot{\phi}S_\theta C_\theta C_\phi \\ \cancel{\dot{r}S_\theta C_\theta C_\phi} + r\dot{\theta}C_\theta^2 C_\phi - r\dot{\phi}S_\theta C_\theta S_\phi - \cancel{\dot{r}C_\theta S_\theta C_\phi} + r\dot{\theta}S_\theta^2 C_\phi \\ \cancel{\dot{r}S_\theta^2 S_\phi C_\phi} + \cancel{r\dot{\theta}C_\theta S_\theta C_\phi S_\phi} + r\dot{\phi}S_\theta^2 C_\phi^2 - \cancel{\dot{r}S_\theta^2 S_\phi C_\phi} - \cancel{r\dot{\theta}C_\theta S_\theta S_\phi C_\phi} + r\dot{\phi}S_\theta^2 S_\phi^2 \end{bmatrix}, \tag{4.83}$$

and finally

$$\begin{bmatrix} -r\dot{\theta}S_\phi - r\dot{\phi}S_\theta C_\theta C_\phi \\ +r\dot{\theta}C_\phi - r\dot{\phi}S_\theta C_\theta S_\phi \\ +r\dot{\phi}S_\theta^2 \end{bmatrix}. \tag{4.84}$$

In matrix form, we have (and can read off M_x, M_y, M_z)

$$\boxed{\mathbf{M} = \frac{1}{2}mr^2 \begin{bmatrix} -2\sin\phi & -\sin(2\theta)\cos\phi \\ 2\cos\phi & -\sin(2\theta)\sin\phi \\ 0 & 1-\cos(2\theta) \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{\phi} \end{bmatrix}.} \tag{4.85}$$

We have also been asked to find $\mathbf{M}^2$ and can write this as a quadratic form

$$\begin{aligned} \mathbf{M}^2 &= \frac{1}{4}m^2r^4 \begin{bmatrix} \dot{\theta} & \dot{\phi} \end{bmatrix} \begin{bmatrix} -2\sin\phi & 2\cos\phi & 0 \\ -\sin(2\theta)\cos\phi & -\sin(2\theta)\sin\phi & 1-\cos(2\theta) \end{bmatrix} \times \\ &\quad \begin{bmatrix} -2\sin\phi & -\sin(2\theta)\cos\phi \\ 2\cos\phi & -\sin(2\theta)\sin\phi \\ 0 & 1-\cos(2\theta) \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{\phi} \end{bmatrix} \\ &= \frac{1}{4}m^2r^4 \begin{bmatrix} \dot{\theta} & \dot{\phi} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2(1-\cos(2\theta)) \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{\phi} \end{bmatrix}. \end{aligned} \tag{4.86}$$

This simplifies surprisingly, leaving only

$$\boxed{\mathbf{M}^2 = m^2r^4\left(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2\right).} \tag{4.87}$$

Solution Part 1. *Spherical coordinates - a smarter way.* Observe that we have no $\dot{r}$ factors in the angular momentum. This makes sense when we consider that the total angular momentum is

$$\mathbf{M} = mr\hat{\mathbf{r}} \times \mathbf{v}, \tag{4.88}$$

so the $\dot{r}\hat{\mathbf{r}}$ term of the velocity is necessarily killed. Let us do that simplification first. We want our velocity completely specified in a $\{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}\}$ basis, and note that our basis vectors are

$$\begin{aligned}
\hat{\mathbf{r}} &= \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix} \\
\hat{\boldsymbol{\theta}} &= \begin{bmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{bmatrix} \\
\hat{\boldsymbol{\phi}} &= \begin{bmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{bmatrix}.
\end{aligned} \tag{4.89}$$

We wish to rewrite

$$\frac{d\hat{\mathbf{r}}}{dt} = \begin{bmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ -\sin\theta & 0 \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{\phi} \end{bmatrix}, \tag{4.90}$$

in terms of these spherical unit vectors and find

$$\begin{aligned}
\frac{d\hat{\mathbf{r}}}{dt} \cdot \hat{\mathbf{r}} &= \hat{\mathbf{r}}^{\mathrm{T}} \frac{d\hat{\mathbf{r}}}{dt} = 0 \\
\frac{d\hat{\mathbf{r}}}{dt} \cdot \hat{\boldsymbol{\theta}} &= \hat{\boldsymbol{\theta}}^{\mathrm{T}} \frac{d\hat{\mathbf{r}}}{dt} = \dot{\theta} \\
\frac{d\hat{\mathbf{r}}}{dt} \cdot \hat{\boldsymbol{\phi}} &= \hat{\boldsymbol{\phi}}^{\mathrm{T}} \frac{d\hat{\mathbf{r}}}{dt} = \dot{\phi}\sin\theta.
\end{aligned} \tag{4.91}$$

So our velocity is

$$\mathbf{v} = \dot{r}\hat{\mathbf{r}} + r\left(\dot{\theta}\hat{\boldsymbol{\theta}} + \dot{\phi}\sin\theta\hat{\boldsymbol{\phi}}\right). \tag{4.92}$$

As an aside, now that we know the Euler-Lagrange methods, we could also compute this velocity from the spherical free particle Lagrangian by writing out the canonical momentum in vector form. We have

$$L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2\theta\right). \tag{4.93}$$

We expect our canonical momentum in vector form to be

$$\begin{aligned}
\mathbf{P} &= \frac{\partial L}{\partial \dot{r}}\hat{\mathbf{r}} + \frac{\partial L}{\partial \dot{\theta}}\frac{\hat{\boldsymbol{\theta}}}{r} + \frac{\partial L}{\partial \dot{\phi}}\frac{\hat{\boldsymbol{\phi}}}{r\sin\theta} \\
&= m\dot{r}\hat{\mathbf{r}} + mr^2\dot{\theta}\frac{\hat{\boldsymbol{\theta}}}{r} + mr^2\sin^2\theta\dot{\phi}\frac{\hat{\boldsymbol{\phi}}}{r\sin\theta} \\
&= m\left(\dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}} + r\sin\theta\dot{\phi}\hat{\boldsymbol{\phi}}\right) \\
&= m\mathbf{v}.
\end{aligned} \tag{4.94}$$

This is consistent with eq. (4.92) calculated hard way, and is a nice verification that the canonical momentum matches the expectation of being nothing more than how to express the momentum in different coordinate systems. Returning to the angular momentum calculation we have

$$\begin{aligned}
\hat{\mathbf{r}}\times\mathbf{v} &= r\hat{\mathbf{r}}\times\left(\dot{\theta}\hat{\boldsymbol{\theta}} + \dot{\phi}\sin\theta\hat{\boldsymbol{\phi}}\right) \\
&= r\left(\dot{\theta}\hat{\boldsymbol{\phi}} - \dot{\phi}\sin\theta\hat{\boldsymbol{\theta}}\right).
\end{aligned} \tag{4.95}$$

Our total angular momentum in vector form is

$$\mathbf{M} = mr^2\left(\dot{\theta}\hat{\boldsymbol{\phi}} - \dot{\phi}\sin\theta\hat{\boldsymbol{\theta}}\right). \tag{4.96}$$

Now, should we wish to extract coordinates with respect to x, y, z we just have to write our vectors $\hat{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{\theta}}$ in the $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ basis and have

$$\begin{aligned}
\mathbf{M} &= mr^2\begin{bmatrix}\hat{\boldsymbol{\phi}} & -\sin\theta\hat{\boldsymbol{\theta}}\end{bmatrix}\begin{bmatrix}\dot{\theta} \\ \dot{\phi}\end{bmatrix} \\
&= mr^2\begin{bmatrix}-\sin\phi & -\sin\theta(\cos\theta\cos\phi) \\ \cos\phi & -\sin\theta(\cos\theta\sin\phi) \\ 0 & \sin^2\theta\end{bmatrix}\begin{bmatrix}\dot{\theta} \\ \dot{\phi}\end{bmatrix}.
\end{aligned} \tag{4.97}$$

This matches eq. (4.85), but all the messy trig is isolated to the calculation of $\mathbf{v}$ in the spherical polar basis.

Solution Part 2. *Cylindrical coordinates.* This one should be easier. To start our position vector is

$$\mathbf{r} = \begin{bmatrix}\rho\cos\phi \\ \rho\sin\phi \\ z\end{bmatrix} = \rho\hat{\boldsymbol{\rho}} + z\hat{\mathbf{z}}. \tag{4.98}$$

Our velocity is

$$\begin{aligned}
\mathbf{v} &= \dot{\rho}\hat{\boldsymbol{\rho}} + \rho\frac{d\hat{\boldsymbol{\rho}}}{dt} + \dot{z}\hat{\mathbf{z}} \\
&= \dot{\rho}\hat{\boldsymbol{\rho}} + \rho\frac{d}{dt}\left(\mathbf{e}_1 e^{i\phi}\right) + \dot{z}\hat{\mathbf{z}} \\
&= \dot{\rho}\hat{\boldsymbol{\rho}} + \rho\dot{\phi}\mathbf{e}_2 e^{i\phi} + \dot{z}\hat{\mathbf{z}} \\
&= \dot{\rho}\hat{\boldsymbol{\rho}} + \rho\dot{\phi}\hat{\boldsymbol{\phi}} + \dot{z}\hat{\mathbf{z}}.
\end{aligned} \tag{4.99}$$

Here, I have used the Clifford algebra representation of $\hat{\boldsymbol{\rho}}$ with the plane bivector $i = \mathbf{e}_1\mathbf{e}_2$. In coordinates we have

$$\hat{\boldsymbol{\phi}} = \mathbf{e}_2\left(\cos\phi + \mathbf{e}_1\mathbf{e}_2\sin\phi\right) = -\mathbf{e}_1\sin\phi + \mathbf{e}_2\cos\phi, \tag{4.100}$$

so our velocity in matrix form is

$$\begin{aligned}
\mathbf{v} &= \dot{\rho}\begin{bmatrix}\cos\phi \\ \sin\phi \\ 0\end{bmatrix} + \rho\dot{\phi}\begin{bmatrix}-\sin\phi \\ \cos\phi \\ 0\end{bmatrix} + \dot{z}\begin{bmatrix}0 \\ 0 \\ 1\end{bmatrix} \\
&= \begin{bmatrix}\dot{\rho}\cos\phi - \rho\dot{\phi}\sin\phi \\ \dot{\rho}\sin\phi + \rho\dot{\phi}\cos\phi \\ \dot{z}\end{bmatrix}.
\end{aligned} \tag{4.101}$$

For our angular momentum we get

$$\begin{aligned}
\mathbf{M} &= \mathbf{r}\times(m\mathbf{v}) \\
&= m\begin{bmatrix}\rho\sin\phi\dot{z} - z\left(\dot{\rho}\sin\phi + \rho\dot{\phi}\cos\phi\right) \\ z\left(\dot{\rho}\cos\phi - \rho\dot{\phi}\sin\phi\right) - \rho\cos\phi\dot{z} \\ \rho\cos\phi\left(\cancel{\dot{\rho}\sin\phi} + \rho\dot{\phi}\cos\phi\right) - \rho\sin\phi\left(\cancel{\dot{\rho}\cos\phi} - \rho\dot{\phi}\sin\phi\right)\end{bmatrix}.
\end{aligned} \tag{4.102}$$

We can now read off M_x, M_y, M_z by inspection

$$\mathbf{M} = m\begin{bmatrix}(\rho\dot{z} - z\dot{\rho})\sin\phi - z\rho\dot{\phi}\cos\phi \\ (z\dot{\rho} - \rho\dot{z})\cos\phi - z\rho\dot{\phi}\sin\phi \\ \rho^2\dot{\phi}\end{bmatrix}. \tag{4.103}$$

We also want the (squared) magnitude, which is

$$\mathbf{M}^2 = m^2\left((\rho\dot{z} - z\dot{\rho})^2 + \rho^2\dot{\phi}^2(z^2 + \rho^2)\right). \tag{4.104}$$

Answer for Exercise 4.6

The center of mass vector is:

$$\mathbf{R} = \frac{1}{3}(\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3). \tag{4.108}$$

This can be used to express each of the position vectors in terms of the $\boldsymbol{\rho}_i$ vectors:

$$\begin{aligned} 3m\mathbf{R} &= m(\boldsymbol{\rho}_1 + \mathbf{r}_3) + m(\boldsymbol{\rho}_2 + \mathbf{r}_3) + m\mathbf{r}_3 \\ &= 2m(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2) + 3m\mathbf{r}_3 \\ \mathbf{r}_3 &= \mathbf{R} - \frac{1}{3}(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2) \\ \mathbf{r}_2 = \boldsymbol{\rho}_2 + \mathbf{r}_3 &= \boldsymbol{\rho}_2 + \mathbf{r}_3 = \frac{2}{3}\boldsymbol{\rho}_2 - \frac{1}{2}\boldsymbol{\rho}_1 + \mathbf{R} \\ \mathbf{r}_1 = \boldsymbol{\rho}_1 + \mathbf{r}_3 &= \frac{2}{3}\boldsymbol{\rho}_1 - \frac{1}{2}\boldsymbol{\rho}_2 + \mathbf{R}. \end{aligned} \tag{4.109}$$

Now, that is enough to specify the part of the Lagrangian from the potentials that act between all the particles

$$L_V = \sum -V_{ij} = g\left(e^{-\mu|\boldsymbol{\rho}_1|} + e^{-\mu|\boldsymbol{\rho}_2|} + e^{-\mu|\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2|}\right). \tag{4.110}$$

Now, we need to calculate the two U potential terms. If we consider with positions $\mathbf{r}_1$, and $\mathbf{r}_2$ to be the ones that can exert a force on the third, the velocities of those masses relative to $\mathbf{r}_3$ are:

$$(\mathbf{r}_3 - \mathbf{r}_k)' = \dot{\boldsymbol{\rho}}_k. \tag{4.111}$$

So, the potential parts of the Lagrangian are

$$L_{U+V} = g\left(e^{-\mu|\boldsymbol{\rho}_1|} + e^{-\mu|\boldsymbol{\rho}_2|} + e^{-\mu|\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2|}\right) + f\left(\mathbf{R} - \frac{1}{3}(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2)\right) \cdot (\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2). \tag{4.112}$$

The kinetic part (omitting the m/2 factor), in terms of the CM and relative vectors is

$$\begin{aligned} (\mathbf{v}_1)^2 + (\mathbf{v}_2)^2 + (\mathbf{v}_3)^2 &= \left(\frac{2}{3}\dot{\boldsymbol{\rho}}_1 - \frac{1}{2}\dot{\boldsymbol{\rho}}_2 + \dot{\mathbf{R}}\right)^2 + \left(\frac{2}{3}\dot{\boldsymbol{\rho}}_2 - \frac{1}{2}\dot{\boldsymbol{\rho}}_1 + \dot{\mathbf{R}}\right)^2 \\ &\quad + \left(\dot{\mathbf{R}} - \frac{1}{3}(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2)\right)^2 \\ &= 3\dot{\mathbf{R}}^2 + (5/9 + 1/4)((\dot{\boldsymbol{\rho}}_1)^2 + (\dot{\boldsymbol{\rho}}_2)^2) \\ &\quad + 2(-2/3 + 1/9)\dot{\boldsymbol{\rho}}_1 \cdot \dot{\boldsymbol{\rho}}_1 + 2(1/3 - 1/2)(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2) \cdot \dot{\mathbf{R}}. \end{aligned} \tag{4.113}$$

So the kinetic part of the Lagrangian is

$$L_K = \frac{3m}{2}\dot{\mathbf{R}}^2 + \frac{29m}{72}((\dot{\boldsymbol{\rho}}_1)^2 + (\dot{\boldsymbol{\rho}}_2)^2) - \frac{5m}{9}\dot{\boldsymbol{\rho}}_1 \cdot \dot{\boldsymbol{\rho}}_2 - \frac{m}{6}(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2) \cdot \dot{\mathbf{R}}. \tag{4.114}$$

and finally, the total Lagrangian is

$$\begin{aligned} L = & \frac{3m}{2}\dot{\mathbf{R}}^2 + \frac{29m}{72}((\dot{\boldsymbol{\rho}}_1)^2 + (\dot{\boldsymbol{\rho}}_2)^2) - \frac{5m}{9}\dot{\boldsymbol{\rho}}_1 \cdot \dot{\boldsymbol{\rho}}_2 - \frac{m}{6}(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2) \cdot \dot{\mathbf{R}} \\ & + g\left(e^{-\mu|\boldsymbol{\rho}_1|} + e^{-\mu|\boldsymbol{\rho}_2|} + e^{-\mu|\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2|}\right) + f\left(\mathbf{R} - \frac{1}{3}(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2)\right) \cdot (\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2)\,. \end{aligned} \tag{4.115}$$

Angular momentum conservation? How about the angular momentum conservation question? How to answer that? One way would be to compute the forces from the Lagrangian, and take cross products but is that really the best way? Perhaps the answer is as simple as observing that there are no external torque's on the system, thus $d\mathbf{L}/dt = 0$, or angular momentum for the system is constant (conserved). Is that actually the case with these velocity dependent potentials? It was suggested to me on PF that I should look at how this Lagrangian transforms under rotation, and use Noether's theorem. The Goldstein book does not explicitly mention this theorem that I can see, and I do not think it was covered yet if it did. Suppose we did know about Noether's theorem for this problem (as I now do with in this revisiting of this problem to complete it), we would have to see if the Lagrangian is invariant under rotation. Suppose that a rigid rotation is introduced, which we can write in GA formalism using dual sided quaternion products

$$\mathbf{x} \to \mathbf{x}' = e^{-i\hat{\mathbf{n}}\alpha/2}\mathbf{x}e^{i\hat{\mathbf{n}}\alpha/2}. \tag{4.116}$$

(could probably also use a matrix formulation, but the parametrization is messier). For all the relative vectors $\boldsymbol{\rho}_k$ we have

$$\left|\boldsymbol{\rho}'_k\right| = \left|\boldsymbol{\rho}_k\right|. \tag{4.117}$$

So all the V potential interactions are invariant. Since the rotation quaternion here is a fixed non-time dependent quantity we have

$$\dot{\boldsymbol{\rho}}'_k = e^{-i\hat{\mathbf{n}}\alpha/2}\dot{\boldsymbol{\rho}}_k e^{i\hat{\mathbf{n}}\alpha/2}, \tag{4.118}$$

so for the dot product in the remaining potential term we have

$$
\begin{aligned}
&\left(\mathbf{R}' - \frac{1}{3}\left(\boldsymbol{\rho}_1' + \boldsymbol{\rho}_2'\right)\right) \cdot \left(\dot{\boldsymbol{\rho}}_1' + \dot{\boldsymbol{\rho}}_2'\right) \\
&\quad = \left(e^{-i\hat{\mathbf{n}}\alpha/2}\left(\mathbf{R} - \frac{1}{3}\left(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\right)\right)e^{i\hat{\mathbf{n}}\alpha/2}\right) \cdot \left(e^{-i\hat{\mathbf{n}}\alpha/2}\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2 e^{i\hat{\mathbf{n}}\alpha/2}\right) \\
&\quad = \left\langle e^{-i\hat{\mathbf{n}}\alpha/2}\left(\mathbf{R} - \frac{1}{3}\left(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\right)\right)e^{i\hat{\mathbf{n}}\alpha/2}e^{-i\hat{\mathbf{n}}\alpha/2}\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2 e^{i\hat{\mathbf{n}}\alpha/2}\right\rangle \\
&\quad = \left\langle e^{-i\hat{\mathbf{n}}\alpha/2}\left(\mathbf{R} - \frac{1}{3}\left(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\right)\right)\left(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2\right)e^{i\hat{\mathbf{n}}\alpha/2}\right\rangle \qquad (4.119) \\
&\quad = \left\langle e^{i\hat{\mathbf{n}}\alpha/2}e^{-i\hat{\mathbf{n}}\alpha/2}\left(\mathbf{R} - \frac{1}{3}\left(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\right)\right)\left(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2\right)\right\rangle \\
&\quad = \left\langle\left(\mathbf{R} - \frac{1}{3}\left(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\right)\right)\left(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2\right)\right\rangle \\
&\quad = \left(\mathbf{R} - \frac{1}{3}\left(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\right)\right) \cdot \left(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2\right).
\end{aligned}
$$

So, presuming I interpreted the $\mathbf{r}$ in $\mathbf{v} \cdot \mathbf{r}$ correctly, all the vector quantities in the Lagrangian are rotation invariant, and by Noether's we should have system angular momentum conservation.

Application of Noether's. Invoking Noether's here seems like cheating, at least without computing the conserved current, so let us do this. To make this easier, suppose we generalize the Lagrangian slightly to get rid of all the peculiar and specific numerical constants. Let $\rho_3 = \mathbf{R}$, then our Lagrangian has the functional form

$$
L = \alpha^{ij}\dot{\boldsymbol{\rho}}_i \cdot \dot{\boldsymbol{\rho}}_j + g^i e^{-\mu|\boldsymbol{\rho}_i|} + g^{ij} e^{-\mu|\boldsymbol{\rho}_i - \boldsymbol{\rho}_j|} + f^i \boldsymbol{\rho}_i \cdot (\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2). \qquad (4.120)
$$

We can then pick specific α^{ij}, f^i, and g^{ij} (not all non-zero), to match the Lagrangian of this problem. This could be expanded in terms of coordinates, producing nine generalized coordinates and nine corresponding velocity terms, but since our Lagrangian transformation is so naturally expressed in vector form this does not seem like a reasonable thing to do. Let us step up the abstraction one more level instead and treat the Noether symmetry in the more general case, supposing that we have a Lagrangian that is invariant under the same rotational transformation applied above, but has the following general form with explicit vector parametrization, where as above, all our vectors come in functions of the dot products (ei-

ther explicit or implied by absolute values) of our vectors or their time derivatives

$$L = f(\mathbf{x}_k \cdot \mathbf{x}_j, \mathbf{x}_k \cdot \dot{\mathbf{x}}_j, \dot{\mathbf{x}}_k \cdot \dot{\mathbf{x}}_j). \tag{4.121}$$

Having all the parametrization being functions of dot products gives the desired rotational symmetry for the Lagrangian. This must be however, not a dot product with an arbitrary vector, but one of the generalized vector parameters of the Lagrangian. Something like the $\mathbf{A} \cdot \mathbf{v}$ term in the Lorentz force Lagrangian does not have this invariance since $\mathbf{A}$ does not transform along with $\mathbf{v}$. Also Note that the absolute values of the $\boldsymbol{\rho}_k$ vectors are functions of dot products. Now we are in shape to compute the conserved "current" for a rotational symmetry. Our vectors and their derivatives are explicitly rotated

$$\begin{aligned} \mathbf{x}'_k &= e^{-i\hat{\mathbf{n}}\alpha/2}\mathbf{x}_k e^{i\hat{\mathbf{n}}\alpha/2} \\ \dot{\mathbf{x}}'_k &= e^{-i\hat{\mathbf{n}}\alpha/2}\dot{\mathbf{x}}_k e^{i\hat{\mathbf{n}}\alpha/2}, \end{aligned} \tag{4.122}$$

and our Lagrangian is assumed, as above with all vectors coming in dot product pairs, to have rotational invariance when all the vectors in the system are rotated

$$L \to L'(\mathbf{x}'_k, \dot{\mathbf{x}}'_j) = L(\mathbf{x}_k, \dot{\mathbf{x}}_j). \tag{4.123}$$

The essence of Noether's theorem was applied chain rule, looking at how the transformed Lagrangian changes with respect to the transformation. In this case we want to calculate

$$\left.\frac{dL'}{d\alpha}\right|_{\alpha=0}. \tag{4.124}$$

First seeing the Noether's derivation, I did not understand why the evaluation at $\alpha = 0$ was required, even after doing this derivation for myself in 8 (after an initial botched attempt), but the reason for it actually became clear with this application, as writing it up will show. Anyways, back to the derivative. One way to evaluate this would be in terms of coordinates, writing $\mathbf{x}'_k = \mathbf{e}^m x'_{km}$,

$$\frac{dL'}{d\alpha}(\mathbf{x}'_k, \dot{\mathbf{x}}'_j) = \sum_{k,m} \frac{\partial L'}{\partial x'_{km}} \frac{\partial x'_{km}}{\partial \alpha} + \frac{\partial L'}{\partial \dot{x}'_{km}} \frac{\partial \dot{x}'_{km}}{\partial \alpha}. \tag{4.125}$$

This is a bit of a mess however, and begs for some shorthand. Let us write

$$\begin{aligned}\boldsymbol{\nabla}_{\mathbf{x}'_k} L' &= e^m \frac{\partial L'}{\partial x'_{km}} \\ \boldsymbol{\nabla}_{\dot{\mathbf{x}}'_k} L' &= e^m \frac{\partial L'}{\partial \dot{x}'_{km}}.\end{aligned} \tag{4.126}$$

Then the chain rule application above becomes

$$\frac{dL'}{d\alpha}(\mathbf{x}'_k, \dot{\mathbf{x}}'_j) = \sum_k \left(\boldsymbol{\nabla}_{\mathbf{x}'_k} L'\right) \cdot \frac{\partial \mathbf{x}'_k}{\partial \alpha} + \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}'_k} L'\right) \cdot \frac{\partial \dot{\mathbf{x}}'_k}{\partial \alpha}. \tag{4.127}$$

Now, while this notational sugar unfortunately has an obscuring effect, it is also practical since we can now work with the transformed position and velocity vectors directly

$$\begin{aligned}\frac{\partial \mathbf{x}'_k}{\partial \alpha} &= (-i\hat{\mathbf{n}}/2)e^{-i\hat{\mathbf{n}}\alpha/2}\mathbf{x}_k e^{i\hat{\mathbf{n}}\alpha/2} + e^{-i\hat{\mathbf{n}}\alpha/2}\mathbf{x}_k e^{i\hat{\mathbf{n}}\alpha/2}(i\hat{\mathbf{n}}/2) \\ &= (-i\hat{\mathbf{n}}/2)\mathbf{x}'_k + \mathbf{x}'_k(i\hat{\mathbf{n}}/2) \\ &= i(\hat{\mathbf{n}} \wedge \mathbf{x}'_k).\end{aligned} \tag{4.128}$$

So we have

$$\frac{dL'}{d\alpha}(\mathbf{x}'_k, \dot{\mathbf{x}}'_j) = \sum_k \left(\boldsymbol{\nabla}_{\mathbf{x}'_k} L'\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \mathbf{x}'_k)\right) + \sum_k \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}'_k} L'\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \dot{\mathbf{x}}'_k)\right). \tag{4.129}$$

Next step is to reintroduce the notational sugar noting that we can vectorize the Euler-Lagrange equations by writing

$$\boldsymbol{\nabla}_{\mathbf{x}_k} L = \frac{d}{dt}\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L. \tag{4.130}$$

We have now a three fold reduction in the number of Euler-Lagrange equations. For each of the generalized vector parameters, we have the Lagrangian gradient with respect to that vector parameter (a generalized force) equals the time rate of change of the velocity gradient. Inserting this we have

$$\frac{dL'}{d\alpha}(\mathbf{x}'_k, \dot{\mathbf{x}}'_j) = \sum_k \left(\frac{d}{dt}\boldsymbol{\nabla}_{\dot{\mathbf{x}}'_k} L'\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \mathbf{x}'_k)\right) + \sum_k \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}'_k} L'\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \dot{\mathbf{x}}'_k)\right). \tag{4.131}$$

Now we can drop the primes in gradient terms because of the Lagrangian invariance for this symmetry, and are left almost with a perfect differential

$$\frac{dL'}{d\alpha}(\mathbf{x}'_k, \dot{\mathbf{x}}'_j) = \sum_k \left(\frac{d}{dt}\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \mathbf{x}'_k)\right) + \sum_k \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \dot{\mathbf{x}}'_k)\right). \tag{4.132}$$

Here is where the evaluation at $\alpha = 0$ comes in, since $\mathbf{x}'_k(\alpha = 0) = \mathbf{x}_k$, and we can now antidifferentiate

$$\begin{aligned}
\frac{dL'}{d\alpha}(\mathbf{x}'_k, \dot{\mathbf{x}}'_j)\bigg|_{\alpha=0} &= \sum_k \left(\frac{d}{dt}\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \mathbf{x}_k)\right) + \sum_k \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \dot{\mathbf{x}}_k)\right) \\
&= \sum_k \frac{d}{dt}\left(\left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) \cdot \left(i(\hat{\mathbf{n}} \wedge \mathbf{x}_k)\right)\right) \\
&= \sum_k \frac{d}{dt}\left\langle \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) i(\hat{\mathbf{n}} \wedge \mathbf{x}_k)\right\rangle \\
&= \sum_k \frac{d}{dt}\frac{1}{2}\left\langle \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) i(\hat{\mathbf{n}}\mathbf{x}_k - \mathbf{x}_k\hat{\mathbf{n}})\right\rangle \\
&= \sum_k \frac{d}{dt}\frac{1}{2}\left\langle \hat{\mathbf{n}} i \left(\mathbf{x}_k \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) - \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) \mathbf{x}_k\right)\right\rangle \\
&= \sum_k \frac{d}{dt}\frac{1}{2}\left\langle \hat{\mathbf{n}} i \left(\mathbf{x}_k \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) - \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) \mathbf{x}_k\right)\right\rangle \\
&= \sum_k \frac{d}{dt}\left\langle \hat{\mathbf{n}} i \left(\mathbf{x}_k \wedge \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right)\right)\right\rangle \\
&= \sum_k \frac{d}{dt}\left\langle \hat{\mathbf{n}} i^2 \left(\mathbf{x}_k \times \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right)\right)\right\rangle \\
&= \sum_k \frac{d}{dt} - \hat{\mathbf{n}} \cdot \left(\mathbf{x}_k \times \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right)\right).
\end{aligned} \tag{4.133}$$

Because of the symmetry this entire derivative is zero, so we have

$$\hat{\mathbf{n}} \cdot \sum_k \left(\mathbf{x}_k \times \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right)\right) = \text{constant}. \tag{4.134}$$

The Lagrangian velocity gradient can be identified as the momentum (ie: the canonical momentum conjugate to $\mathbf{x}_k$)

$$\mathbf{p}_k \equiv \boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L. \tag{4.135}$$

Also noting that this is constant for any $\hat{\mathbf{n}}$, we finally have the conserved "current" for a rotational symmetry of a system of particles

$$\sum_k \mathbf{x}_k \times \mathbf{p}_k = \text{constant}. \tag{4.136}$$

This digression to Noether's appears to be well worth it for answering the angular momentum question of the problem. Glibly saying "yes angular momentum is conserved", just because the Lagrangian has a rotational symmetry is not enough. We have seen in this particular problem that the Lagrangian, having only dot products has the rotational symmetry, but because of the velocity dependent potential terms $f^i \dot{\boldsymbol{\rho}}_k \cdot \dot{\boldsymbol{\rho}}_j$, the normal Kinetic energy momentum vectors are not equal to the canonical conjugate momentum vectors. Only when the angular momentum is generalized, and written in terms of the canonical conjugate momentum is the total system angular momentum conserved. Namely, the generalized angular momentum for this problem is conserved

$$\sum_k \mathbf{x}_k \times \left(\boldsymbol{\nabla}_{\dot{\mathbf{x}}_k} L\right) = \text{constant}. \tag{4.137}$$

but the "traditional" angular momentum $\sum_k \mathbf{x}_k \times m\dot{\mathbf{x}}_k$, is not.

Answer for Exercise 4.7

Barbell shape, equal masses. center of rod between masses constrained to circular motion. Assuming motion in a plane, the equation for the center of the rod is:

$$c = ae^{i\theta}, \tag{4.138}$$

and the two mass points positions are:

$$\begin{aligned} q_1 &= c + (l/2)e^{i\alpha} \\ q_2 &= c - (l/2)e^{i\alpha}. \end{aligned} \tag{4.139}$$

taking derivatives:

$$\begin{aligned} \dot{q}_1 &= ai\dot{\theta}e^{i\theta} + (l/2)i\dot{\alpha}e^{i\alpha} \\ \dot{q}_2 &= ai\dot{\theta}e^{i\theta} - (l/2)i\dot{\alpha}e^{i\alpha}. \end{aligned} \tag{4.140}$$

and squared magnitudes:

$$\begin{aligned} \dot{q}_\pm &= \left|a\dot{\theta} \pm (l/2)\dot{\alpha}e^{i(\alpha-\theta)}\right|^2 \\ &= \left(a\dot{\theta} \pm \frac{1}{2}l\dot{\alpha}\cos(\alpha-\theta)\right)^2 + \left(\frac{1}{2}l\dot{\alpha}\sin(\alpha-\theta)\right)^2. \end{aligned} \tag{4.141}$$

Summing the kinetic terms yields

$$K = m\left(a\dot{\theta}\right)^2 + m\left(\frac{1}{2}l\dot{\alpha}\right)^2. \tag{4.142}$$

Summing the potential energies, presuming that the motion is vertical, we have:

$$V = mg(l/2)\cos\theta - mg(l/2)\cos\theta, \tag{4.143}$$

so the Lagrangian is just the Kinetic energy.

Taking derivatives to get the EOMs we have:

$$\begin{aligned}\left(ma^2\dot{\theta}\right)' &= 0\\ \left(\frac{1}{4}ml^2\dot{\alpha}\right)' &= 0.\end{aligned} \tag{4.144}$$

This is surprising seeming. Is this correct?

Answer for Exercise 4.8

I found it helpful to clarify for myself what was meant by $g_{ab}(q^c)$. This is a function of all the generalized coordinates:

$$g_{ab}(q^c) = g_{ab}(q^1, q^2, \ldots, q^N) = g_{ab}(\mathbf{q}). \tag{4.146}$$

So I think that a vector parameter reminder is helpful.

$$L = \frac{1}{2}\sum g_{bc}(\mathbf{q})\dot{q}^b\dot{q}^c, \tag{4.147}$$

$$\frac{\partial L}{\partial q^a} = \frac{1}{2}\sum \dot{q}^b\dot{q}^c\frac{\partial g_{bc}(\mathbf{q})}{\partial q^a}. \tag{4.148}$$

Now, proceed to calculate the generalize momentums:

$$\begin{aligned}\frac{\partial L}{\partial \dot{q}^a} &= \frac{1}{2}\sum g_{bc}(\mathbf{q})\frac{\partial\left(\dot{q}^b\dot{q}^c\right)}{\partial \dot{q}^a}\\ &= \frac{1}{2}\sum g_{ac}(\mathbf{q})\dot{q}^c + g_{ba}(\mathbf{q})\dot{q}^b\\ &= \sum g_{ab}(\mathbf{q})\dot{q}^b.\end{aligned} \tag{4.149}$$

For

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} = \sum \frac{\partial g_{ab}}{\partial q^d}\dot{q}^d\dot{q}^b + g_{ba}\ddot{q}^b. \tag{4.150}$$

Taking the difference of eq. (4.148) and eq. (4.150) we have:

$$\begin{aligned}
0 &= \sum \frac{1}{2}\dot{q}^b\dot{q}^c\frac{\partial g_{bc}}{\partial q^a} - \frac{\partial g_{ab}}{\partial q^d}\dot{q}^d\dot{q}^b - g_{ba}\ddot{q}^b \\
&= \sum \dot{q}^b\dot{q}^c\left(\frac{1}{2}\frac{\partial g_{bc}}{\partial q^a} - \frac{\partial g_{ab}}{\partial q^c}\right) - g_{ba}\ddot{q}^b \\
&= \sum \dot{q}^b\dot{q}^c\left(-\frac{1}{2}\frac{\partial g_{bc}}{\partial q^a} + \frac{1}{2}\frac{\partial g_{ab}}{\partial q^c} + \frac{1}{2}\frac{\partial g_{ab}}{\partial q^c}\right) + g_{ba}\ddot{q}^b \\
&= \sum \frac{1}{2}\dot{q}^b\dot{q}^c\left(-\frac{\partial g_{bc}}{\partial q^a} + \frac{\partial g_{ab}}{\partial q^c} + \frac{\partial g_{ac}}{\partial q^b}\right) + g_{ba}\ddot{q}^b.
\end{aligned} \tag{4.151}$$

Here a split of the symmetric expression

$$X = \sum \dot{q}^b\dot{q}^c\frac{\partial g_{ab}}{\partial q^c} = \frac{1}{2}(X + X), \tag{4.152}$$

was used, and then an interchange of dummy indices b, c. Now multiply this whole sum by g^{ba}, and sum to remove the metric term from the generalized acceleration

$$\begin{aligned}
\sum g^{da}g_{ba}\ddot{q}^b &= -\frac{1}{2}\sum \dot{q}^b\dot{q}^c g^{da}\left(-\frac{\partial g_{bc}}{\partial q^a} + \frac{\partial g_{ab}}{\partial q^c} + \frac{\partial g_{ac}}{\partial q^b}\right) \\
\sum \delta^d{}_b\ddot{q}^b &= \\
\ddot{q}^d &=
\end{aligned} \tag{4.153}$$

Swapping a, and d indices to get form stated in the problem we have

$$\begin{aligned}
0 &= \ddot{q}^a + \frac{1}{2}\sum \dot{q}^b\dot{q}^c g^{ad}\left(-\frac{\partial g_{bc}}{\partial q^d} + \frac{\partial g_{db}}{\partial q^c} + \frac{\partial g_{dc}}{\partial q^b}\right) \\
&= \ddot{q}^a + \sum \dot{q}^b\dot{q}^c\Gamma^a{}_{bc} \\
\Gamma^a{}_{bc} &= \frac{1}{2}g^{ad}\left(-\frac{\partial g_{bc}}{\partial q^d} + \frac{\partial g_{db}}{\partial q^c} + \frac{\partial g_{dc}}{\partial q^b}\right).
\end{aligned} \tag{4.154}$$

Answer for Exercise 4.9

Digging in

$$\begin{aligned}\frac{\partial L}{\partial x} &= \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} \\ m\dot{x}^2 V_x - 2VV_x &= \frac{d}{dt}\left(\frac{1}{3}m^2\dot{x}^3 + 2m\dot{x}V\right).\end{aligned} \tag{4.156}$$

When taking the time derivative of V, $dV/dt \neq 0$, despite no explicit time dependence. Take an example, such as $V = mgx$, where the positional parameter is dependent on time, so the chain rule is required:

$$\frac{dV}{dt} = \frac{dV}{dx}\frac{dx}{dt} = \dot{x}V_x. \tag{4.157}$$

Perhaps that is obvious, but I made that mistake first doing this problem (which would have been harder to make if I had used an example potential) the first time. I subsequently constructed an alternate Lagrangian $\left(L = \frac{1}{12}m^2\dot{x}^4 - m\dot{x}^2V + V^2\right)$ that worked when this mistake was made, and emailed the author suggesting that I believed he had a sign typo in his problem set. Anyways, continuing with the calculation:

$$\begin{aligned}m\dot{x}^2 V_x - 2VV_x &= m^2\dot{x}^2\ddot{x} + 2m\ddot{x}V + 2m\dot{x}^2V_x \\ m\dot{x}^2 V_x - 2VV_x - 2m\dot{x}^2V_x &= m\ddot{x}\left(m\dot{x}^2 + 2V\right) \\ -\left(2V + m\dot{x}^2\right)V_x &=\end{aligned} \tag{4.158}$$

Canceling left and right common factors, which perhaps not coincidentally equal $2E = V + \frac{1}{2}mv^2$ we have:

$$m\ddot{x} = -V_x. \tag{4.159}$$

This is what we would get for our standard kinetic and position dependent Lagrangian too:

$$L = \frac{1}{2}m\dot{x}^2 - V. \tag{4.160}$$

$$\begin{aligned}\frac{\partial L}{\partial x} &= \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} \\ -V_x &= \frac{d(m\dot{x})}{dt} \\ -V_x &= m\ddot{x}.\end{aligned} \tag{4.161}$$

Answer for Exercise 4.10

The first thing to observe here is that for $|\mathbf{v}| << c$, this is our familiar kinetic energy Lagrangian

$$\begin{aligned} L &= -mc^2\left(1 - \frac{1}{2}\mathbf{v}^2/c^2 + \frac{1}{2}\frac{1}{-2}\frac{1}{2!}(\mathbf{v}/c)^4 + \cdots\right) - V(\mathbf{r}) \\ &\approx -mc^2 + \frac{1}{2}m\mathbf{v}^2 - V(\mathbf{r}). \end{aligned} \tag{4.163}$$

The constant term $-mc^2$ will not change the equations of motion and we can perhaps think of this as an additional potential term (quite large as we see from atomic fusion and fission). For small $\mathbf{v}$ we recover the Newtonian Kinetic energy term, and therefore expect the results will be equivalent to the Newtonian equations in that limit. Moving on to the calculations we have:

$$\begin{aligned} \frac{\partial L}{\partial x^i} &= \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} \\ -\frac{\partial V}{\partial x^i} &= -c^2\frac{d}{dt}m\frac{\partial L}{\partial \dot{x}^i}\sqrt{1 - \sum\left(\dot{x}^j\right)^2/c^2} \\ &= -c^2\frac{d}{dt}m\frac{1}{2}\frac{1}{\sqrt{1-\mathbf{v}^2/c^2}}\frac{\partial L}{\partial \dot{x}^i}\left(1 - \sum\left(\dot{x}^j\right)^2/c^2\right) \\ &= -c^2\frac{d}{dt}m\frac{1}{2}\frac{1}{\sqrt{1-\mathbf{v}^2/c^2}}(-2)\dot{x}^i/c^2 \\ &= \frac{d}{dt}m\frac{1}{\sqrt{1-\mathbf{v}^2/c^2}}\dot{x}^i \\ &= \frac{d}{dt}m\gamma\dot{x}^i \\ \implies & \\ -\left(\sum \mathbf{e}_i\frac{\partial}{\partial x^i}\right)V &= \frac{d}{dt}m\gamma\sum \mathbf{e}_i\dot{x}^i \\ -\nabla V &= \frac{d(m\gamma\mathbf{v})}{dt}. \end{aligned} \tag{4.164}$$

For $v << c$, $gamma \approx 1$, so we get our Newtonian result in the limiting case. Now, I found this result very impressive result, buried in a couple line problem statement. I subsequently used this as the starting point for guessing about how to formulate the Lagrange equations in a proper time form, as well as a proper velocity form for this Kinetic and potential term. Those turn out to make it possible to express Maxwell's law and the

Lorentz force law together in a particularly nice compact covariant form. This catches me a up a bit in terms of my understanding and think that I am now at least learning and rediscovering things known since the early 1900s;)

Answer for Exercise 4.11

First consider a single pendulum (fixed length l).

$$\begin{aligned} x &= l\exp(i\theta) \\ \dot{x} &= li\dot{\theta}\exp(i\theta) \\ |\dot{x}|^2 &= l^2\dot{\theta}^2. \end{aligned} \tag{4.165}$$

Now, if $\theta = 0$ represents the downwards position at rest, the height above that rest point is $h = l - l\cos\theta$. Therefore the Lagrangian is:

$$\begin{aligned} L &= \frac{1}{2}mv^2 - mgh \\ &= \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1-\cos\theta). \end{aligned} \tag{4.166}$$

The constant term can be dropped resulting in the equivalent Lagrangian:

$$L' = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta. \tag{4.167}$$

This amounts to a difference in the reference point for the potential energy, so instead of measuring the potential energy $V = mgh$ from a reference position below the mass, one could consider that the potential has a maximum of zero at the highest position, and decreases from there as:

$$V' = 0 - mgl\cos\theta. \tag{4.168}$$

Moving back to the EOMs that result from either form of Lagrangian, we have after taking our derivatives:

$$-mgl\sin\theta = \frac{d}{dt}ml^2\dot{\theta} = ml^2\ddot{\theta}. \tag{4.169}$$

Dividing out the ml^2 we are left with

$$\ddot{\theta} = -g/l\sin\theta. \tag{4.170}$$

This is consistent with our expectations, and recovers the familiar small angle SHM equation:

$$\ddot{\theta} \approx -g/l\theta. \tag{4.171}$$

Now, move on to the double pendulum, and compute the Kinetic energies of the two particles:

$$\begin{aligned} x_1 &= l_1 \exp(i\theta_1) \\ \dot{x}_1 &= l_1 i\dot{\theta}_1 \exp(i\theta_1) \\ |\dot{x}_1|^2 &= {l_1}^2\dot{\theta}_1^2, \end{aligned} \tag{4.172}$$

$$\begin{aligned} x_2 &= x_1 + l_2 \exp(i\theta_2) \\ \dot{x}_2 &= \dot{x}_1 + l_2 i\dot{\theta}_2 \exp(i\theta_2) \\ &= l_1 i\dot{\theta}_1 \exp(i\theta_1) + l_2 i\dot{\theta}_2 \exp(i\theta_2) \\ |\dot{x}_2|^2 &= (l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + l_1 i\dot{\theta}_1 \exp(i\theta_1) l_2(-i)\dot{\theta}_2 \exp(-i\theta_2) \\ &\quad + l_1(-i)\dot{\theta}_1 \exp(-i\theta_1) l_2 i\dot{\theta}_2 \exp(i\theta_2) \\ &= (l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + l_1 l_2 \dot{\theta}_1\dot{\theta}_2 \left(\exp(i(\theta_1-\theta_2)) + \exp(-i(\theta_1-\theta_2))\right) \\ &= (l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + 2 l_1 l_2 \dot{\theta}_1\dot{\theta}_2 \cos(\theta_1-\theta_2). \end{aligned} \tag{4.173}$$

Now calculate the potential energies for the two masses. The first has potential of

$$V_1 = m_1 g l_1 (1 - \cos\theta_1). \tag{4.174}$$

and the potential energy of the second mass, relative to the position of the first mass is:

$$V_2' = m_2 g l_2 (1 - \cos\theta_2). \tag{4.175}$$

But that is the potential only if the first mass is at rest. The total difference in height from the dual rest position is:

$$l_1(1 - \cos\theta_1) + l_2(1 - \cos\theta_2). \tag{4.176}$$

So, the potential energy for the second mass is:

$$V_2 = m_2 g \left(l_1(1 - \cos\theta_1) + l_2(1 - \cos\theta_2)\right). \tag{4.177}$$

Dropping constant terms the total Lagrangian for the system is:

$$\begin{aligned} L &= \frac{1}{2}m_1{v_1}^2 + \frac{1}{2}m_2{v_2}^2 + m_1 g l_1 \cos\theta_1 + m_2 g\left(l_1\cos\theta_1 + l_2\cos\theta_2\right) \\ &= \frac{1}{2}m_1{l_1}^2\dot{\theta}_1^2 + \frac{1}{2}m_2\left((l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + 2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1-\theta_2)\right) \\ &+ m_1 g l_1\cos\theta_1 + m_2 g\left(l_1\cos\theta_1 + l_2\cos\theta_2\right). \end{aligned} \tag{4.178}$$

Again looking at the resulting Lagrangian, we see that it would have been more natural to measure the potential energy from a reference point of zero potential at the horizontal position, and measure downwards from there:

$$\begin{aligned} V_1' &= 0 - m_1 g l_1\cos\theta_1 \\ V_2' &= 0 - m_2 g\left(l_1\cos\theta_1 + l_2\cos\theta_2\right). \end{aligned} \tag{4.179}$$

N coupled pendulums. Now, with just two masses it is not too messy to expand out those kinetic energy terms, but for more the trig gets too messy. With the K_2 term of the Lagrangian in complex form we have:

$$\begin{aligned} L &= \frac{1}{2}m_1{l_1}^2\dot{\theta}_1^2 + \frac{1}{2}m_2\left|l_1\dot{\theta}_1 + l_2\dot{\theta}_2\exp(i(\theta_2-\theta_1))\right|^2 \\ &+ m_1 g l_1\cos\theta_1 + m_2 g\left(l_1\cos\theta_1 + l_2\cos\theta_2\right) \\ &= \frac{1}{2}m_1\left|l_1\dot{\theta}_1\exp(i\theta_1)\right|^2 + \frac{1}{2}m_2\left|l_1\dot{\theta}_1\exp(i\theta_1) + l_2\dot{\theta}_2\exp(i\theta_2)\right|^2 \\ &+ m_1 g l_1\cos\theta_1 + m_2 g\left(l_1\cos\theta_1 + l_2\cos\theta_2\right). \end{aligned} \tag{4.180}$$

By inspection we can also write the Lagrangian for the N particle variant:

$$L = \frac{1}{2}\sum_{j=1}^{N} m_j\left|\sum_{k=1}^{j} l_k\dot{\theta}_k\exp(i\theta_k)\right|^2 + g\sum_{j=1}^{N} l_j\cos\theta_j\sum_{k=j}^{N} m_k. \tag{4.181}$$

Can this be used to derive the wave equation? If each of the N masses is a fraction $m_j = \Delta m = M/N$ of the total mass, and the lengths are

all uniformly divided into segments of length $l_j = \Delta l = L/N$, then the Lagrangian becomes:

$$\begin{aligned} L &= \frac{\Delta l}{2g} \sum_{j=1}^{N} \left| \sum_{k=1}^{j} \dot{\theta}_k \exp(i\theta_k) \right|^2 + \sum_{j=1}^{N} \cos\theta_j \sum_{k=j}^{N} 1 \\ &= \frac{\Delta l}{2g} \sum_{j=1}^{N} \left| \sum_{k=1}^{j} \dot{\theta}_k \exp(i\theta_k) \right|^2 + (N - j + 1) \sum_{j=1}^{N} \cos\theta_j. \end{aligned} \tag{4.182}$$

FIXME: return to this later?

Double pendulum. First consider a single pendulum (fixed length l).

$$\begin{aligned} x &= l \exp(i\theta) \\ \dot{x} &= li\dot{\theta} \exp(i\theta) \\ |\dot{x}|^2 &= l^2\dot{\theta}^2. \end{aligned} \tag{4.183}$$

Now, if $\theta = 0$ represents the downwards position at rest, the height above that rest point is $h = l - l\cos\theta$. Therefore the Lagrangian is:

$$\begin{aligned} L &= \frac{1}{2}mv^2 - mgh \\ &= \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos\theta). \end{aligned} \tag{4.184}$$

The constant term can be dropped resulting in the equivalent Lagrangian:

$$L' = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta. \tag{4.185}$$

This amounts to a difference in the reference point for the potential energy, so instead of measuring the potential energy $V = mgh$ from a reference position below the mass, one could consider that the potential has a maximum of zero at the highest position, and decreases from there as:

$$V' = 0 - mgl\cos\theta. \tag{4.186}$$

Moving back to the EOMs that result from either form of Lagrangian, we have after taking our derivatives:

$$-mgl\sin\theta = \frac{d}{dt}ml^2\dot{\theta} = ml^2\ddot{\theta}. \tag{4.187}$$

Dividing out the ml^2 we are left with

$$\ddot{\theta} = -g/l \sin\theta. \tag{4.188}$$

This is consistent with our expectations, and recovers the familiar small angle SHM equation:

$$\ddot{\theta} \approx -g/l\theta. \tag{4.189}$$

Now, move on to the double pendulum, and compute the Kinetic energies of the two particles:

$$\begin{aligned} x_1 &= l_1 \exp(i\theta_1) \\ \dot{x}_1 &= l_1 i\dot{\theta}_1 \exp(i\theta_1) \\ |\dot{x}_1|^2 &= {l_1}^2\dot{\theta}_1^2. \end{aligned} \tag{4.190}$$

$$\begin{aligned} x_2 &= x_1 + l_2 \exp(i\theta_2) \\ \dot{x}_2 &= \dot{x}_1 + l_2 i\dot{\theta}_2 \exp(i\theta_2) \\ &= l_1 i\dot{\theta}_1 \exp(i\theta_1) + l_2 i\dot{\theta}_2 \exp(i\theta_2) \\ |\dot{x}_2|^2 &= (l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + l_1 i\dot{\theta}_1 \exp(i\theta_1) l_2(-i)\dot{\theta}_2 \exp(-i\theta_2) \\ &\quad + l_1(-i)\dot{\theta}_1 \exp(-i\theta_1) l_2 i\dot{\theta}_2 \exp(i\theta_2) \\ &= (l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + l_1 l_2 \dot{\theta}_1\dot{\theta}_2 \left(\exp(i(\theta_1-\theta_2)) + \exp(-i(\theta_1-\theta_2))\right) \\ &= (l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + 2 l_1 l_2 \dot{\theta}_1\dot{\theta}_2 \cos(\theta_1-\theta_2). \end{aligned} \tag{4.191}$$

Now calculate the potential energies for the two masses. The first has potential of

$$V_1 = m_1 g l_1 (1 - \cos\theta_1). \tag{4.192}$$

and the potential energy of the second mass, relative to the position of the first mass is:

$$V_2' = m_2 g l_2 (1 - \cos\theta_2). \tag{4.193}$$

But that is the potential only if the first mass is at rest. The total difference in height from the dual rest position is:

$$l_1(1 - \cos\theta_1) + l_2(1 - \cos\theta_2). \tag{4.194}$$

So, the potential energy for the second mass is:

$$V_2 = m_2 g \left(l_1(1 - \cos\theta_1) + l_2(1 - \cos\theta_2) \right). \tag{4.195}$$

Dropping constant terms the total Lagrangian for the system is:

$$\begin{aligned} L &= \frac{1}{2} m_1 {v_1}^2 + \frac{1}{2} m_2 {v_2}^2 + m_1 g l_1 \cos\theta_1 + m_2 g \left(l_1 \cos\theta_1 + l_2 \cos\theta_2 \right) \\ &= \frac{1}{2} m_1 {l_1}^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 \left((l_1 \dot{\theta}_1)^2 + (l_2 \dot{\theta}_2)^2 + 2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2) \right) \\ &+ m_1 g l_1 \cos\theta_1 + m_2 g \left(l_1 \cos\theta_1 + l_2 \cos\theta_2 \right). \end{aligned} \tag{4.196}$$

Again looking at the resulting Lagrangian, we see that it would have been more natural to measure the potential energy from a reference point of zero potential at the horizontal position, and measure downwards from there:

$$\begin{aligned} V_1' &= 0 - m_1 g l_1 \cos\theta_1 \\ V_2' &= 0 - m_2 g \left(l_1 \cos\theta_1 + l_2 \cos\theta_2 \right). \end{aligned} \tag{4.197}$$

Answer for Exercise 4.12

Cylindrical Polar Coordinates. The next two parts of question 6 require cylindrical polar coordinates. I found a digression was useful (or at least interesting), to see if the gradient followed from the Lagrangian as was the case with non-orthonormal constant frame basis vectors. The first step required for this calculation (and the later parts of the problem) is to express the KE in terms of the polar coordinates. We need the velocity to do so:

$$\begin{aligned} \mathbf{r} &= \mathbf{e}_3 z + \mathbf{e}_1 r e^{i\theta} \\ \dot{\mathbf{r}} &= \mathbf{e}_3 \dot{z} + \mathbf{e}_1 (\dot{r} + r\dot{\theta} i) e^{i\theta} \\ |\dot{\mathbf{r}}| &= \dot{z}^2 + \left| \dot{r} + r\dot{\theta} i \right|^2 \\ &= \dot{z}^2 + \dot{r}^2 + \left(r\dot{\theta} \right)^2. \end{aligned} \tag{4.198}$$

Now, form the Lagrangian of a point particle with a non-velocity dependent potential:

$$L = \frac{1}{2} m \left(\dot{z}^2 + \dot{r}^2 + \left(r\dot{\theta} \right)^2 \right) - \phi. \tag{4.199}$$

and calculate the equations of motion:

$$\begin{aligned}\frac{\partial L}{\partial z} &= \left(\frac{\partial L}{\partial \dot{z}}\right)' \\ -\frac{\partial \phi}{\partial z} &= (m\dot{z})'\,.\end{aligned} \tag{4.200}$$

$$\begin{aligned}\frac{\partial L}{\partial r} &= \left(\frac{\partial L}{\partial \dot{r}}\right)' \\ -\frac{\partial \phi}{\partial r} + mr\dot{\theta}^2 &= (m\dot{r})'\,.\end{aligned} \tag{4.201}$$

$$\begin{aligned}\frac{\partial L}{\partial \theta} &= \left(\frac{\partial L}{\partial \dot{\theta}}\right)' \\ -\frac{\partial \phi}{\partial \theta} &= \left(mr^2\dot{\theta}\right)'\,.\end{aligned} \tag{4.202}$$

There are a few things to observe about these equations. One is that we can assign physically significance to an expression such as $mr^2\dot{\theta}$. If the potential has no θ dependence this is a conserved quantity (angular momentum). The other thing to observe here is that the dimensions for the θ coordinate equation result has got an extra length factor in the numerator. Thus we can not multiply these with our respective frame vectors and sum. We can however scale that last equation by a factor of $1/r$ and then sum:

$$\hat{\mathbf{z}}(m\dot{z})' + \hat{\mathbf{r}}\left((m\dot{r})' - mr\dot{\theta}^2\right) + \frac{1}{r}(mr^2\dot{\theta})' = -\left(\hat{\mathbf{z}}\frac{\partial}{\partial z} + \hat{\mathbf{r}}\frac{\partial}{\partial r} + \frac{1}{r}\frac{\partial}{\partial \theta}\right)\phi. \tag{4.203}$$

For constant mass this is:

$$m\left(\hat{\mathbf{z}}\ddot{z} + \hat{\mathbf{r}}\left(\ddot{r} - r\dot{\theta}^2\right) + \frac{1}{r}\left(2r\dot{r}\dot{\theta} + r^2\ddot{\theta}\right)\right) = -\left(\hat{\mathbf{z}}\frac{\partial}{\partial z} + \hat{\mathbf{r}}\frac{\partial}{\partial r} + \frac{1}{r}\hat{\boldsymbol{\theta}}\frac{\partial}{\partial \theta}\right)\phi. \tag{4.204}$$

However, is such a construction have a meaningful physical quantity? One can easily imagine more complex generalized coordinates where guessing scale factors in this fashion would not be possible. Let us compare this to a calculation of acceleration in cylindrical coordinates.

$$\begin{aligned}\ddot{\mathbf{r}} &= \mathbf{e}_3\ddot{z} + \mathbf{e}_1\left(\ddot{r} + r\ddot{\theta}i + \dot{r}\dot{\theta}i + \left(\dot{r} + r\dot{\theta}i\right)i\dot{\theta}\right)e^{i\theta} \\ &= \mathbf{e}_3\ddot{z} + \mathbf{e}_1\left(\ddot{r} + r\ddot{\theta}i + 2\dot{r}\dot{\theta}i - r\dot{\theta}^2\right)e^{i\theta} \\ &= \hat{\mathbf{z}}\ddot{z} + \hat{\mathbf{r}}\left(\ddot{r} - r\dot{\theta}^2\right) + \hat{\boldsymbol{\theta}}\left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right).\end{aligned} \tag{4.205}$$

Sure enough, the ad-hoc vector that was constructed matches the acceleration vector for the constant mass case, so the right hand side must also define the gradient in cylindrical coordinates.

$$\begin{aligned}\nabla &= \hat{\mathbf{z}}\frac{\partial}{\partial z} + \hat{\mathbf{r}}\frac{\partial}{\partial r} + \frac{1}{r}\hat{\boldsymbol{\theta}}\frac{\partial}{\partial \theta} \\ &= \hat{\mathbf{z}}\frac{\partial}{\partial z} + \hat{\mathbf{r}}\left(\frac{\partial}{\partial r} + \frac{i}{r}\frac{\partial}{\partial \theta}\right).\end{aligned} \tag{4.206}$$

Very cool result. Seeing this I finally understand when and where statements like "angular momentum is conserved" is true. Specifically it requires a potential that has no angular dependence (ie: like gravity acting between two point masses.) I never found that making such an angular momentum conservation "law" statement to be obvious, even once the acceleration was expressed in a radial decomposition. This is something that can be understood without the Lagrangian formulation. To do so the missing factor is that before a conservation statement like this can be claimed one has to first express the gradient in cylindrical form, and then look at the coordinates with respect to the generalized frame vectors. Conservation of angular momentum depends on an appropriately well behaved potential function! Intuitively, I understood that something else was required to make this statement, but it took the form of an unproven axiom in most elementary texts. FIXME: generalize this and prove to myself that angular momentum is conserved in a N-body problem and/or with a rigid body rotation constraint on $N-1$ of the masses.

(i). The Lagrangian for this problem is:

$$L = \frac{1}{2}m\mathbf{v}^2 - e\mathbf{A}\cdot\mathbf{v}. \tag{4.207}$$

Given a cylindrical decomposition, our velocity is:

$$\begin{aligned}\mathbf{r} &= z\hat{\mathbf{z}} + r\hat{\mathbf{r}} \\ \dot{\mathbf{r}} &= \dot{z}\hat{\mathbf{z}} + r\dot{\hat{\mathbf{r}}} + \dot{r}\hat{\mathbf{r}} \\ &= \dot{z}\hat{\mathbf{z}} + \hat{\mathbf{r}}\left(\dot{r} + r\dot{\theta}i\right) \\ &= \dot{z}\hat{\mathbf{z}} + \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}.\end{aligned} \tag{4.208}$$

The specific potential for the problem, using (z, θ, r) coordinates is:

$$\mathbf{A} = \hat{\boldsymbol{\theta}}\frac{f(r)}{r}. \tag{4.209}$$

Therefore the Lagrangian is:

$$L = \frac{1}{2}m\left(\dot{z}^2 + \dot{r}^2 + (r\dot{\theta})^2\right) - e\frac{f(r)}{r}r\dot{\theta}. \tag{4.210}$$

so the equations of motion for the z, θ, and r coordinates (respectively) are:

$$\begin{aligned} (m\dot{z})' &= 0 \\ (mr^2\dot{\theta} - ef(r))' &= 0 \\ (m\dot{r})' &= mr\dot{\theta}^2 - ef'(r)\dot{\theta}. \end{aligned} \tag{4.211}$$

From second of these equations we have:

$$mr^2\dot{\theta} - ef(r) = K. \tag{4.212}$$

In particular this is true for $r = r(t_0) = r_0$, so

$$m{r_0}^2\dot{\theta}_0 - ef(r_0) = K. \tag{4.213}$$

Equating K's and rearranging, we have

$$\dot{\theta}(t) - \left(\frac{r_0}{r}\right)^2\dot{\theta}(t_0) = \frac{e}{mr^2}\left(f(r) - f(r_0)\right). \tag{4.214}$$

Now, the problem is to show that

$$\dot{\theta} = \frac{e}{mr^2}\left(f(r) - f(r_0)\right). \tag{4.215}$$

I do not see how that follows? Ah, I see, the velocity is in the (r, z) plane for $t = 0$, so $\dot{\theta}(t_0) = 0$.

(ii). The potential for this problem is

$$\mathbf{A} = rg(z)\hat{\boldsymbol{\theta}}. \tag{4.216}$$

Therefore the Lagrangian is:

$$L = \frac{1}{2}m\left(\dot{z}^2 + \dot{r}^2 + (r\dot{\theta})^2\right) - er^2g(z)\dot{\theta}. \tag{4.217}$$

Taking θ, r, z derivatives:

$$0 = \left(mr^2\dot{\theta} - er^2g(z)\right)' \tag{4.218}$$

$$mr\dot{\theta}^2 - 2erg(z)\dot{\theta} = (m\dot{r})' \tag{4.219}$$

$$-er^2 g'\dot{\theta} = (m\dot{z})'. \tag{4.220}$$

One constant of motion is:

$$mr^2\dot{\theta} - er^2 g(z) = K. \tag{4.221}$$

Looking at Tong's solutions another is the Hamiltonian.

$$\dot{\theta} = (e/m)g(z) + (r_0/r)^2\left(\dot{\theta}_0 - (e/m)g(z_0)\right). \tag{4.222}$$

With $\dot{\theta}_0 = 2eg(z_0)/m$ this is:

$$\dot{\theta} = \frac{e}{m}\left(g(z) + \left(\frac{r_0}{r}\right)^2 g(z_0)\right). \tag{4.223}$$

Answer for Exercise 4.13

Jackson gives a tip to use the convective derivative (yet another name for the chain rule), and using this in the Euler-Lagrange equations we have

$$\boldsymbol{\nabla} L = \frac{d}{dt}\boldsymbol{\nabla}_{\mathbf{u}} L = \left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\boldsymbol{\nabla}\right)\mathbf{e}_a \frac{\partial L}{\partial \dot{x}^a}. \tag{4.225}$$

where $\{\mathbf{e}_a\}$ is the spatial basis. The first order of business is calculating the gradient and conjugate momenta. For the latter we have

$$\begin{aligned}\mathbf{e}_a \frac{\partial L}{\partial \dot{x}^a} &= \mathbf{e}_a\left(-mc^2\gamma\frac{1}{2}(-2)\dot{x}^a/c^2 + \frac{e}{c}A^a\right) \\ &= m\gamma\mathbf{u} + \frac{e}{c}\mathbf{A} \\ &\equiv \mathbf{p} + \frac{e}{c}\mathbf{A}.\end{aligned} \tag{4.226}$$

Applying the convective derivative we have

$$\frac{d}{dt}\mathbf{e}_a \frac{\partial L}{\partial \dot{x}^a} = \frac{d\mathbf{p}}{dt} + \frac{e}{c}\frac{\partial \mathbf{A}}{\partial t} + \frac{e}{c}\mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{A}. \tag{4.227}$$

For the gradient we have

$$\mathbf{e}_a \frac{\partial L}{\partial x^a} = e\left(\frac{1}{c}\dot{x}^b \boldsymbol{\nabla} A^b - \boldsymbol{\nabla}\phi\right). \tag{4.228}$$

Rearranging eq. (4.225) for this Lagrangian we have

$$\frac{d\mathbf{p}}{dt} = e\left(-\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \frac{1}{c}\mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{A} + \frac{1}{c}\dot{x}^b \boldsymbol{\nabla} A^b\right). \tag{4.229}$$

The first two terms are the electric field

$$\mathbf{E} \equiv -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}. \tag{4.230}$$

So it remains to be shown that the remaining two equal $(\mathbf{u}/c)\times\mathbf{B} = (\mathbf{u}/c)\times(\boldsymbol{\nabla}\times\mathbf{A})$. Using the Hestenes notation using primes to denote what the gradient is operating on, we have

$$\begin{aligned}
\dot{x}^b \boldsymbol{\nabla} A^b - \mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{A} &= \boldsymbol{\nabla}'\mathbf{u}\cdot\mathbf{A}' - \mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{A} \\
&= -\mathbf{u}\cdot(\boldsymbol{\nabla}\wedge\mathbf{A}) \\
&= \frac{1}{2}\left((\boldsymbol{\nabla}\wedge\mathbf{A})\mathbf{u} - \mathbf{u}(\boldsymbol{\nabla}\wedge\mathbf{A})\right) \\
&= \frac{I}{2}\left((\boldsymbol{\nabla}\times\mathbf{A})\mathbf{u} - \mathbf{u}(\boldsymbol{\nabla}\times\mathbf{A})\right) \\
&= -I(\mathbf{u}\wedge\mathbf{B}) \\
&= -II(\mathbf{u}\times\mathbf{B}) \\
&= \mathbf{u}\times\mathbf{B}.
\end{aligned} \tag{4.231}$$

I have used the geometric algebra identities I am familiar with to regroup things, but this last bit can likely be done with index manipulation too. The exercise is complete, and we have from the Lagrangian

$$\frac{d\mathbf{p}}{dt} = e\left(\mathbf{E} + \frac{1}{c}\mathbf{u}\times\mathbf{B}\right). \tag{4.232}$$

Answer for Exercise 4.14

Let's put the electric field in the $\hat{\mathbf{x}}$ direction ($\theta = 0$), so that the potential acting on charge i is given implicitly by

$$\mathbf{F}_i = q_i\mathcal{E}\hat{\mathbf{x}} = -\nabla\phi_i = -\hat{\mathbf{x}}\frac{d\phi_i}{dx}. \tag{4.234}$$

or

$$\phi_i = -q_i(x_i - x_0). \tag{4.235}$$

Our positions, and velocities are

$$\mathbf{r}_{1,2} = \pm\frac{r}{2}\hat{\mathbf{x}}e^{\hat{\mathbf{x}}\hat{\mathbf{y}}\theta}. \tag{4.236a}$$

$$\frac{d\mathbf{r}_{1,2}}{dt} = \pm\frac{r}{2}\dot{\theta}\hat{\mathbf{y}}e^{\hat{\mathbf{x}}\hat{\mathbf{y}}\theta}. \tag{4.236b}$$

Our kinetic energy is

$$\begin{aligned} T &= \frac{1}{2}\sum_i m_i\left(\frac{d\mathbf{r}_i}{dt}\right)^2 \\ &= \frac{1}{2}\sum_i m_i\left(\frac{r}{2}\right)^2\dot{\theta}^2 \\ &= \frac{1}{2}(m_1+m_2)\left(\frac{r}{2}\right)^2\dot{\theta}^2. \end{aligned} \tag{4.237}$$

For our potential energies we require the x component of the position vectors, which are

$$\begin{aligned} x_i &= \mathbf{r}_i\cdot\hat{\mathbf{x}} \\ &= \pm\left\langle\frac{r}{2}\hat{\mathbf{x}}e^{\hat{\mathbf{x}}\hat{\mathbf{y}}\theta}\hat{\mathbf{x}}\right\rangle \\ &= \pm\frac{r}{2}\cos\theta. \end{aligned} \tag{4.238}$$

Our potentials are

$$\phi_1 = -q_1\mathcal{E}\frac{r}{2}\cos\theta + \phi_0. \tag{4.239a}$$

$$\phi_2 = q_2\mathcal{E}\frac{r}{2}\cos\theta + \phi_0. \tag{4.239b}$$

Our system Lagrangian, after dropping the constant reference potential that doesn't effect the dynamics is

$$L = \frac{1}{2}(m_1+m_2)\left(\frac{r}{2}\right)^2\dot{\theta}^2 + q_1\mathcal{E}\frac{r}{2}\cos\theta - q_2\mathcal{E}\frac{r}{2}\cos\theta. \tag{4.240}$$

For this problem we had two equal masses and equal magnitude charges $m = m_1 = m_2$ and $q = q_1 = -q_2$

$$L = \frac{1}{4}mr^2\dot{\theta}^2 + qr\mathcal{E}\cos\theta. \tag{4.241}$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = \frac{1}{2}mr^2\dot{\theta}. \tag{4.242}$$

$$\begin{aligned} \frac{\partial L}{\partial \theta} = -qr\mathcal{E}\sin\theta &= \frac{dp_\theta}{dt} \\ &= \frac{1}{2}mr^2\ddot{\theta}. \end{aligned} \tag{4.243}$$

Putting these together, with $p = qr$, we have the result stated in the text

$$F_\theta = \frac{dp_\theta}{dt} = -p\mathcal{E}\sin\theta. \tag{4.244}$$

5 CONSTRAINTS.

Exercise 5.1 Pendulum on a rotating wheel. (*[24] p5.*)

Lagrangian and equations of motion for pendulum with pivot moving on a circle.

Exercise 5.2 Two circular constrained paths. (*[24] p7*)

Masses connected by a spring.

Exercise 5.3 Masses on string, one dangling. (*[24] p8*)

Two particles connected by string, one on table, the other dangling.

Exercise 5.4 Pendulum with support moving in circle. (*[16] p1.3.*)

Attempting a mechanics problem from Landau I get a different answer. I wrote up my solution to see if I can spot either where I went wrong, or demonstrate the error, and then posted it to physicsforums . I wasn't wrong, but the text wasn't either. The complete result is given below, where the problem (§1 problem 3a) of [16] is to calculate the Lagrangian of a pendulum where the point of support is moving in a circle (figure and full text for problem in this Google books reference)

Exercise 5.5 Pendulum with support moving in line. (*[16] p1.3b.*)

This problem like the last, but with the point of suspension moving in a horizontal line $x = a \cos \gamma t$.

Exercise 5.6 Pendulum with support moving in vertical line. (*[16] p1.3c.*)

As above, but with the support point moving up and down as $a \cos \gamma t$.

Exercise 5.7 Coupled hoop and spring system.

Find the Langrangian for the system sketched in fig. 5.1, where one mass is connected between two springs to a bar. That bar moves up and down as forced by the motion of the other mass along a immovable hoop.

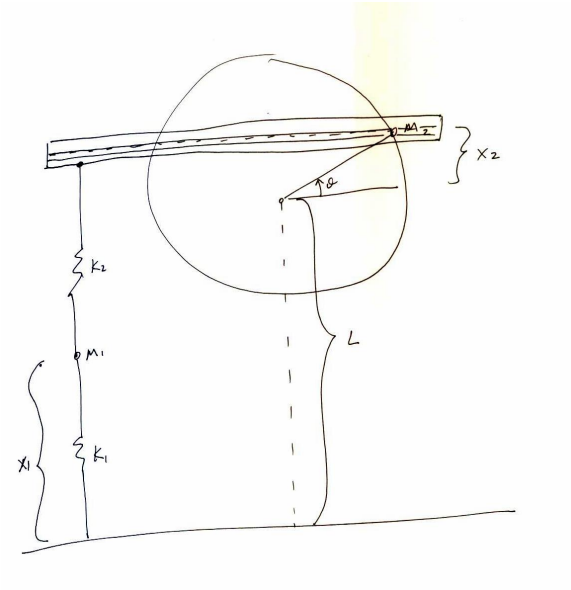


Figure 5.1: Coupled hoop and spring system.

5.1 SOLUTIONS.

Answer for Exercise 5.1

Express the position of the pivot point on the wheel with:

$$q_1 = Re^{-i\omega t}. \tag{5.1}$$

The position of the mass is then:

$$q_2 = Re^{-i\omega t} - ile^{i\theta}. \tag{5.2}$$

The velocity of the mass is then:

$$\dot{q}_2 = -i(\dot{\omega}t + \omega)Re^{-i\omega t} + l\dot{\theta}e^{i\theta}. \tag{5.3}$$

Let $\omega t = \alpha$, we have a Kinetic energy of:

$$\begin{aligned}
\frac{1}{2}m|\dot{q}_2|^2 &= \frac{1}{2}m\left|-i\dot{\alpha}Re^{-i\omega t} + l\dot{\theta}e^{i\theta}\right|^2 \\
&= \frac{1}{2}m\left(R^2\dot{\alpha}^2 + l^2\dot{\theta}^2 + 2Rl\dot{\alpha}\dot{\theta}\,\mathrm{Re}\left(-ie^{-i\alpha-i\theta}\right)\right) \\
&= \frac{1}{2}m\left(R^2\dot{\alpha}^2 + l^2\dot{\theta}^2 + 2Rl\dot{\alpha}\dot{\theta}\cos(-\alpha-\theta-\pi/2)\right) \\
&= \frac{1}{2}m\left(R^2\dot{\alpha}^2 + l^2\dot{\theta}^2 - 2Rl\dot{\alpha}\dot{\theta}\sin(\alpha+\theta)\right).
\end{aligned} \tag{5.4}$$

The potential energy in the Lagrangian does not depend on the position of the pivot, only the angle from vertical, so it is thus:

$$\begin{aligned}
V &= mgl(1-\cos\theta) \\
V' &= 0 - mgl\cos\theta.
\end{aligned} \tag{5.5}$$

Depending on whether one measures the potential up from the lowest potential point, or measures decreasing potential from zero at the horizontal. Either way, combining the kinetic and potential terms, and dividing through by ml^2 we have the Lagrangian of:

$$L = \frac{1}{2}\left((R/l)^2\dot{\alpha}^2 + \dot{\theta}^2 - 2(R/l)\dot{\alpha}\dot{\theta}\sin(\alpha+\theta)\right) + (g/l)\cos\theta. \tag{5.6}$$

Digression. Reduction of the Lagrangian. Now, in Tong's solutions for this problem (which he emailed me since I questioned problem 2), he had $\dot{\alpha} = \omega =$ constant, which allows the Lagrangian above to be expressed as:

$$L = \frac{1}{2}\left((R/l)^2\omega^2 + \dot{\theta}^2\right) + \frac{d}{dt}((R/l)\cos(\omega t+\theta)) + \omega(R/l)\sin(\omega t+\theta) + (g/l)\cos\theta. \tag{5.7}$$

and he made the surprising step of removing that cosine term completely, with a statement that it would not effect the dynamics because it was a time derivative. That turns out to be a generalized result, but I had to prove it to myself. I also asked around on PF about this, and it was not any named property of Lagrangians, but was a theorem in some texts. First consider the simple example of a Lagrangian with such a cosine derivative term added to it:

$$L' = L + \frac{d}{dt}A\cos(\omega t+\theta). \tag{5.8}$$

and compute the equations of motion from this:

$$
\begin{aligned}
0 &= \frac{\partial L'}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{\theta}}\right) \\
&= \frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) + \frac{\partial}{\partial \theta}\frac{d}{dt}A\cos(\omega t + \theta) - \frac{d}{dt}\frac{\partial}{\partial \dot{\theta}}\frac{d}{dt}A\cos(\omega t + \theta) \\
&= \frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial}{\partial \theta}A\dot{\theta}\sin(\omega t + \theta) + \frac{d}{dt}\frac{\partial}{\partial \dot{\theta}}A\dot{\theta}\sin(\omega t + \theta) \\
&= \frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial}{\partial \theta}A\dot{\theta}\sin(\omega t + \theta) + \frac{d}{dt}A\sin(\omega t + \theta) \\
&= \frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - A\dot{\theta}\cos(\omega t + \theta) + A\dot{\theta}\cos(\omega t + \theta) \\
&= \frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right).
\end{aligned}
\tag{5.9}
$$

Now consider the general case, altering a Lagrangian by adding the time derivative of a positional dependent function:

$$
L' = L + \frac{df}{dt}. \tag{5.10}
$$

and compute the equations of motion from this more generally altered function:

$$
\begin{aligned}
0 &= \frac{\partial L'}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L'}{\partial q^i}\right) \\
&= \frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial q^i}\right) + \frac{\partial}{\partial q^i}\frac{df}{dt} - \frac{d}{dt}\frac{\partial}{\partial q^i}\frac{df}{dt}.
\end{aligned}
\tag{5.11}
$$

Now, if $f(q^j, \dot{q}^j, t) = f(q^j, t)$ we have:

$$
\frac{df}{dt} = \sum \frac{\partial f}{\partial q^j}\dot{q}^j + \frac{\partial f}{\partial t}. \tag{5.12}
$$

We want to see if the following sums to zero:

$$
\begin{aligned}
\frac{\partial}{\partial q^i}\frac{df}{dt} - \frac{d}{dt}\frac{\partial}{\partial q^i}\frac{df}{dt} &= \sum \frac{\partial}{\partial q^i}\frac{\partial f}{\partial q^j}\left(\dot{q}^j + \frac{\partial f}{\partial t}\right) - \frac{d}{dt}\frac{\partial}{\partial \dot{q}^i}\left(\sum \frac{\partial f}{\partial q^j}\dot{q}^j + \frac{\partial f}{\partial t}\right) \\
&= \sum \frac{\partial^2 f}{\partial q^i \partial q^j}\dot{q}^j + \frac{\partial^2 f}{\partial q^i \partial t} - \frac{d}{dt}\left(\sum \delta_{ij}\frac{\partial f}{\partial q^j} + \frac{\partial^2 f}{\partial \dot{q}^i \partial t}\right) \\
&= \sum \frac{\partial^2 f}{\partial q^i \partial q^j}\dot{q}^j + \frac{\partial^2 f}{\partial q^i \partial t} - \frac{d}{dt}\frac{\partial f}{\partial q^i} \\
&= \sum \frac{\partial^2 f}{\partial q^i \partial q^j}\dot{q}^j + \frac{\partial^2 f}{\partial q^i \partial t} - \sum \dot{q}^j \frac{\partial^2 f}{\partial q^j \partial q^i} - \frac{\partial^2 f}{\partial t \partial q^i}.
\end{aligned}
$$

(5.13)

Therefore provided the function is sufficiently continuous that all mixed pairs of mixed partials are equal, this is zero, and the df/dt addition does not change the equations of motion that the Lagrangian generates.

Back to the problem. Now, return to the Lagrangian for this problem, and compute the equations of motion. Writing $\mu = R/l$, we have:

$$L = \frac{1}{2}\left(\mu^2\dot{\alpha}^2 + \dot{\theta}^2 - 2\mu\dot{\alpha}\dot{\theta}\sin(\alpha+\theta)\right) + (g/l)\cos\theta. \tag{5.14}$$

$$\begin{aligned} 0 &= \frac{d}{dt}\frac{\partial -}{\partial\dot{\theta}}\frac{\partial L}{\partial\theta} \\ &= \frac{d}{dt}\left(\dot{\theta} - \mu\dot{\alpha}\sin(\alpha+\theta)\right) + \mu\dot{\alpha}\dot{\theta}\cos(\alpha+\theta) + (g/l)\sin\theta \\ &= \ddot{\theta} - \mu\ddot{\alpha}\sin(\alpha+\theta) - \mu\dot{\alpha}\cos(\alpha+\theta)(\dot{\alpha}+\dot{\theta}) + \mu\dot{\alpha}\dot{\theta}\cos(\alpha+\theta) + (g/l)\sin\theta. \end{aligned} \tag{5.15}$$

Sure enough we have a cancellation of terms for constant ω. In general we are left with:

$$\ddot{\theta} = \mu\ddot{\alpha}\sin(\alpha+\theta) + \mu\dot{\alpha}^2\cos(\alpha+\theta) - (g/l)\sin\theta. \tag{5.16}$$

This expands to

$$\ddot{\theta} = \mu(\ddot{\omega}t + 2\dot{\omega})\sin(\omega t+\theta) + \mu(\dot{\omega}t+\omega)^2\cos(\omega t+\theta) - (g/l)\sin\theta. \tag{5.17}$$

For constant ω, this is just:

$$\ddot{\theta} = \mu\omega^2\cos(\omega t+\theta) - (g/l)\sin\theta. \tag{5.18}$$

Answer for Exercise 5.2

With $i = \mathbf{e}_1 \wedge \mathbf{e}_2$, the paths, (squared) speeds and separation of the masses can be written:

$$\begin{aligned}
q_1 &= \mathbf{e}_1 R_1 e^{i\theta} \\
q_2 &= c\mathbf{e}_3 + \mathbf{e}_1 \left(ai + R_2 e^{i\alpha}\right) \\
|\dot{q}_1|^2 &= \left(R_1\dot{\theta}\right)^2 \\
|\dot{q}_2|^2 &= \left(R_2\dot{\alpha}\right)^2 \\
d^2 &= \left(q_1 - q_2\right)^2 \\
&= c^2 + \left|ai + R_2 e^{i\alpha} - R_1 e^{i\theta}\right|^2 \\
&= c^2 + a^2 + {R_2}^2 + {R_1}^2 + ai\left(R_2 e^{-i\alpha} - R_1 e^{-i\theta} - R_2 e^{i\alpha} + R_1 e^{i\theta}\right) \\
&\quad - R_1 R_2 \left(e^{i\alpha} e^{-i\theta} + e^{-i\alpha} e^{i\theta}\right) \\
&= c^2 + a^2 + {R_2}^2 + {R_1}^2 + 2a(R_2 \sin\alpha - R_1 \sin\theta) - 2R_1 R_2 \cos(\alpha - \theta).
\end{aligned} \tag{5.19}$$

With the given potential:

$$V = \frac{1}{2}\omega^2 d^2. \tag{5.20}$$

We have the following Lagrangian (where the constant terms in the separation have been dropped) :

$$\begin{aligned}
L = &\frac{1}{2}m_1(R_1\dot{\theta})^2 + \frac{1}{2}m_2(R_2\dot{\alpha})^2 \\
&+ \omega^2\left(a(R_2 \sin\alpha - R_1 \sin\theta) - R_1 R_2 \cos(\alpha - \theta)\right).
\end{aligned} \tag{5.21}$$

The last part of the problem was to show that there is an additional conserved quantity when $a = 0$. The Lagrangian in that case is:

$$L = \frac{1}{2}m_1\left(R_1\dot{\theta}\right)^2 + \frac{1}{2}m_2\left(R_2\dot{\alpha}\right)^2 - R_1 R_2 \omega^2 \cos(\alpha - \theta). \tag{5.22}$$

Evaluating the Lagrange equations, for this condition one has:

$$\begin{aligned}
-R_1 R_2 \omega^2 \sin(\alpha - \theta) &= \left(m_1 {R_1}^2 \dot{\theta}\right)' \\
R_1 R_2 \omega^2 \sin(\alpha - \theta) &= \left(m_2 {R_2}^2 \dot{\alpha}\right)'.
\end{aligned} \tag{5.23}$$

Summing these one has:

$$\left(m_1 {R_1}^2 \dot{\theta}\right)' + \left(m_2 {R_2}^2 \dot{\alpha}\right)' = 0. \tag{5.24}$$

Therefore the additional conserved quantity is:

$$m_1 {R_1}^2 \dot{\theta} + m_2 {R_2}^2 \dot{\alpha} = K. \tag{5.25}$$

FIXME: Is there a way to identify such a conserved quantity without evaluating the derivatives? Noether's?

Spring Potential? Small digression. Let us take the gradient of this spring potential and see if this matches our expectations for a $-kx$ spring force.

$$-\nabla_d V = -\omega^2 d\hat{\mathbf{d}} = -\omega^2 \mathbf{d}. \tag{5.26}$$

Okay, this works, $\omega^2 = k$, which just expresses the positiveness of this constant.

Answer for Exercise 5.3

Part (i). The second particle hangs straight down (also Goldstein problem 9, also example 2.3 in Hestenes NFCM.) First mass m_1 on the table, and second, hanging. The kinetic term for the mass on the table was calculated above in problem 7, so adding that and the KE term for the dangling mass we have:

$$K = \frac{1}{2} m_1 \left(\dot{r}^2 + (r\dot{\psi})^2 \right) + \frac{1}{2} m_2 \dot{r}^2. \tag{5.27}$$

Our potential, measuring down is:

$$V = 0 - m_2 g (l - r). \tag{5.28}$$

Combining the KE and PE terms and dropping constant terms we have:

$$L = \frac{1}{2} m_1 \left(\dot{r}^2 + (r\dot{\psi})^2 \right) + \frac{1}{2} m_2 \dot{r}^2 - m_2 g r. \tag{5.29}$$

The ignorable coordinate is ψ since it has only derivatives in the Lagrangian. EOMs are:

$$\begin{aligned} 0 &= \left(m_1 r^2 \dot{\psi} \right)' \\ m_1 r \dot{\psi}^2 - m_2 g &= \left((m_1 + m_2) \dot{r} \right)' = M \ddot{r}. \end{aligned} \tag{5.30}$$

The first equation here is a conservation of angular momentum statement, whereas the second equation has all the force terms that lie along the string

(radially above the table, and downwards below). We see the $r\dot{\psi}^2 = r\omega^2$ angular acceleration component when calculating radial and non-radial component of acceleration. Goldstein asks here for the equations of motion as a second order equation, and to integrate once. We can go all the way, but only implicitly, as we can write $t = t(r)$, using $\dot{r}$ as an integrating factor:

$$m_1 r^2 \dot{\psi} = m_1 {r_0}^2 \omega_0 \tag{5.31a}$$

$$\dot{\psi} = \left(\frac{r_0}{r}\right)^2 \omega_0 \tag{5.31b}$$

$$m_1 \frac{{r_0}^4}{r^3} {\omega_0}^2 - m_2 g = M\ddot{r} \tag{5.31c}$$

$$\begin{aligned} m_1 \dot{r} \frac{{r_0}^4}{r^3} {\omega_0}^2 - m_2 g \dot{r} &= M\dot{r}\ddot{r} \\ -m_1 {r_0}^2 \left(\frac{1}{r^2}\right)' {\omega_0}^2 - m_2 g \dot{r} &= M\left(\dot{r}^2\right)' \\ K - m_1 {r_0}^4 \frac{1}{r^2} {\omega_0}^2 - m_2 g r &= M\dot{r}^2 \end{aligned} \tag{5.31d}$$

$$K = m_1 {r_0}^2 {\omega_0}^2 + m_2 g r_0 + M\dot{r}_0^2. \tag{5.31e}$$

$$m_1 {\omega_0}^2 {r_0}^2 \left(1 - \frac{{r_0}^2}{r^2}\right) + M\dot{r}_0^2 - m_2 g \left(r - r_0\right) = M\dot{r}^2 \tag{5.31f}$$

$$t = \int \frac{dr}{\sqrt{\frac{m_1}{M} {\omega_0}^2 {r_0}^2 \left(1 - \frac{{r_0}^2}{r^2}\right) + \dot{r}_0^2 - \frac{m_2}{M} g \left(r - r_0\right)}}. \tag{5.31g}$$

We can also write $\psi = \psi(r)$, but that does not look like it is any easier to solve:

$$\begin{aligned}
\dot{\psi} &= \frac{d\psi}{dr}\frac{dr}{dt} \\
&\Longrightarrow \\
\frac{d\psi}{dr} &= \frac{dt}{dr}\left(\frac{r_0}{r}\right)^2 \omega_0 \\
\psi &= \int \frac{{r_0}^2 \omega_0 dr}{r^2 \sqrt{\frac{m_1}{M}{\omega_0}^2 {r_0}^2 \left(1 - \frac{{r_0}^2}{r^2}\right) + \dot{r}_0^2 - \frac{m_2}{M} g\left(r - r_0\right)}}.
\end{aligned} \tag{5.32}$$

(ii). Motion of dangling mass not restricted to straight down. This part of the problem treats the dangling mass as a spherical pendulum. If θ is the angle from the vertical and α is the angle in the horizontal plane of motion, we can describe the coordinate of the dangler (pointing $\hat{\mathbf{z}} = \hat{\mathbf{g}}$ downwards), as:

$$q_2 = R(\sin\theta\cos\alpha, \sin\theta\sin\alpha, \cos\theta). \tag{5.33}$$

and the velocity as:

$$\begin{aligned}
\dot{q}_2 &= \dot{R}(\sin\theta\cos\alpha, \sin\theta\sin\alpha, \cos\theta) \\
&\quad + R(\cos\theta\cos\alpha, \cos\theta\sin\alpha, -\sin\theta)\dot{\theta} \\
&\quad + R(-\sin\theta\sin\alpha, \sin\theta\cos\alpha, 0)\dot{\alpha}.
\end{aligned} \tag{5.34}$$

and can then attempt to square this mess to get the squared speed that we need for the kinetic energy term of the Lagrangian. Instead, lets choose an alternate parametrization:

$$\begin{aligned}
q_2 &= R\cos\theta\hat{\mathbf{z}} + \mathbf{e}_1 R\sin\theta e^{i\alpha} \\
\dot{q}_2 &= \left(\dot{R}\cos\theta - R\sin\theta\dot{\theta}\right)\hat{\mathbf{z}} + \mathbf{e}_1 e^{i\alpha}\left(\dot{R}\sin\theta + R\cos\theta\dot{\theta} + R\sin\theta i\dot{\alpha}\right) \\
|\dot{q}_2|^2 &= \left(\dot{R}\cos\theta - R\sin\theta\dot{\theta}\right)^2 + \left(\dot{R}\sin\theta + R\cos\theta\dot{\theta}\right)^2 + (R\sin\theta\dot{\alpha})^2 \\
&= \dot{R}^2 + (R\dot{\theta})^2 + (R\sin\theta\dot{\alpha})^2.
\end{aligned} \tag{5.35}$$

Our potential is

$$V = 0 - m_2 g(l - r)\cos\theta, \tag{5.36}$$

so, the Lagrangian is therefore:

$$L = \frac{1}{2}m_2\left(\dot{r}^2 + (l-r)^2\left(\dot{\theta}^2 + \sin\theta\dot{\alpha}\right)^2\right) + \frac{1}{2}m_1\left(\dot{r}^2 + (r\dot{\psi})^2\right) + m_2 g(l-r)\cos\theta. \tag{5.37}$$

Answer for Exercise 5.4

The coordinates of the mass are

$$p = ae^{i\gamma t} + ile^{i\phi}, \tag{5.38}$$

or in coordinates

$$p = (a\cos\gamma t + l\sin\phi, -a\sin\gamma t + l\cos\phi). \tag{5.39}$$

The velocity is

$$\dot{p} = (-a\gamma\sin\gamma t + l\dot{\phi}\cos\phi, -a\gamma\cos\gamma t - l\dot{\phi}\sin\phi), \tag{5.40}$$

and in the square

$$\begin{aligned}\dot{p}^2 &= a^2\gamma^2 + l^2\dot{\phi}^2 - 2a\gamma\dot{\phi}\sin\gamma t\cos\phi + 2a\gamma l\dot{\phi}\cos\gamma t\sin\phi \\ &= a^2\gamma^2 + l^2\dot{\phi}^2 + 2a\gamma l\dot{\phi}\sin(\gamma t - \phi).\end{aligned} \tag{5.41}$$

For the potential our height above the minimum is

$$h = 2a + l - a(1-\cos\gamma t) - l\cos\phi = a(1+\cos\gamma t) + l(1-\cos\phi). \tag{5.42}$$

In the potential the total derivative $\cos\gamma t$ can be dropped, as can all the constant terms, leaving

$$U = -mgl\cos\phi, \tag{5.43}$$

so by the above the Lagrangian should be (after also dropping the constant term $ma^2\gamma^2/2$

$$L = \frac{1}{2}m\left(l^2\dot{\phi}^2 + 2a\gamma l\dot{\phi}\sin(\gamma t - \phi)\right) + mgl\cos\phi. \tag{5.44}$$

This is almost the stated value in the text

$$L = \frac{1}{2}m\left(l^2\dot{\phi}^2 + 2a\gamma^2 l\sin(\gamma t - \phi)\right) + mgl\cos\phi. \tag{5.45}$$

We have what appears to be an innocent looking typo (text putting in a γ instead of a $\dot{\phi}$), but the subsequent text also didn't make sense. That referred to the omission of the total derivative $mla\gamma\cos(\phi - \gamma t)$, which isn't even a term that I have in my result.

In the physicsforums response it was cleverly pointed out by Dickfore that eq. (5.44) can be recast into a total derivative

$$\begin{aligned} & mal\gamma\dot{\phi}\sin(\gamma t - \phi) \\ & \quad = mal\gamma(\dot{\phi} - \gamma)\sin(\gamma t - \phi) + mal\gamma^2\sin(\gamma t - \phi) \\ & \quad = \frac{d}{dt}\left(mal\gamma\cos(\gamma t - \phi)\right) + mal\gamma^2\sin(\gamma t - \phi), \end{aligned} \tag{5.46}$$

which resolves the conundrum!

Answer for Exercise 5.5

Our mass point has coordinates

$$\begin{aligned} p &= a\cos\gamma t + lie^{-i\phi} \\ &= a\cos\gamma t + li(\cos\phi - i\sin\phi) \\ &= (a\cos\gamma t + l\sin\phi, l\cos\phi), \end{aligned} \tag{5.47}$$

so that the velocity is

$$\dot{p} = (-a\gamma\sin\gamma t + l\dot{\phi}\cos\phi, -l\dot{\phi}\sin\phi). \tag{5.48}$$

Our squared velocity is

$$\begin{aligned} \dot{p}^2 &= a^2\gamma^2\sin^2\gamma t + l^2\dot{\phi}^2 - 2a\gamma l\dot{\phi}\sin\gamma t\cos\phi \\ &= \frac{1}{2}a^2\gamma^2\frac{d}{dt}\left(t - \frac{1}{2\gamma}\sin 2\gamma t\right) + l^2\dot{\phi}^2 - a\gamma l\dot{\phi}(\sin(\gamma t + \phi) + \sin(\gamma t - \phi)). \end{aligned} \tag{5.49}$$

In the last term, we can reduce the sum of sines, finding a total derivative term and a remainder as in the previous problem. That is

$$\begin{aligned} & \dot{\phi}(\sin(\gamma t + \phi) + \sin(\gamma t - \phi)) \\ & \quad = (\dot{\phi} + \gamma)\sin(\gamma t + \phi) - \gamma\sin(\gamma t + \phi) + (\dot{\phi} - \gamma)\sin(\gamma t - \phi) + \gamma\sin(\gamma t - \phi) \\ & \quad = \frac{d}{dt}\left(-\cos(\gamma t + \phi) + \cos(\gamma t - \phi)\right) + \gamma(\sin(\gamma t - \phi) - \sin(\gamma t + \phi)) \\ & \quad = \frac{d}{dt}\left(-\cos(\gamma t + \phi) + \cos(\gamma t - \phi)\right) - 2\gamma\cos\gamma t\sin\phi. \end{aligned}$$

(5.50)

Putting all the pieces together and dropping the total derivatives we have the stated solution

$$L = \frac{1}{2}m\left(l^2\dot{\phi}^2 + 2a\gamma^2 l\cos\gamma t\sin\phi\right) + mgl\cos\phi. \tag{5.51}$$

Answer for Exercise 5.6

Our mass point is

$$p = a\cos\gamma t + le^{i\phi}. \tag{5.52}$$

with velocity

$$\begin{aligned}\dot{p} &= -a\gamma\sin\gamma t + li\dot{\phi}e^{i\phi}\\ &= (-a\gamma\sin\gamma t - l\dot{\phi}\sin\phi, l\dot{\phi}\cos\phi).\end{aligned} \tag{5.53}$$

In the square this is

$$|\dot{p}|^2 = a^2\gamma^2\sin^2\gamma t + l^2\dot{\phi}^2\sin^2\phi + 2al\gamma\dot{\phi}\sin\gamma t\sin\phi. \tag{5.54}$$

Having done the simplification in the last problem in a complicated way, let's try it, knowing what our answer is

$$\begin{aligned}\dot{\phi}\sin\gamma t\sin\phi &= \dot{\phi}\sin\gamma t\sin\phi - \gamma\cos\gamma t\cos\phi + \gamma\cos\gamma t\cos\phi\\ &= \sin\gamma t\frac{d}{dt}(-\cos\phi) + \left(\frac{d}{dt}(-\sin\gamma t)\right)\cos\phi + \gamma\cos\gamma t\cos\phi\\ &= \gamma\cos\gamma t\cos\phi - \frac{d}{dt}(\sin\gamma t\cos\phi).\end{aligned} \tag{5.55}$$

With the height of the particle above the lowest point given by

$$h = a + l - a\cos\gamma t - l\cos\phi, \tag{5.56}$$

we can write the Lagrangian immediately (dropping all the total derivative terms)

$$L = \frac{1}{2}m\left(l^2\dot{\phi}^2\sin^2\phi + 2al\gamma^2\cos\gamma t\cos\phi\right) + mgl\cos\phi. \tag{5.57}$$

Answer for Exercise 5.7

The Lagrangian can be written by inspection. Writing $x = x_1$, and $x_2 = R\sin\theta$, we have

$$L = \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2R^2\dot{\theta}^2 - \frac{1}{2}k_1x^2 - \frac{1}{2}k_2(L+R\sin\theta - x)^2 - m_1gx - m_2g(L+R\sin\theta). \tag{5.58}$$

Evaluation of the Euler-Lagrange equations gives

$$\begin{aligned} m_1\ddot{x} &= -k_1x + k_2(L+R\sin\theta - x) - m_1g \\ m_2R^2\ddot{\theta} &= -k_2(L+R\sin\theta - x)R\cos\theta - m_2gR\cos\theta, \end{aligned} \tag{5.59}$$

or

$$\begin{aligned} \ddot{x} &= -x\frac{k_1+k_2}{m_1} + \frac{k_2R\sin\theta}{m_1} - g + \frac{k_2L}{m_1} \\ \ddot{\theta} &= -\frac{1}{R}\left(\frac{k_2}{m_2}\left(L+R\sin\theta - x\right) + g\right)\cos\theta. \end{aligned} \tag{5.60}$$

Just like any other coupled pendulum system, this one is non-linear. There is no obvious way to solve this in closed form, but we could determine a solution in the neighborhood of a point $(x,\theta) = (x_0,\theta_0)$. Let us switch our dynamical variables to ones that express the deviation from the initial point $\delta x = x - x_0$, and $\delta\theta = \theta - \theta_0$, with $u = (\delta x)'$, and $v = (\delta\theta)'$. Our system then takes the form

$$\begin{aligned} u' &= f(x,\theta) = -x\frac{k_1+k_2}{m_1} + \frac{k_2R\sin\theta}{m_1} - g + \frac{k_2L}{m_1} \\ v' &= g(x,\theta) = -\frac{1}{R}\left(\frac{k_2}{m_2}\left(L+R\sin\theta - x\right) + g\right)\cos\theta \\ (\delta x)' &= u \\ (\delta\theta)' &= v. \end{aligned} \tag{5.61}$$

We can use a first order Taylor approximation of the form $f(x,\theta) = f(x_0,\theta_0) + f_x(x_0,\theta_0)(\delta x) + f_\theta(x_0,\theta_0)(\delta\theta)$. So, to first order, our system has the approximation

$$\begin{aligned}
u' &= -x_0\frac{k_1+k_2}{m_1} + \frac{k_2 R\sin\theta_0}{m_1} \\
&\quad - g + \frac{k_2 L}{m_1} - (\delta x)\frac{k_1+k_2}{m_1} + \frac{k_2 R\cos\theta_0}{m_1}(\delta\theta) \\
v' &= -\frac{1}{R}\left(\frac{k_2}{m_2}\left(L + R\sin\theta_0 - x_0\right) + g\right)\cos\theta_0 + \frac{k_2\cos\theta_0}{m_2 R}(\delta x) \\
&\quad - \frac{1}{R}\left(\frac{k_2}{m_2}\left(\left(L - x_0\right)\sin\theta_0 + R\right) + g\sin\theta_0\right)(\delta\theta) \\
(\delta x)' &= u \\
(\delta\theta)' &= v.
\end{aligned} \tag{5.62}$$

This would be tidier in matrix form with $\mathbf{x} = (u, v, \delta x, \delta\theta)$

$$\begin{aligned}
\mathbf{x}' &= \begin{bmatrix} -x_0\frac{k_1+k_2}{m_1} + \frac{k_2 R\sin\theta_0}{m_1} - g + \frac{k_2 L}{m_1} \\ -\frac{1}{R}\left(\frac{k_2}{m_2}\left(L + R\sin\theta_0 - x_0\right) + g\right)\cos\theta_0 \\ 0 \\ 0 \end{bmatrix} \\
&\quad + \begin{bmatrix} 0 & 0 & -\frac{k_1+k_2}{m_1} & \frac{k_2 R\cos\theta_0}{m_1} \\ 0 & 0 & \frac{k_2\cos\theta_0}{m_2 R} & -\frac{1}{R}\left(\frac{k_2}{m_2}\left(\left(L - x_0\right)\sin\theta_0 + R\right) + g\sin\theta_0\right) \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\mathbf{x}.
\end{aligned} \tag{5.63}$$

This reduces the problem to the solutions of first order equations of the form

$$\begin{aligned}
\mathbf{x}' &= \mathbf{a} + \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix} \\
\mathbf{x} &= \mathbf{a} + \mathbf{B}\mathbf{x},
\end{aligned} \tag{5.64}$$

where $\mathbf{a}$, and A are constant matrices. Such a matrix equation has the solution

$$\mathbf{x} = e^{Bt}\mathbf{x}_0 + (e^{Bt} - I)B^{-1}\mathbf{a}, \tag{5.65}$$

but the zeros in B should allow the exponential and inverse to be calculated with less work. That inverse is readily verified to be

$$B^{-1} = \begin{bmatrix} 0 & I \\ A^{-1} & 0 \end{bmatrix}. \tag{5.66}$$

It is also not hard to show that

$$\begin{aligned} B^{2n} &= \begin{bmatrix} A^n & 0 \\ 0 & A^n \end{bmatrix} \\ B^{2n+1} &= \begin{bmatrix} 0 & A^{n+1} \\ A^n & 0 \end{bmatrix}. \end{aligned} \tag{5.67}$$

Together this allows for the power series expansion

$$e^{Bt} = \begin{bmatrix} \cosh(t\sqrt{A}) & \sinh(t\sqrt{A}) \\ \sinh(t\sqrt{A})\frac{1}{\sqrt{A}} & \cosh(t\sqrt{A}) \end{bmatrix}. \tag{5.68}$$

All of the remaining sub matrix expansions should be straightforward to calculate provided the eigenvalues and vectors of A are calculated. Specifically, suppose that we have

$$A = U \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U^{-1}. \tag{5.69}$$

Then all the perhaps non-obvious functions of matrices expand to just

$$\begin{aligned} A^{-1} &= U \begin{bmatrix} \lambda_1^{-1} & 0 \\ 0 & \lambda_2^{-1} \end{bmatrix} U^{-1} \\ \sqrt{A} &= U \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix} U^{-1} \\ \cosh(t\sqrt{A}) &= U \begin{bmatrix} \cosh(t\sqrt{\lambda_1}) & 0 \\ 0 & \cosh(t\sqrt{\lambda_2}) \end{bmatrix} U^{-1} \\ \sinh(t\sqrt{A}) &= U \begin{bmatrix} \sinh(t\sqrt{\lambda_1}) & 0 \\ 0 & \sinh(t\sqrt{\lambda_2}) \end{bmatrix} U^{-1} \\ \sinh(t\sqrt{A})\frac{1}{\sqrt{A}} &= U \begin{bmatrix} \sinh(t\sqrt{\lambda_1})/\sqrt{\lambda_1} & 0 \\ 0 & \sinh(t\sqrt{\lambda_2})/\sqrt{\lambda_2} \end{bmatrix} U^{-1}. \end{aligned} \tag{5.70}$$

6 SPACE TIME ALGEBRA (STA.)

6.1 OVERVIEW.

The STA and geometric algebra ideas used here are not complete to learn from in isolation. The reader is referred to [2] for a more complete exposition of both STA and geometric algebra.

6.1.1 *Conventions.*

Definition 6.1: Index conventions.

Latin indexes $i, j, k, r, s, t, \cdots$ are used to designate values in the range $\{1, 2, 3\}$. Greek indexes are $\alpha, \beta, \mu, \nu, \cdots$ are used for indexes of space-time quantities $\{0, 1, 2, 3\}$. The Einstein convention of implied summation for mixed upper and lower Greek indexes will be used, for example

$$x^\alpha x_\alpha \equiv \sum_{\alpha=0}^{3} x^\alpha x_\alpha.$$

6.1.2 *Space Time Algebra (STA.)*

In the geometric algebra literature, the Dirac algebra of quantum field theory has been rebranded Space Time Algebra (STA). The differences between STA and the Dirac theory that uses matrices (γ_μ) are as follows

- STA completely omits any representation of the Dirac basis vectors γ_μ. In particular, any possible matrix representation is irrelevant.
- STA provides a rich set of fundamental operations (grade selection, generalized dot and wedge products for multivector elements, rotation and reflection operations, ...)

- Matrix trace, and commutator and anticommutator operations are nowhere to be found in STA, as geometrically grounded equivalents are available instead.
- The "slashed" quantities from Dirac theory, such as $\not{p} = \gamma_\mu p^\mu$ are nothing more than vectors in their entirety in STA (where the basis is no longer implicit, as is the case for coordinates.)

Our basis vectors have the following properties.

Definition 6.2: Standard basis.

Let the four-vector standard basis be designated $\{\gamma_0, \gamma_1, \gamma_2, \gamma_3\}$, where the basis vectors satisfy $\gamma_0^2 = -\gamma_i^2 = 1$, and $\gamma_\alpha \cdot \gamma_\beta = 0, \forall \alpha \neq \beta$.

Exercise 6.1 Commutator properties of the STA basis.

In Dirac theory, the commutator properties of the Dirac matrices is considered fundamental, namely

$$\left\{\gamma_\mu, \gamma_\nu\right\} = 2\eta_{\mu\nu}.$$

Show that this follows from the axiomatic assumptions of geometric algebra, and describe how the dot and wedge products are related to the anticommutator and commutator products of Dirac theory.

Definition 6.3: Pseudoscalar.

The pseudoscalar for the space is denoted $I = \gamma_0\gamma_1\gamma_2\gamma_3$.

Exercise 6.2 Pseudoscalar.

Show that the STA pseudoscalar I defined by definition 6.2 satisfies

$$\tilde{I} = I,$$

where the tilde operator designates reversion. Also show that I has the properties of an imaginary number

$$I^2 = -1.$$

Finally, show that, unlike the spatial pseudoscalar that commutes with all grades, I anticommutes with any vector or trivector, and commutes with any bivector.

Definition 6.4: Reciprocal basis.

The reciprocal basis $\left\{\gamma^0, \gamma^1, \gamma^2, \gamma^3\right\}$ is defined , such that the property $\gamma^\alpha \cdot \gamma_\beta = \delta^\alpha{}_\beta$ holds.

Observe that, $\gamma^0 = \gamma_0$ and $\gamma^i = -\gamma_i$.

Theorem 6.1: Coordinates.

Coordinates are defined in terms of dot products with the standard basis, or reciprocal basis

$$\begin{aligned} x^\alpha &= x \cdot \gamma^\alpha \\ x_\alpha &= x \cdot \gamma_\alpha, \end{aligned}$$

Proof. Suppose that a coordinate representation of the following form is assumed

$$x = x^\alpha \gamma_\alpha = x_\beta \gamma^\beta. \tag{6.12}$$

We wish to determine the representation of the x^α or x_β coordinates in terms of x and the basis elements. Taking the dot product with any standard basis element, we find

$$\begin{aligned} x \cdot \gamma_\mu &= (x_\beta \gamma^\beta) \cdot \gamma_\mu \\ &= x_\beta \delta^\beta{}_\mu \\ &= x_\mu, \end{aligned} \tag{6.13}$$

as claimed. Similarly, dotting with a reciprocal frame vector, we find

$$\begin{aligned} x \cdot \gamma^\mu &= (x^\beta \gamma_\beta) \cdot \gamma^\mu \\ &= x^\beta \delta_\beta{}^\mu \\ &= x^\mu. \end{aligned} \tag{6.14}$$

□

Observe that raising or lowering the index of a spatial index toggles the sign of a coordinate, but timelike indexes are left unchanged.

$$\begin{aligned} x^0 &= x_0 \\ x^i &= -x_i \end{aligned} \tag{6.15}$$

Definition 6.5: Spacetime gradient.

The spacetime gradient operator is

$$\nabla = \gamma^\mu \partial_\mu = \gamma_\nu \partial^\nu,$$

where

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

and

$$\partial^\mu = \frac{\partial}{\partial x_\mu}.$$

This definition of gradient is consistent with the Dirac gradient (sometimes denoted $\not{\partial}$).

Definition 6.6: Timelike and spacelike components of a four-vector.

Given a four vector $x = \gamma_\mu x^\mu$, that would be designated $x^\mu = \left\{x^0, \mathbf{x}\right\}$ in conventional special relativity, we write

$$x^0 = x \cdot \gamma_0,$$

and

$$\mathbf{x} = x \wedge \gamma_0,$$

or

$$x = (x^0 + \mathbf{x})\gamma_0.$$

The spacetime split of a four-vector x is relative to the frame. In the relativistic lingo, one would say that it is "observer dependent", as the same operations with $\gamma_0{}'$, the timelike basis vector for a different frame, would yield a different set of coordinates.

While the dot and wedge products above provide an effective mechanism to split a four vector into a set of timelike and spacelike quantities,

the spatial component of a vector has a bivector representation in STA. Consider the following coordinate expansion of a spatial vector

$$\begin{aligned}\mathbf{x} &= x \wedge \gamma_0 \\ &= \left(x^\mu \gamma_\mu\right) \wedge \gamma_0 \\ &= \sum_{k=1}^{3} x^k \gamma_k \gamma_0.\end{aligned} \tag{6.16}$$

Definition 6.7: Spatial basis.

We designate $\mathbf{e}_i = \gamma_i \gamma_0$ as the standard basis vectors for $\mathbb{R}^3$.

In the literature, this bivector representation of the spatial basis may be designated $\sigma_i = \gamma_i \gamma_0$, as these bivectors have the properties of the Pauli matrices σ_i. Because this book a number of purely non-relativistic applications too, the Pauli notation will not be used here.

Exercise 6.3 Orthonormality of the spatial basis.

Show that the spatial basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, defined by definition 6.7, is orthonormal.

Exercise 6.4 Spatial pseudoscalar.

Show that the STA pseudoscalar $I = \gamma_0 \gamma_1 \gamma_2 \gamma_3$ equals the spatial pseudoscalar $I = \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3$.

Exercise 6.5 Characteristics of the Pauli matrices.

The Pauli matrices obey the following anticommutation relations:

$$\{\sigma_a, \sigma_b\} = 2\delta_{ab}, \tag{6.19}$$

and commutation relations:

$$\left[\sigma_a, \sigma_b\right] = 2i\epsilon_{abc}\, \sigma_c, \tag{6.20}$$

Show how these relate to the geometric algebra dot and wedge products, and determine the geometric algebra representation of the imaginary i above.

6.2 SOLUTIONS.

Answer for Exercise 6.1

The anticommutator is defined as symmetric sum of products

$$\left\{\gamma_\mu, \gamma_\nu\right\} \equiv \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu, \tag{6.1}$$

but this is just twice the dot product in its geometric algebra form $ab = (ab + ba)/2$. Observe that the properties of the basis vectors defined in definition 6.2 may be summarized as

$$\gamma_\mu \cdot \gamma_\nu = \eta_{\mu\nu}, \tag{6.2}$$

where $\eta_{\mu\nu} = \text{diag}(+,-,-,-) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ is the conventional metric tensor. This means

$$\gamma_\mu \cdot \gamma_\nu = \eta_{\mu\nu} = 2\left\{\gamma_\mu, \gamma_\nu\right\}, \tag{6.3}$$

as claimed.

Similarly, observe that the commutator, defined as the antisymmetric sum of products

$$\left[\gamma_\mu, \gamma_\nu\right] \equiv \gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu, \tag{6.4}$$

is twice the wedge product $a \wedge b = (ab - ba)/2$. This provides geometric identifications for the respective anti-commutator and commutator products respectively

$$\begin{aligned} \left\{\gamma_\mu, \gamma_\nu\right\} &= 2\gamma_\mu \cdot \gamma_\nu \\ \left[\gamma_\mu, \gamma_\nu\right] &= 2\gamma_\mu \wedge \gamma_\nu, \end{aligned} \tag{6.5}$$

Answer for Exercise 6.2

Since $\gamma_\alpha\gamma_\beta = -\gamma_\beta\gamma_\alpha$ for any $\alpha \neq \beta$, any permutation of the factors of I changes the sign once. In particular

$$\begin{aligned} I &= \gamma_0\gamma_1\gamma_2\gamma_3 \\ &= -\gamma_1\gamma_2\gamma_3\gamma_0 \\ &= -\gamma_2\gamma_3\gamma_1\gamma_0 \\ &= +\gamma_3\gamma_2\gamma_1\gamma_0 = \tilde{I}. \end{aligned} \tag{6.6}$$

Using this, we have

$$\begin{aligned} I^2 &= I\tilde{I} \\ &= (\gamma_0\gamma_1\gamma_2\gamma_3)(\gamma_3\gamma_2\gamma_1\gamma_0) \\ &= (\gamma_0)^2(\gamma_1)^2(\gamma_2)^2(\gamma_3)^2 \\ &= (+1)(-1)(-1)(-1) \\ &= -1. \end{aligned} \tag{6.7}$$

To illustrate the anticommutation property with any vector basis element, consider the following two examples:

$$\begin{aligned} I\gamma_0 &= \gamma_0\gamma_1\gamma_2\gamma_3\gamma_0 \\ &= -\gamma_0\gamma_0\gamma_1\gamma_2\gamma_3 \\ &= -\gamma_0 I, \end{aligned} \tag{6.8}$$

$$\begin{aligned} I\gamma_2 &= \gamma_0\gamma_1\gamma_2\gamma_3\gamma_2 \\ &= -\gamma_0\gamma_1\gamma_2\gamma_2\gamma_3 \\ &= -\gamma_2\gamma_0\gamma_1\gamma_2\gamma_3 \\ &= -\gamma_2 I. \end{aligned} \tag{6.9}$$

A total of three sign swaps is required to "percolate" any given γ_α through the factors of I, resulting in an overall sign change of -1.

For any bivector basis element $\alpha \neq \beta$

$$\begin{aligned} I\gamma_\alpha\gamma_\beta &= -\gamma_\alpha I\gamma_\beta \\ &= +\gamma_\alpha\gamma_\beta I. \end{aligned} \tag{6.10}$$

Similarly for any trivector basis element $\alpha \neq \beta \neq \sigma$

$$\begin{aligned} I\gamma_\alpha\gamma_\beta\gamma_\sigma &= -\gamma_\alpha I\gamma_\beta\gamma_\sigma \\ &= +\gamma_\alpha\gamma_\beta I\gamma_\sigma \\ &= -\gamma_\alpha\gamma_\beta\gamma_\sigma I. \end{aligned} \tag{6.11}$$

Answer for Exercise 6.3

$$\begin{aligned} \mathbf{e}_i \cdot \mathbf{e}_j &= \left\langle \gamma_i\gamma_0\gamma_j\gamma_0 \right\rangle \\ &= -\left\langle \gamma_i\gamma_j \right\rangle \\ &= -\gamma_i \cdot \gamma_j. \end{aligned} \tag{6.17}$$

This is zero for all $i \neq j$, and unity for any $i = j$.

Answer for Exercise 6.4

The spatial pseudoscalar, expanded in terms of the STA basis vectors, is

$$\begin{aligned} I &= \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3 \\ &= (\gamma_1\gamma_0)(\gamma_2\gamma_0)(\gamma_3\gamma_0) \\ &= (\gamma_1\gamma_0)\gamma_2(\gamma_0\gamma_3)\gamma_0 \\ &= (-\gamma_0\gamma_1)\gamma_2(-\gamma_3\gamma_0)\gamma_0 \\ &= \gamma_0\gamma_1\gamma_2\gamma_3(\gamma_0\gamma_0) \\ &= \gamma_0\gamma_1\gamma_2\gamma_3, \end{aligned} \tag{6.18}$$

as claimed.

6.3 LORENTZ TRANSFORMATIONS IN STA.

6.3.1 *Motivation.*

One of the remarkable features of geometric algebra are the complex exponential sandwiches that can be used to encode rotations in any dimension, or rotation like operations like Lorentz transformations in Minkowski spaces. In this post, we show some examples that unpack the geometric algebra expressions for Lorentz transformations operations of this sort. In particular, we will look at the exponential sandwich operations for spatial rotations and Lorentz boosts in the Dirac algebra, known as Space Time Algebra (STA) in geometric algebra circles, and demonstrate that these sandwiches do have the desired effects.

6.3.2 *Lorentz transformations.*

Theorem 6.2: Lorentz transformation.

The transformation

$$x \to e^B x e^{-B} = x',$$

where $B = a \wedge b$, is an STA 2-blade for any two linearly independent four-vectors a, b, is a norm preserving, that is

$$x^2 = x'^2.$$

Proof. The proof is disturbingly trivial in this geometric algebra form

$$\begin{aligned} x'^2 &= e^B x e^{-B} e^B x e^{-B} \\ &= e^B x x e^{-B} \\ &= x^2 e^B e^{-B} \\ &= x^2. \end{aligned} \tag{6.21}$$

□

In particular, observe that we did not need to construct the usual infinitesimal representations of rotation and boost transformation matrices or tensors in order to demonstrate that we have spacetime invariance for the transformations. The rough idea of such a transformation is that the exponential commutes with components of the four-vector that lie off the spacetime plane specified by the bivector B, and anticommutes with components of the four-vector that lie in the plane. The end result is that the sandwich operation simplifies to

$$x' = x_{\|} e^{-B} + x_{\perp}, \tag{6.22}$$

where $x = x_{\perp} + x_{\|}$ and $x_{\perp} \cdot B = 0$, and $x_{\|} \wedge B = 0$. In particular, using $x = xBB^{-1} = \left(x \cdot B + x \wedge B\right) B^{-1}$, we find that

$$\begin{aligned} x_{\|} &= \left(x \cdot B\right) B^{-1} \\ x_{\perp} &= \left(x \wedge B\right) B^{-1}. \end{aligned} \tag{6.23}$$

When B is a spacetime plane $B = b \wedge \gamma_0$, then this exponential has a hyperbolic nature, and we end up with a Lorentz boost. When B is a spatial bivector, we end up with a single complex exponential, encoding our plane old 3D rotation. More general B's that encode composite boosts and rotations are also possible, but B must be invertible (it should have no lightlike factors.) The rough geometry of these projections is illustrated in fig. 6.1, where the spacetime plane is represented by B.

What is not so obvious is how to pick B's that correspond to specific rotation axes or boost directions. Let's consider each of those cases in turn.

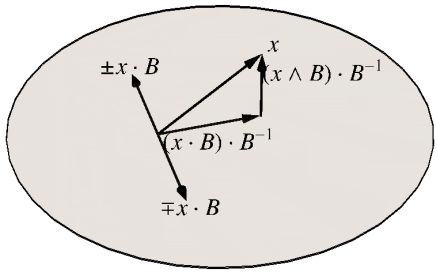


Figure 6.1: Projection and rejection geometry.

Theorem 6.3: Boost.

The boost along a direction vector $\hat{\mathbf{v}}$ and rapidity α is given by

$$x' = e^{-\hat{\mathbf{v}}\alpha/2} x e^{\hat{\mathbf{v}}\alpha/2},$$

where $\hat{\mathbf{v}} = \gamma_{k0} \cos\theta^k$ is an STA bivector representing a spatial direction with direction cosines $\cos\theta^k$.

Proof. We want to demonstrate that this is equivalent to the usual boost formulation. We can start with decomposition of the four-vector x into components that lie in and off of the spacetime plane $\hat{\mathbf{v}}$.

$$\begin{aligned} x &= \left(x^0 + \mathbf{x}\right)\gamma_0 \\ &= \left(x^0 + \mathbf{x}\hat{\mathbf{v}}^2\right)\gamma_0 \\ &= \left(x^0 + (\mathbf{x}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}} + (\mathbf{x}\wedge\hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\gamma_0, \end{aligned} \tag{6.24}$$

where $\mathbf{x} = x \wedge \gamma_0$. The first two components lie in the boost plane, whereas the last is the spatial component of the vector that lies perpendicular to the boost plane. Observe that $\hat{\mathbf{v}}$ anticommutes with the dot product term and commutes with he wedge product term, so we have

$$\begin{aligned} x' &= \left(x^0 + (\mathbf{x}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\gamma_0 e^{\hat{\mathbf{v}}\alpha/2} e^{\hat{\mathbf{v}}\alpha/2} + (\mathbf{x}\wedge\hat{\mathbf{v}})\,\hat{\mathbf{v}}\gamma_0 e^{-\hat{\mathbf{v}}\alpha/2} e^{\hat{\mathbf{v}}\alpha/2} \\ &= \left(x^0 + (\mathbf{x}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\gamma_0 e^{\hat{\mathbf{v}}\alpha} + (\mathbf{x}\wedge\hat{\mathbf{v}})\,\hat{\mathbf{v}}\gamma_0. \end{aligned} \tag{6.25}$$

Noting that $\hat{\mathbf{v}}^2 = 1$, we may expand the exponential in hyperbolic functions, and find that the boosted portion of the vector expands as

$$\begin{aligned}\left(x^0 + (\mathbf{x}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\gamma_0 e^{\hat{\mathbf{v}}\alpha} &= \left(x^0 + (\mathbf{x}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\gamma_0\left(\cosh\alpha + \hat{\mathbf{v}}\sinh\alpha\right)\\ &= \left(x^0 + (\mathbf{x}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\left(\cosh\alpha - \hat{\mathbf{v}}\sinh\alpha\right)\gamma_0\\ &= \left(x^0\cosh\alpha - (\mathbf{x}\cdot\hat{\mathbf{v}})\sinh\alpha\right)\gamma_0\\ &\quad + \left(-x^0\sinh\alpha + (\mathbf{x}\cdot\hat{\mathbf{v}})\cosh\alpha\right)\hat{\mathbf{v}}\gamma_0.\end{aligned} \tag{6.26}$$

We are left with

$$\begin{aligned}x' &= \left(x^0\cosh\alpha - (\mathbf{x}\cdot\hat{\mathbf{v}})\sinh\alpha\right)\gamma_0\\ &\quad + \left((\mathbf{x}\cdot\hat{\mathbf{v}})\cosh\alpha - x^0\sinh\alpha\right)\hat{\mathbf{v}}\gamma_0 + (\mathbf{x}\wedge\hat{\mathbf{v}})\,\hat{\mathbf{v}}\gamma_0\\ &= \begin{bmatrix}\gamma_0 & \hat{\mathbf{v}}\gamma_0\end{bmatrix}\begin{bmatrix}\cosh\alpha & -\sinh\alpha\\ -\sinh\alpha & \cosh\alpha\end{bmatrix}\begin{bmatrix}x^0\\ \mathbf{x}\cdot\hat{\mathbf{v}}\end{bmatrix} + (\mathbf{x}\wedge\hat{\mathbf{v}})\,\hat{\mathbf{v}}\gamma_0,\end{aligned} \tag{6.27}$$

which has the desired Lorentz boost structure. Of course, this is usually seen with $\hat{\mathbf{v}} = \gamma_{10}$ so that the components in the coordinate column vector are (ct, x). □

Theorem 6.4: Spatial rotation.

Given two linearly independent spatial bivectors $\mathbf{a} = a^k\gamma_{k0}, \mathbf{b} = b^k\gamma_{k0}$, a rotation of θ radians in the plane of $\mathbf{a}, \mathbf{b}$ from $\mathbf{a}$ towards $\mathbf{b}$, is given by

$$x' = e^{-i\theta}xe^{i\theta},$$

where $i = (\mathbf{a}\wedge\mathbf{b})/|\mathbf{a}\wedge\mathbf{b}|$, is a unit (spatial) bivector.

Proof. Without loss of generality, we may pick $i = \hat{\mathbf{a}}\hat{\mathbf{b}}$, where $\hat{\mathbf{a}}^2 = \hat{\mathbf{b}}^2 = 1$, and $\hat{\mathbf{a}}\cdot\hat{\mathbf{b}} = 0$. With such an orthonormal basis for the plane, we can decompose our four vector into portions that lie in and off the plane

$$\begin{aligned}x &= \left(x^0 + \mathbf{x}\right)\gamma_0\\ &= \left(x^0 + \mathbf{x}ii^{-1}\right)\gamma_0\\ &= \left(x^0 + (\mathbf{x}\cdot i)\,i^{-1} + (\mathbf{x}\wedge i)\,i^{-1}\right)\gamma_0.\end{aligned} \tag{6.28}$$

The projective term lies in the plane of rotation, whereas the timelike and spatial rejection term are perpendicular. That is

$$\begin{aligned} x_{\parallel} &= (\mathbf{x} \cdot i)\, i^{-1} \gamma_0 \\ x_{\perp} &= \left(x^0 + (\mathbf{x} \wedge i)\, i^{-1}\right) \gamma_0, \end{aligned} \tag{6.29}$$

where $x_{\parallel} \wedge i = 0$, and $x_{\perp} \cdot i = 0$. The plane pseudoscalar i anticommutes with $x_{\parallel}$, and commutes with $x_{\perp}$, so

$$\begin{aligned} x' &= e^{-i\theta/2} \left(x_{\parallel} + x_{\perp}\right) e^{i\theta/2} \\ &= x_{\parallel} e^{i\theta} + x_{\perp}. \end{aligned} \tag{6.30}$$

However

$$\begin{aligned} (\mathbf{x} \cdot i)\, i^{-1} &= \left(\mathbf{x} \cdot \left(\hat{\mathbf{a}} \wedge \hat{\mathbf{b}}\right)\right) \hat{\mathbf{b}}\hat{\mathbf{a}} \\ &= (\mathbf{x} \cdot \hat{\mathbf{a}})\, \hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{a}} - \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right) \hat{\mathbf{a}}\hat{\mathbf{b}}\hat{\mathbf{a}} \\ &= (\mathbf{x} \cdot \hat{\mathbf{a}})\, \hat{\mathbf{a}} + \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right) \hat{\mathbf{b}}, \end{aligned} \tag{6.31}$$

so

$$\begin{aligned} x_{\parallel} e^{i\theta} &= \left((\mathbf{x} \cdot \hat{\mathbf{a}})\, \hat{\mathbf{a}} + \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right) \hat{\mathbf{b}}\right) \gamma_0 \left(\cos\theta + \hat{\mathbf{a}}\hat{\mathbf{b}} \sin\theta\right) \\ &= \hat{\mathbf{a}} \left((\mathbf{x} \cdot \hat{\mathbf{a}}) \cos\theta - \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right) \sin\theta\right) \gamma_0 + \hat{\mathbf{b}} \left((\mathbf{x} \cdot \hat{\mathbf{a}}) \sin\theta + \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right) \cos\theta\right) \gamma_0, \end{aligned} \tag{6.32}$$

so

$$x' = \begin{bmatrix} \hat{\mathbf{a}} & \hat{\mathbf{b}} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \mathbf{x} \cdot \hat{\mathbf{a}} \\ \mathbf{x} \cdot \hat{\mathbf{b}} \end{bmatrix} \gamma_0 + (x \wedge i)\, i^{-1} \gamma_0. \tag{6.33}$$

Observe that this rejection term can be explicitly expanded to

$$(\mathbf{x} \wedge i)\, i^{-1} \gamma_0 = x - (\mathbf{x} \cdot \hat{\mathbf{a}})\, \hat{\mathbf{a}}\gamma_0 - \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right) \hat{\mathbf{b}}\gamma_0. \tag{6.34}$$

This is the timelike component of the vector, plus the spatial component that is normal to the plane. This exponential sandwich transformation rotates only the portion of the vector that lies in the plane, and leaves the rest (timelike and normal) untouched. □

6.3.3 *Problems.*

Exercise 6.6 Verify components relative to boost direction.

In eq. (6.24) the vector x was expanded in terms of the spacetime split. An alternate approach, is to expand as

$$\begin{aligned} x &= x\hat{\mathbf{v}}^2 \\ &= (x \cdot \hat{\mathbf{v}} + x \wedge \hat{\mathbf{v}})\,\hat{\mathbf{v}} \\ &= (x \cdot \hat{\mathbf{v}})\,\hat{\mathbf{v}} + (x \wedge \hat{\mathbf{v}})\,\hat{\mathbf{v}}. \end{aligned} \tag{6.35}$$

Show that

$$(x \cdot \hat{\mathbf{v}})\,\hat{\mathbf{v}} = \left(x^0 + (\mathbf{x} \cdot \hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\gamma_0, \tag{6.36}$$

and

$$(x \wedge \hat{\mathbf{v}})\,\hat{\mathbf{v}} = (\mathbf{x} \wedge \hat{\mathbf{v}})\,\hat{\mathbf{v}}\gamma_0. \tag{6.37}$$

Exercise 6.7 Rotation transformation components.

Given a unit spatial bivector $i = \hat{\mathbf{a}}\hat{\mathbf{b}}$, where $\hat{\mathbf{a}} \cdot \hat{\mathbf{b}} = 0$ and $i^2 = -1$, show that

$$\begin{aligned} (x \cdot i)\,i^{-1} &= (\mathbf{x} \cdot i)\,i^{-1}\gamma_0 \\ &= (\mathbf{x} \cdot \hat{\mathbf{a}})\,\hat{\mathbf{a}}\gamma_0 + \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right)\hat{\mathbf{b}}\gamma_0, \end{aligned} \tag{6.45}$$

and

$$\begin{aligned} (x \wedge i)\,i^{-1} &= (\mathbf{x} \wedge i)\,i^{-1}\gamma_0 \\ &= x - (\mathbf{x} \cdot \hat{\mathbf{a}})\,\hat{\mathbf{a}}\gamma_0 - \left(\mathbf{x} \cdot \hat{\mathbf{b}}\right)\hat{\mathbf{b}}\gamma_0. \end{aligned} \tag{6.46}$$

Also show that i anticommutes with $(x \cdot i)\,i^{-1}$ and commutes with $(x \wedge i)\,i^{-1}$.

6.4 CURVILINEAR COORDINATES, GRADIENT, AND RECIPROCAL FRAMES.

6.4.1 *Motivation.*

I started pondering some aspects of spacetime integration theory, and found that there were some aspects of the concepts of reciprocal frames

that were not clear to me. In the process of sorting those ideas out for myself, I wrote up the following notes.

In the notes below, I will introduce the many of the prerequisite ideas that are needed to express and apply the fundamental theorem of geometric calculus in a 4D relativistic context. The focus will be the Dirac's algebra of special relativity, known as STA (Space Time Algebra) in geometric algebra parlance. If desired, it should be clear how to apply these ideas to lower or higher dimensional spaces, and to plain old Euclidean metrics.

On notation. In Euclidean space we use bold face reciprocal frame vectors $\mathbf{x}^i \cdot \mathbf{x}_j = \delta^i{}_j$, which nicely distinguishes them from the generalized coordinates x_i, x^j associated with the basis or the reciprocal frame, that is

$$\mathbf{x} = x^i \mathbf{x}_i = x_j \mathbf{x}^j. \tag{6.47}$$

On the other hand, it is conventional to use non-bold face for both the four-vectors and their coordinates in STA, such as the following standard basis decomposition

$$x = x^\mu \gamma_\mu = x_\mu \gamma^\mu. \tag{6.48}$$

If we use non-bold face x^μ, x_ν for the coordinates with respect to a specified frame, then we cannot also use non-bold face for the curvilinear basis vectors. To resolve this notational ambiguity, I've chosen to use bold face $\mathbf{x}^\mu, \mathbf{x}_\nu$ symbols as the curvilinear basis elements in this relativistic context, as we do for Euclidean spaces.

6.4.2 *Basis and coordinates.*

Definition 6.8: Standard Dirac basis.

The Dirac basis elements are $\{\gamma_0, \gamma_1, \gamma_2, \gamma_3\}$, satisfying

$$\gamma_0^2 = 1 = -\gamma_k^2, \quad \forall k = 1, 2, 3,$$

and

$$\gamma_\mu \cdot \gamma_\nu = 0, \quad \forall \mu \neq \nu. \tag{6.49}$$

A conventional way of summarizing these orthogonality relationships is $\gamma_\mu \cdot \gamma_\nu = \eta_{\mu\nu}$, where $\eta_{\mu\nu}$ are the elements of the metric $G = \text{diag}(+, -, -, -)$.

Definition 6.9: Reciprocal basis for the standard Dirac basis.

We define a reciprocal basis $\left\{\gamma^0, \gamma^1, \gamma^2, \gamma^3\right\}$ satisfying $\gamma^\mu \cdot \gamma_\nu = \delta^\mu{}_\nu, \forall \mu, \nu \in 0, 1, 2, 3$.

Theorem 6.5: Reciprocal basis uniqueness.

This reciprocal basis is unique, and for our choice of metric has the values

$$\gamma^0 = \gamma_0, \quad \gamma^k = -\gamma_k, \quad \forall k = 1, 2, 3.$$

Proof is left to the reader.

Definition 6.10: Coordinates.

We define the coordinates of a vector with respect to the standard basis as x^μ satisfying

$$x = x^\mu \gamma_\mu,$$

and define the coordinates of a vector with respect to the reciprocal basis as x_μ satisfying

$$x = x_\mu \gamma^\mu,$$

Theorem 6.6: Coordinates.

Given the definitions above, we may compute the coordinates of a vector, simply by dotting with the basis elements

$$x^\mu = x \cdot \gamma^\mu,$$

and

$$x_\mu = x \cdot \gamma_\mu,$$

Proof. This follows by straightforward computation

$$\begin{aligned} x \cdot \gamma^\mu &= (x^\nu \gamma_\nu) \cdot \gamma^\mu \\ &= x^\nu (\gamma_\nu \cdot \gamma^\mu) \\ &= x^\nu \delta_\nu{}^\mu \\ &= x^\mu, \end{aligned} \tag{6.50}$$

and

$$\begin{aligned} x \cdot \gamma_\mu &= (x_\nu \gamma^\nu) \cdot \gamma_\mu \\ &= x_\nu \left(\gamma^\nu \cdot \gamma_\mu\right) \\ &= x_\nu \delta^\nu{}_\mu \\ &= x_\mu. \end{aligned} \tag{6.51}$$

□

6.4.3 *Derivative operators.*

We'd like to determine the form of the (spacetime) gradient operator. The gradient can be defined in terms of coordinates directly, but we choose an implicit definition, in terms of the directional derivative.

Definition 6.11: Directional derivative and gradient.

Let $F = F(x)$ be a four-vector parameterized multivector. The directional derivative of F with respect to the (four-vector) direction a is denoted

$$(a \cdot \nabla) F = \lim_{\epsilon \to 0} \frac{F(x + \epsilon a) - F(x)}{\epsilon},$$

where ∇ is called the space time gradient.

Theorem 6.7: Gradient.

The standard basis representation of the gradient is

$$\nabla = \gamma^\mu \partial_\mu,$$

where

$$\partial_\mu = \frac{\partial}{\partial x^\mu}.$$

Proof. The Dirac gradient pops naturally out of the coordinate representation of the directional derivative, as we can see by expanding $F(x + \epsilon a)$ in Taylor series

$$\begin{aligned} F(x + \epsilon a) &= F(x) + \epsilon \frac{dF(x + \epsilon a)}{d\epsilon} + O(\epsilon^2) \\ &= F(x) + \epsilon \frac{\partial F}{\partial(x^\mu + \epsilon a^\mu)} \frac{\partial(x^\mu + \epsilon a^\mu)}{\partial \epsilon} \\ &= F(x) + \epsilon \frac{\partial F}{\partial(x^\mu + \epsilon a^\mu)} a^\mu. \end{aligned} \tag{6.52}$$

The directional derivative is

$$\begin{aligned} \lim_{\epsilon \to 0} \frac{F(x + \epsilon a) - F(x)}{\epsilon} &= \lim_{\epsilon \to 0} a^\mu \frac{\partial F}{\partial(x^\mu + \epsilon a^\mu)} \\ &= a^\mu \frac{\partial F}{\partial x^\mu} \\ &= (a^\nu \gamma_\nu) \cdot \gamma^\mu \frac{\partial F}{\partial x^\mu} \\ &= a \cdot \left(\gamma^\mu \partial_\mu\right) F. \end{aligned} \tag{6.53}$$

□

6.4.4 *Curvilinear bases.*

Curvilinear bases are the foundation of the fundamental theorem of multivector calculus. This form of integral calculus is defined over parameterized surfaces (called manifolds) that satisfy some specific non-degeneracy and continuity requirements.

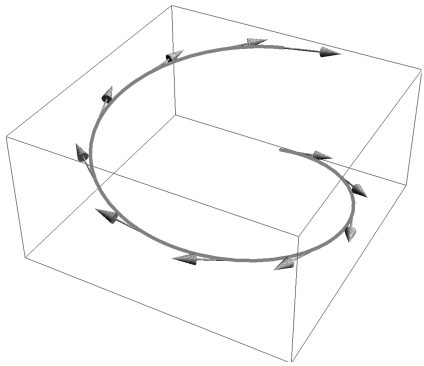
A parameterized vector $x(u, v, \cdots w)$ can be thought of as tracing out a hypersurface (curve, surface, volume, ...), where the dimension of the hypersurface depends on the number of parameters. At each point, a bases can be constructed from the differentials of the parameterized vector. Such a basis is called the tangent space to the surface at the point in question. Our curvilinear bases will be related to these differentials. We will also be interested in a dual basis that is restricted to the span of the tangent space.

This dual basis will be called the reciprocal frame, and line the basis of the tangent space itself, also varies from point to point on the surface.

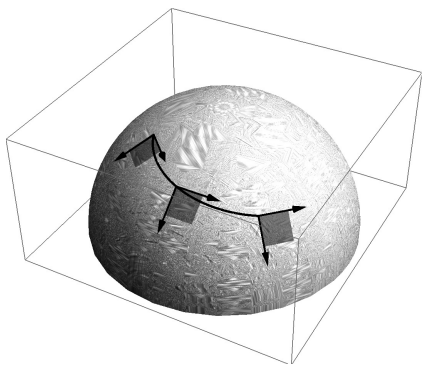
One and two parameter spaces are illustrated in fig. 6.2, and the tangent space basis at a specific point of a two parameter surface, $x(u^0, u^1)$, is also illustrated in fig. 6.3 . The differential directions that span the tangent space are

$$\begin{aligned} d\mathbf{x}_0 &= \frac{\partial x}{\partial u^0} du^0 \\ d\mathbf{x}_1 &= \frac{\partial x}{\partial u^1} du^1, \end{aligned} \tag{6.54}$$

and the tangent space itself is span $\{d\mathbf{x}_0, d\mathbf{x}_1\}$. We may form an oriented surface area element $d\mathbf{x}_0 \wedge d\mathbf{x}_1$ over this surface. Tangent spaces associ-



(a)



(b)

Figure 6.2: One and two parameter curves, with illustration of tangent spaces.

ated with 3 or more parameters cannot be easily visualized in three dimensions, but the idea generalizes algebraically without trouble.

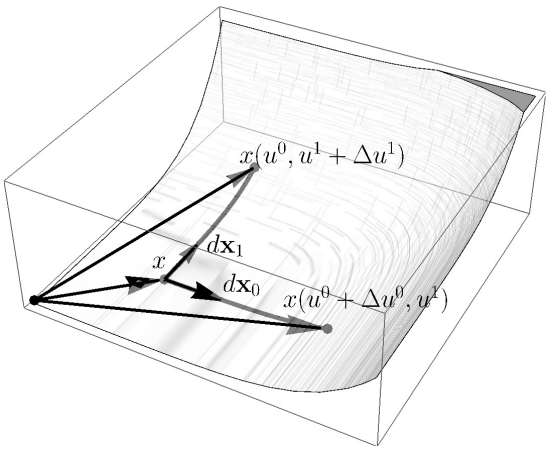


Figure 6.3: Two parameter surface.

Definition 6.12: Tangent basis and space.

Given a parameterization $x = x(u^0, \cdots, u^N)$, where $N < 4$, the span of the vectors

$$\mathbf{x}_\mu = \frac{\partial x}{\partial u^\mu},$$

is called the tangent space for the hypersurface associated with the parameterization, and it's basis is $\left\{\mathbf{x}_\mu\right\}$.

Later we will see that parameterization constraints must be imposed, as not all surfaces generated by a set of parameterizations are useful for integration theory. In particular, degenerate parameterizations for which the wedge products of the tangent space basis vectors are zero, or those wedge products cannot be inverted, are not physically meaningful. We require the functional form of parameterizations to be differentiable, to provide a one to one invertible mapping from the parameter space to the vector space. Properly behaved surfaces of this sort are called manifolds.

Having introduced curvilinear coordinates associated with a parameterization, we can now determine the form of the gradient with respect to a parameterization of spacetime.

Theorem 6.8: Gradient, curvilinear representation.

Given a spacetime parameterization $x = x(u^0, u^1, u^2, u^3)$, the gradient with respect to the parameters u^μ is

$$\nabla = \sum_\mu \mathbf{x}^\mu \frac{\partial}{\partial u^\mu},$$

where

$$\mathbf{x}^\mu = \nabla u^\mu.$$

The vectors $\mathbf{x}^\mu$ are called the reciprocal frame vectors, and the ordered set $\left\{\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3\right\}$ is called the reciprocal basis.

It is convenient to define $\partial_\mu \equiv \partial/\partial u^\mu$, so that the gradient can be expressed in mixed index representation

$$\nabla = \mathbf{x}^\mu \partial_\mu.$$

This introduces some notational ambiguity, since we used $\partial_\mu = \partial/\partial x^\mu$ for the standard basis derivative operators too, but we will be careful to be explicit when there is any doubt about what is intended.

Proof. The proof follows by application of the chain rule.

$$\begin{aligned}
\nabla F &= \gamma^\alpha \frac{\partial F}{\partial x^\alpha} \\
&= \gamma^\alpha \frac{\partial u^\mu}{\partial x^\alpha} \frac{\partial F}{\partial u^\mu} \\
&= (\nabla u^\mu) \frac{\partial F}{\partial u^\mu} \\
&= \mathbf{x}^\mu \frac{\partial F}{\partial u^\mu}.
\end{aligned} \tag{6.55}$$

□

Theorem 6.9: Reciprocal relationship.

The vectors $\mathbf{x}^\mu = \nabla u^\mu$, and $\mathbf{x}_\mu = \partial x/\partial u^\mu$ satisfy the reciprocal relationship

$$\mathbf{x}^\mu \cdot \mathbf{x}_\nu = \delta^\mu{}_\nu.$$

Proof.

$$\begin{aligned}
\mathbf{x}^\mu \cdot \mathbf{x}_\nu &= \nabla u^\mu \cdot \frac{\partial x}{\partial u^\nu} \\
&= \left(\gamma^\alpha \frac{\partial u^\mu}{\partial x^\alpha}\right) \cdot \left(\frac{\partial x^\beta}{\partial u^\nu}\gamma_\beta\right) \\
&= \delta^\alpha{}_\beta \frac{\partial u^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial u^\nu} \\
&= \frac{\partial u^\mu}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial u^\nu} \\
&= \frac{\partial u^\mu}{\partial u^\nu} \\
&= \delta^\mu{}_\nu.
\end{aligned} \tag{6.56}$$

□

The following example parameterization scales the proper time parameter, and uses polar coordinates in the x, y plane.

Exercise 6.8 Compute the curvilinear and reciprocal basis.

Given $x(t, \rho, \theta, z) = ct\gamma_0 + \gamma_1 \rho e^{-i\theta} + z\gamma_3$, where $i = \gamma_1\gamma_2$, compute the curvilinear frame vectors and their reciprocals.

Despite being a fairly simple parameterization, it was still fairly difficult to solve for the gradients when the parameterization introduced coupling between the coordinates. Because all the tangent space vectors are mutually orthogonal, we didn't need to go through all this trouble, since we could have just computed the inverses of all the tangent space vectors

$$\begin{aligned}
\mathbf{x}^0 &= \frac{1}{\mathbf{x}_0} = \frac{1}{c\gamma_0} = \frac{1}{c}\gamma^0 \\
\mathbf{x}^1 &= \frac{1}{\mathbf{x}_1} = \frac{1}{\gamma_1 e^{-i\alpha}} = \frac{1}{e^{-i\alpha}}\frac{1}{\gamma_1} = e^{i\alpha}\gamma^1 = \gamma^1 e^{-i\alpha} \\
\mathbf{x}^2 &= \frac{1}{\mathbf{x}_2} = \frac{1}{\rho\gamma_2 e^{-i\alpha}} = \frac{1}{\rho}\frac{1}{e^{-i\alpha}}\frac{1}{\gamma_2} = \frac{1}{\rho}e^{i\alpha}\gamma^2 = \frac{1}{\rho}\gamma^2 e^{-i\alpha} \\
\mathbf{x}^3 &= \frac{1}{\mathbf{x}_3} = \frac{1}{\gamma_3} = \gamma^3.
\end{aligned} \tag{6.68}$$

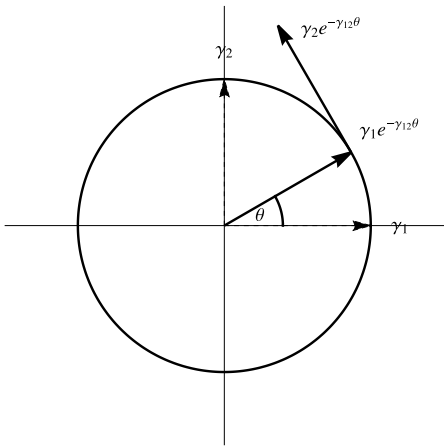


Figure 6.4: Tangent space direction vectors.

We will not generally be this lucky, and want a less labor intensive strategy to find the reciprocal frame that works for the general case.

There is one additional special cawe. When we have a full parameterization of spacetime, then we can do this with nothing more than a matrix inversion.

Theorem 6.10: Reciprocal frame matrix equations.

Given a spacetime basis $\{\mathbf{x}_0, \cdots \mathbf{x}_3\}$, let $[\mathbf{x}_\mu]$ and $[\mathbf{x}^\nu]$ be column matrices with the coordinates of these vectors and their reciprocals, with respect to the standard basis $\{\gamma_0, \gamma_1, \gamma_2, \gamma_3\}$. Let

$$A = \begin{bmatrix} [\mathbf{x}_0] & \cdots & [\mathbf{x}_3] \end{bmatrix}, \qquad X = \begin{bmatrix} [\mathbf{x}^0] & \cdots & [\mathbf{x}^3] \end{bmatrix}.$$

The coordinates of the reciprocal frame vectors can be found by solving

$$A^{\mathrm{T}} G X = 1,$$

where $G = \mathrm{diag}(1, -1, -1, -1)$ and the RHS is an 4×4 identity matrix.

Proof. Let $\mathbf{x}_\mu = a_\mu{}^\alpha \gamma_\alpha, \mathbf{x}^\nu = b^{\nu\beta}\gamma_\beta$, so that

$$A = \begin{bmatrix} a_\nu{}^\mu \end{bmatrix}, \tag{6.69}$$

and

$$X = \left[b^{\nu\mu}\right], \tag{6.70}$$

where $\mu \in [0, 3]$ are the row indexes and $\nu \in [0, N-1]$ are the column indexes. The reciprocal frame satisfies $\mathbf{x}_\mu \cdot \mathbf{x}^\nu = \delta_\mu{}^\nu$, which has the coordinate representation of

$$\begin{aligned}\mathbf{x}_\mu \cdot \mathbf{x}^\nu &= \left(a_\mu{}^\alpha \gamma_\alpha\right) \cdot \left(b^{\nu\beta}\gamma_\beta\right) \\ &= a_\mu{}^\alpha \eta_{\alpha\beta} b^{\nu\beta} \\ &= [A^{\mathrm{T}} G B]_\mu{}^\nu,\end{aligned} \tag{6.71}$$

where μ is the row index and ν is the column index. □

Exercise 6.9 Matrix inversion reciprocals.

For the parameterization of exercise 6.8, find the reciprocal frame vectors by matrix inversion.

There will be circumstances where we parameterize only a subset of spacetime, and are interested in calculating quantities associated with such a surface. For example, suppose that

$$x(\rho, \theta) = \gamma_1 \rho e^{-i\theta}, \tag{6.78}$$

where $i = \gamma_1\gamma_2$ as before. We are now parameterizing only the $x - y$ plane. We will still find

$$\begin{aligned}\mathbf{x}_1 &= \gamma_1 e^{-i\theta} \\ \mathbf{x}_2 &= \gamma_2 \rho e^{-i\theta}.\end{aligned} \tag{6.79}$$

We can compute the reciprocals of these vectors using the gradient method. It's possible to state matrix equations representing the reciprocal relationship of theorem 6.9, which, in this case, is $X^{\mathrm{T}} G Y = 1$, where the RHS is a 2×2 identity matrix, and X, Y are 4×2 matrices of coordinates, with

$$X = \begin{bmatrix} 0 & 0 \\ C & -\rho S \\ S & \rho C \\ 0 & 0 \end{bmatrix}. \tag{6.80}$$

We no longer have a square matrix problem to solve, and our solution set is multivalued. In particular, this matrix equation has solutions

$$\begin{aligned}\mathbf{x}^1 &= \gamma^1 e^{-i\theta} + \alpha\gamma^0 + \beta\gamma^3 \\ \mathbf{x}^2 &= \frac{\gamma^2}{\rho} e^{-i\theta} + \alpha'\gamma^0 + \beta'\gamma^3.\end{aligned} \tag{6.81}$$

where $\alpha, \alpha', \beta, \beta'$ are arbitrary constants. In the example we considered, we saw that our ρ, θ parameters were functions of only x^1, x^2, so taking gradients could not introduce any γ^0, γ^3 dependence in $\mathbf{x}^1, \mathbf{x}^2$. It seems reasonable to assert that we seek an algebraic method of computing a set of vectors that satisfies the reciprocal relationships, where that set of vectors is restricted to the tangent space. We will need to figure out how to prove that this reciprocal construction is identical to the parameter gradients, but let's start with figuring out what such a tangent space restricted solution looks like.

Theorem 6.11: Reciprocal frame for two parameter subspace.

Given two vectors, $\mathbf{x}_1, \mathbf{x}_2$, the vectors $\mathbf{x}^1, \mathbf{x}^2 \in \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$ such that $\mathbf{x}^\mu \cdot \mathbf{x}_\nu = \delta^\mu{}_\nu$ are given by

$$\begin{aligned}\mathbf{x}^1 &= \mathbf{x}_2 \cdot \frac{1}{\mathbf{x}_1 \wedge \mathbf{x}_2} \\ \mathbf{x}^2 &= -\mathbf{x}_1 \cdot \frac{1}{\mathbf{x}_1 \wedge \mathbf{x}_2},\end{aligned}$$

provided $\mathbf{x}_1 \wedge \mathbf{x}_2 \neq 0$ and $(\mathbf{x}_1 \wedge \mathbf{x}_2)^2 \neq 0$.

Proof. The most general set of vectors that satisfy the span constraint are

$$\begin{aligned}\mathbf{x}^1 &= a\mathbf{x}_1 + b\mathbf{x}_2 \\ \mathbf{x}^2 &= c\mathbf{x}_1 + d\mathbf{x}_2.\end{aligned} \tag{6.82}$$

We can use wedge products with either $\mathbf{x}_1$ or $\mathbf{x}_2$ to eliminate the other from the RHS

$$\begin{aligned}\mathbf{x}^1 \wedge \mathbf{x}_2 &= a\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \\ \mathbf{x}^1 \wedge \mathbf{x}_1 &= -b\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \\ \mathbf{x}^2 \wedge \mathbf{x}_2 &= c\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \\ \mathbf{x}^2 \wedge \mathbf{x}_1 &= -d\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right),\end{aligned} \tag{6.83}$$

and then dot both sides with $\mathbf{x}_1 \wedge \mathbf{x}_2$ to produce four scalar equations

$$\begin{aligned} a\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2 &= \left(\mathbf{x}^1 \wedge \mathbf{x}_2\right) \cdot \left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \\ &= \left(\mathbf{x}_2 \cdot \mathbf{x}_1\right) \cancel{\mathbf{x}^1 \cdot \mathbf{x}_2} - \left(\mathbf{x}_2 \cdot \mathbf{x}_2\right)\left(\mathbf{x}^1 \cdot \mathbf{x}_1\right) \\ &= -\mathbf{x}_2 \cdot \mathbf{x}_2 \end{aligned} \tag{6.84a}$$

$$\begin{aligned} -b\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2 &= \left(\mathbf{x}^1 \wedge \mathbf{x}_1\right) \cdot \left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \\ &= \cancel{\mathbf{x}^1 \cdot \mathbf{x}_2}\left(\mathbf{x}_1 \cdot \mathbf{x}_1\right) - \left(\mathbf{x}^1 \cdot \mathbf{x}_1\right)\left(\mathbf{x}_1 \cdot \mathbf{x}_2\right) \\ &= -\mathbf{x}_1 \cdot \mathbf{x}_2 \end{aligned} \tag{6.84b}$$

$$\begin{aligned} c\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2 &= \left(\mathbf{x}^2 \wedge \mathbf{x}_2\right) \cdot \left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \\ &= \left(\mathbf{x}_2 \cdot \mathbf{x}_1\right)\left(\mathbf{x}^2 \cdot \mathbf{x}_2\right) - \left(\mathbf{x}_2 \cdot \mathbf{x}_2\right) \cancel{\mathbf{x}^2 \cdot \mathbf{x}_1} \\ &= \mathbf{x}_2 \cdot \mathbf{x}_1 \end{aligned} \tag{6.84c}$$

$$\begin{aligned} -d\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2 &= \left(\mathbf{x}^2 \wedge \mathbf{x}_1\right) \cdot \left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \\ &= \left(\mathbf{x}_1 \cdot \mathbf{x}_1\right)\left(\mathbf{x}^2 \cdot \mathbf{x}_2\right) - \left(\mathbf{x}_1 \cdot \mathbf{x}_2\right) \cancel{\mathbf{x}^2 \cdot \mathbf{x}_1} \\ &= \mathbf{x}_1 \cdot \mathbf{x}_1 . \end{aligned} \tag{6.84d}$$

Putting the pieces together we have

$$\begin{aligned} \mathbf{x}^1 &= \frac{-\left(\mathbf{x}_2 \cdot \mathbf{x}_2\right)\mathbf{x}_1 + \left(\mathbf{x}_1 \cdot \mathbf{x}_2\right)\mathbf{x}_2}{\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2} \\ &= \frac{\mathbf{x}_2 \cdot \left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)}{\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2} \\ &= \mathbf{x}_2 \cdot \frac{1}{\mathbf{x}_1 \wedge \mathbf{x}_2} \end{aligned} \tag{6.85a}$$

$$\begin{aligned} \mathbf{x}^2 &= \frac{\left(\mathbf{x}_1 \cdot \mathbf{x}_2\right)\mathbf{x}_1 - \left(\mathbf{x}_1 \cdot \mathbf{x}_1\right)\mathbf{x}_2}{\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2} \\ &= \frac{-\mathbf{x}_1 \cdot \left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)}{\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right)^2} \\ &= -\mathbf{x}_1 \cdot \frac{1}{\mathbf{x}_1 \wedge \mathbf{x}_2} \end{aligned} \tag{6.85b}$$

□

Lemma 6.1: Distribution identity.

Given k-vectors B, C and a vector a, where the grade of C is greater than that of B, then

$$(a \wedge B) \cdot C = a \cdot (B \cdot C) .$$

See [15] for a proof.

Theorem 6.12: Higher order tangent space reciprocals.

Given an N parameter tangent space with basis $\{\mathbf{x}_0, \mathbf{x}_1, \cdots \mathbf{x}_{N-1}\}$, the reciprocals are given by

$$\mathbf{x}^\mu = (-1)^\mu \left(\mathbf{x}_0 \wedge \cdots \check{\mathbf{x}_\mu} \cdots \wedge \mathbf{x}_{N-1}\right) \cdot I_N^{-1},$$

where the checked term $(\check{\mathbf{x}_\mu})$ indicates that all terms are included in the wedges except the $\mathbf{x}_\mu$ term, and $I_N = \mathbf{x}_0 \wedge \cdots \mathbf{x}_{N-1}$ is the pseudoscalar for the tangent space.

Proof. I'll outline the proof for the three parameter tangent space case, from which the pattern will be clear. The motivation for this proof is a re-examination of the algebraic structure of the two vector solution. Suppose we have a tangent space basis $\{\mathbf{x}_0, \mathbf{x}_1\}$, for which we've shown that

$$\mathbf{x}^0 = \mathbf{x}_1 \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1} = \frac{\mathbf{x}_1 \cdot (\mathbf{x}_0 \wedge \mathbf{x}_1)}{(\mathbf{x}_0 \wedge \mathbf{x}_1)^2}. \tag{6.86}$$

If we dot with $\mathbf{x}_0$ and $\mathbf{x}_1$ respectively, we find

$$\mathbf{x}_0 \cdot \mathbf{x}^0 = \mathbf{x}_0 \cdot \frac{\mathbf{x}_1 \cdot (\mathbf{x}_0 \wedge \mathbf{x}_1)}{(\mathbf{x}_0 \wedge \mathbf{x}_1)^2} = (\mathbf{x}_0 \wedge \mathbf{x}_1) \cdot \frac{\mathbf{x}_0 \wedge \mathbf{x}_1}{(\mathbf{x}_0 \wedge \mathbf{x}_1)^2}. \tag{6.87}$$

We end up with unity as expected. Here the "factored" out vector is reincorporated into the pseudoscalar using the distribution identity lemma 6.1. Similarly, dotting with $\mathbf{x}_1$, we find

$$\mathbf{x}_1 \cdot \mathbf{x}^0 = \mathbf{x}_1 \cdot \frac{\mathbf{x}_1 \cdot (\mathbf{x}_0 \wedge \mathbf{x}_1)}{(\mathbf{x}_0 \wedge \mathbf{x}_1)^2} = (\mathbf{x}_1 \wedge \mathbf{x}_1) \cdot \frac{\mathbf{x}_0 \wedge \mathbf{x}_1}{(\mathbf{x}_0 \wedge \mathbf{x}_1)^2}. \tag{6.88}$$

This is zero, since wedging a vector with itself is zero. We can perform such an operation in reverse, taking the square of the tangent space pseudoscalar, and factoring out one of the basis vectors. After this, division by that squared pseudoscalar will normalize things.

For a three parameter tangent space with basis $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2\}$, we can factor out any of the tangent vectors like so

$$\begin{aligned}
\left(\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2\right)^2 &= \mathbf{x}_0 \cdot \left(\left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \cdot \left(\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2\right)\right) \\
&= (-1)\mathbf{x}_1 \cdot \left(\left(\mathbf{x}_0 \wedge \mathbf{x}_2\right) \cdot \left(\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2\right)\right) \\
&= (-1)^2\mathbf{x}_2 \cdot \left(\left(\mathbf{x}_0 \wedge \mathbf{x}_1\right) \cdot \left(\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2\right)\right).
\end{aligned} \tag{6.89}$$

The toggling of sign reflects the number of permutations required to move the vector of interest to the front of the wedge sequence. Having factored out any one of the vectors, we can rearrange to find that vector that is it's inverse and perpendicular to all the others.

$$\begin{aligned}
\mathbf{x}^0 &= (-1)^0 \left(\mathbf{x}_1 \wedge \mathbf{x}_2\right) \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2} \\
\mathbf{x}^1 &= (-1)^1 \left(\mathbf{x}_0 \wedge \mathbf{x}_2\right) \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2} \\
\mathbf{x}^2 &= (-1)^2 \left(\mathbf{x}_0 \wedge \mathbf{x}_1\right) \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2}.
\end{aligned} \tag{6.90}$$

□

In the fashion above, should we want the reciprocal frame for all of spacetime given dimension 4 tangent space, we can state it trivially

$$\begin{aligned}
\mathbf{x}^0 &= (-1)^0 \left(\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \mathbf{x}_3\right) \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \mathbf{x}_3} \\
\mathbf{x}^1 &= (-1)^1 \left(\mathbf{x}_0 \wedge \mathbf{x}_2 \wedge \mathbf{x}_3\right) \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \mathbf{x}_3} \\
\mathbf{x}^2 &= (-1)^2 \left(\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_3\right) \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \mathbf{x}_3} \\
\mathbf{x}^3 &= (-1)^3 \left(\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2\right) \cdot \frac{1}{\mathbf{x}_0 \wedge \mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \mathbf{x}_3}.
\end{aligned} \tag{6.91}$$

This is probably not an efficient way to compute all these reciprocals, since we can utilize a single matrix inversion to solve them in one shot. However, there are theoretical advantages to this construction that will be useful when we get to integration theory.

On degeneracy. A small mention of degeneracy was mentioned above. Regardless of metric, $\mathbf{x}_0 \wedge \mathbf{x}_1 = 0$ means that this pair of vectors are colinear. A tangent space with such a pseudoscalar is clearly undesirable, and we must construct parameterizations for which the area element is non-zero in all regions of interest.

Things get more interesting in mixed signature spaces where we can have vectors that square to zero (i.e. lightlike). If the tangent space pseudoscalar has a lightlike factor, then that pseudoscalar will not be invertible. Such a degeneracy will will likely lead to many other troubles, and parameterizations of this sort should be avoided.

This following problem illustrates an example of this sort of degenerate parameterization.

Exercise 6.10 Degenerate surface parameterization.

Given a spacetime plane parameterization $x(u, v) = ua + vb$, where

$$a = \gamma_0 + \gamma_1 + \gamma_2 + \gamma_3, \tag{6.92}$$

$$b = \gamma_0 - \gamma_1 + \gamma_2 - \gamma_3, \tag{6.93}$$

show that this is a degenerate parameterization, and find the bivector that represents the tangent space. Are these vectors lightlike, spacelike, or timelike? Comment on whether this parameterization represents a physically relevant spacetime surface.

Final notes. There are a few loose ends above. In particular, we haven't conclusively proven that the set of reciprocal vectors $\mathbf{x}^\mu = \nabla u^\mu$ are exactly those obtained through algebraic means. For a full parameterization of spacetime, they are necessarily the same, since both are unique. So we know that eq. (6.91) must equal the reciprocals obtained by evaluating the gradient for a full parameterization (and this must also equal the reciprocals that we can obtain through matrix inversion.) We have also not proved explicitly that the three parameter construction of the reciprocals in eq. (6.90) is in the tangent space, but that is a fairly trivial observation, so that can be left as an exercise for the reader dismissal. Some additional thought about this is probably required, but it seems reasonable to put that on the back burner and move on to some applications.

6.4.5 *More examples.*

new blog post: Here are a few additional examples of reciprocal frame calculations.

Exercise 6.11 Unidirectional arbitrary functional dependence.

Let

$$x = af(u),$$

where a is a constant vector and $f(u)$ is some arbitrary differentiable function with a non-zero derivative in the region of interest.

Exercise 6.12 Linear two variable parameterization.

Let $x = au + bv$, where $x \wedge a \wedge b = 0$ represents spacetime plane (also the tangent space.) Find the curvilinear coordinates and their reciprocals.

Exercise 6.13 Quadratic two variable parameterization.

Now consider a variation of the previous problem, with $x = au^2 + bv^2$. Find the curvilinear coordinates and their reciprocals.

Exercise 6.14 Reciprocal frame for generalized cylindrical parameterization.

Let the vector parameterization be $x(\rho, \theta) = \rho e^{-i\theta/2} x(\rho_0, \theta_0) e^{i\theta}$, where $i^2 = \pm 1$ is a unit bivector ($+1$ for a boost, and -1 for a rotation), and where θ, ρ are scalars. Find the tangent space vectors and their reciprocals.

Note that this is cylindrical parameterization for the rotation case, and traces out hyperbolic regions for the boost case. The boost case is illustrated in fig. 6.5 where hyperbolas in the light cone are found for boosts of γ_0 with various values of ρ, and the spacelike hyperbolas are boosts of γ_1, again for various values of ρ.

6.4.6 *Parameterization of a general linear transformation.*

Given N parameters $u^0, u^1, \cdots u^{N-1}$, a general linear transformation from the parameter space to the vector space has the form

$$x = a^\alpha{}_\beta \gamma_\alpha u^\beta, \tag{6.116}$$

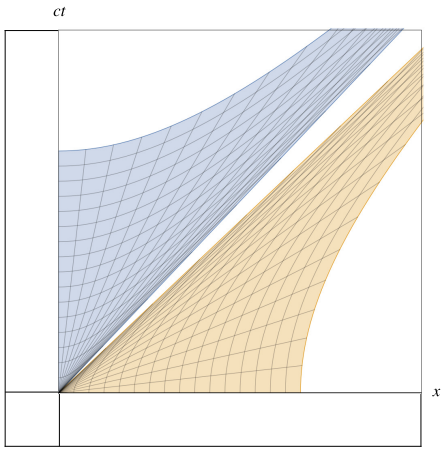


Figure 6.5: "Cylindrical" boost parameterization.

where $\beta \in [0, \cdots, N-1]$ and $\alpha \in [0, 3]$. For such a general transformation, observe that the curvilinear basis vectors are

$$\begin{aligned}
\mathbf{x}_\mu &= \frac{\partial x}{\partial u^\mu} \\
&= \frac{\partial}{\partial u^\mu} a^\alpha{}_\beta \gamma_\alpha u^\beta \\
&= a^\alpha{}_\mu \gamma_\alpha.
\end{aligned} \tag{6.117}$$

We find an interpretation of $a^\alpha{}_\mu$ by dotting $\mathbf{x}_\mu$ with the reciprocal frame vectors of the standard basis

$$\begin{aligned}
\mathbf{x}_\mu \cdot \gamma^\nu &= a^\alpha{}_\mu (\gamma_\alpha \cdot \gamma^\nu) \\
&= a^\nu{}_\mu,
\end{aligned} \tag{6.118}$$

so

$$x = \mathbf{x}_\mu u^\mu. \tag{6.119}$$

We are able to reinterpret eq. (6.116) as a contraction of the tangent space vectors with the parameters, scaling and summing these direction vectors to characterize all the points in the tangent plane.

Theorem 6.13: Projecting onto the tangent space.

Let T represent the tangent space. The projection of a vector onto the tangent space has the form

$$\mathrm{Proj}_T\, y = (y \cdot \mathbf{x}^\mu)\, \mathbf{x}_\mu = (y \cdot \mathbf{x}_\mu)\, \mathbf{x}^\mu.$$

Proof. Let's designate a as the portion of the vector y that lies outside of the tangent space

$$y = y^\mu \mathbf{x}_\mu + a. \tag{6.120}$$

If we knew the coordinates y^μ, we would have a recipe for the projection. Algebraically, requiring that a lies outside of the tangent space, is equivalent to stating $a \cdot \mathbf{x}_\mu = a \cdot \mathbf{x}^\mu = 0$. We use that fact, and then take dot products

$$\begin{aligned} y \cdot \mathbf{x}^\nu &= \left(y^\mu \mathbf{x}_\mu + a\right) \cdot \mathbf{x}^\nu \\ &= y^\nu, \end{aligned} \tag{6.121}$$

so

$$y = (y \cdot \mathbf{x}^\mu)\, \mathbf{x}_\mu + a. \tag{6.122}$$

Similarly, the tangent space projection can be expressed as a linear combination of reciprocal basis elements

$$y = y_\mu \mathbf{x}^\mu + a. \tag{6.123}$$

Dotting with $\mathbf{x}_\mu$, we have

$$\begin{aligned} y \cdot \mathbf{x}^\mu &= (y_\alpha \mathbf{x}^\alpha + a) \cdot \mathbf{x}_\mu \\ &= y_\mu, \end{aligned} \tag{6.124}$$

so

$$y = (y \cdot \mathbf{x}^\mu)\, \mathbf{x}_\mu + a. \tag{6.125}$$

We find the two stated ways of computing the projection.

Observe that, for the special case that all of $\left\{\mathbf{x}_\mu\right\}$ are orthogonal, the equivalence of these two projection methods follows directly, since

$$\begin{aligned} (y \cdot \mathbf{x}^\mu)\, \mathbf{x}_\mu &= \left(y \cdot \frac{1}{\mathbf{x}_\mu}\right) \frac{1}{\mathbf{x}^\mu} \\ &= \left(y \cdot \frac{\mathbf{x}_\mu}{\left(\mathbf{x}_\mu\right)^2}\right) \frac{\mathbf{x}^\mu}{(\mathbf{x}^\mu)^2} \\ &= \left(y \cdot \mathbf{x}_\mu\right) \mathbf{x}^\mu. \end{aligned} \tag{6.126}$$

□

6.5 FUNDAMENTAL THEOREM OF GEOMETRIC CALCULUS.

6.5.1 *Motivation.*

I've been slowly working my way towards a statement of the fundamental theorem of integral calculus, where the functions being integrated are elements of the Dirac algebra (space time multivectors in the geometric algebra parlance.)

This is interesting because we want to be able to do line, surface, 3-volume and 4-volume space time integrals. We have many $\mathbb{R}^3$ integral theorems

$$\int_A^B d\mathbf{l} \cdot \boldsymbol{\nabla} f = f(B) - f(A), \tag{6.127a}$$

$$\int_S dA\, \hat{\mathbf{n}} \times \boldsymbol{\nabla} f = \oint_{\partial S} d\mathbf{x}\, f, \tag{6.127b}$$

$$\int_S dA\, \hat{\mathbf{n}} \cdot (\boldsymbol{\nabla} \times \mathbf{f}) = \oint_{\partial S} d\mathbf{x} \cdot \mathbf{f}, \tag{6.127c}$$

$$\int_S dxdy \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right) = \oint_{\partial S} P dx + Q dy, \tag{6.127d}$$

$$\int_V dV\, \boldsymbol{\nabla} f = \int_{\partial V} dA\, \hat{\mathbf{n}} f, \tag{6.127e}$$

$$\int_V dV\, \boldsymbol{\nabla} \times \mathbf{f} = \int_{\partial V} dA\, \hat{\mathbf{n}} \times \mathbf{f}, \tag{6.127f}$$

$$\int_V dV\, \boldsymbol{\nabla} \cdot \mathbf{f} = \int_{\partial V} dA\, \hat{\mathbf{n}} \cdot \mathbf{f}, \tag{6.127g}$$

and want to know how to generalize these to four dimensions and also make sure that we are handling the relativistic mixed signature correctly. If our starting point was the mess of equations above, we'd be in trouble, since it is not obvious how these generalize. All the theorems with

unit normals have to be handled completely differently in four dimensions since we don't have a unique normal to any given spacetime plane. What comes to our rescue is the Fundamental Theorem of Geometric Calculus (FTGC), which has the form

$$\int F d^n\mathbf{x} \overleftrightarrow{\partial} G = \int F d^{n-1}\mathbf{x} G, \tag{6.128}$$

where F, G are multivectors functions (i.e. sums of products of vectors.) We've seen ([19], [15]) that all the identities above are special cases of the fundamental theorem.

Do we need any special care to state the FTGC correctly for our relativistic case? It turns out that the answer is no! Tangent and reciprocal frame vectors do all the heavy lifting, and we can use the fundamental theorem as is, even in our mixed signature space. The only real change that we need to make is use spacetime gradient and vector derivative operators instead of their spatial equivalents. We will see how this works below. Note that instead of starting with eq. (6.128) directly, I will attempt to build up to that point in a progressive fashion that is hopefully does not require the reader to make too many unjustified mental leaps.

6.5.2 *Multivector line integrals.*

We want to define multivector line integrals to start with. Recall that in $\mathbb{R}^3$ we would say that for scalar functions f, the integral

$$\int d\mathbf{x} f = \int f d\mathbf{x},$$

is a line integral. Also, for vector functions $\mathbf{f}$ we call

$$\int d\mathbf{x} \cdot \mathbf{f} = \frac{1}{2} \int d\mathbf{x}\, \mathbf{f} + \mathbf{f} d\mathbf{x}.$$

a line integral. In order to generalize line integrals to multivector functions, we will allow our multivector functions to be placed on either or both sides of the differential.

Definition 6.13: Line integral.

Given a single variable parameterization $x = x(u)$, we write $d^1\mathbf{x} = \mathbf{x}_u du$, and call

$$\int F d^1\mathbf{x}\, G,$$

a *line integral*, where F, G are arbitrary multivector functions.

We must be careful not to reorder any of the factors in the integrand, since the differential may not commute with either F or G. Here is a simple example where the integrand has a product of a vector and differential.

Exercise 6.15 Circular parameterization.

Given a circular parameterization $x(\theta) = \gamma_1 e^{-i\theta}$, where $i = \gamma_1\gamma_2$, the unit bivector for the x, y plane. Compute the line integral

$$\int_0^{\pi/4} F(\theta)\, d^1\mathbf{x}\, G(\theta), \tag{6.129}$$

where $F(\theta) = \mathbf{x}^\theta + \gamma_3 + \gamma_1\gamma_0$ is a multivector valued function, and $G(\theta) = \gamma_0$ is vector valued.

Exercise 6.16 Line integral for boosted time direction vector.

Let $x = e^{\hat{\mathbf{v}}\alpha/2}\gamma_0 e^{-\hat{\mathbf{v}}\alpha/2}$ represent the spacetime curve of all the boosts of γ_0 along a specific velocity direction vector, where $\hat{\mathbf{v}} = (v \wedge \gamma_0)/\|v \wedge \gamma_0\|$ is a unit spatial bivector for any constant vector v. Compute the line integral

$$\int x\, d^1\mathbf{x}. \tag{6.132}$$

6.5.3 *Perfect differentials.*

Having seen a couple examples of multivector line integrals, let's now move on to figure out the structure of a line integral that has a "perfect" differential integrand. We can take a hint from the $\mathbb{R}^3$ vector result that we already know, namely

$$\int_A^B d\mathbf{l} \cdot \boldsymbol{\nabla} f = f(B) - f(A). \tag{6.137}$$

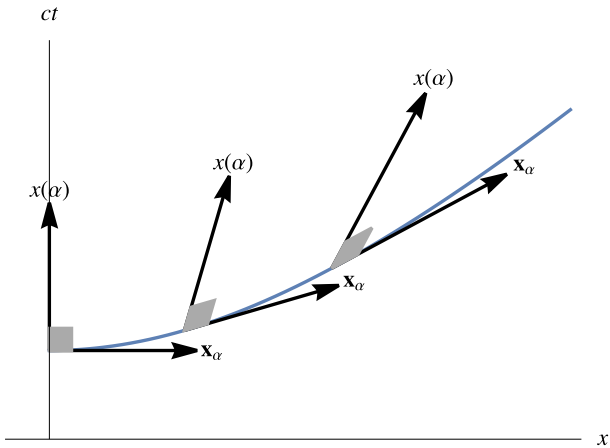


Figure 6.6: Tangent perpendicularity in mixed metric.

It seems reasonable to guess that the relativistic generalization of this is

$$\int_A^B dx \cdot \nabla f = f(B) - f(A). \tag{6.138}$$

Let's check that, by expanding in coordinates

$$\begin{aligned}\int_A^B dx \cdot \nabla f &= \int_A^B d\tau \frac{dx^\mu}{d\tau} \partial_\mu f \\ &= \int_A^B d\tau \frac{dx^\mu}{d\tau} \frac{\partial f}{\partial x^\mu} \\ &= \int_A^B d\tau \frac{df}{d\tau} \\ &= f(B) - f(A).\end{aligned} \tag{6.139}$$

If we drop the dot product, will we have such a nice result? Let's see:

$$\begin{aligned}\int_A^B dx \nabla f &= \int_A^B d\tau \frac{dx^\mu}{d\tau} \gamma_\mu \gamma^\nu \partial_\nu f \\ &= \int_A^B d\tau \frac{dx^\mu}{d\tau} \frac{\partial f}{\partial x^\mu} + \int_A^B d\tau \sum_{\mu \neq \nu} \gamma_\mu \gamma^\nu \frac{dx^\mu}{d\tau} \frac{\partial f}{\partial x^\nu}.\end{aligned} \tag{6.140}$$

This scalar component of this integrand is a perfect differential, but the bivector part of the integrand is a complete mess, that we have no hope of generally integrating. It happens that if we consider one of the simplest

parameterization examples, we can get a strong hint of how to generalize the differential operator to one that ends up providing a perfect differential. In particular, let's integrate over a linear constant path, such as $x(\tau) = \tau\gamma_0$. For this path, we have

$$\begin{aligned}\int_A^B dx\nabla f &= \int_A^B \gamma_0 d\tau \left(\gamma^0\partial_0 + \gamma^1\partial_1 + \gamma^2\partial_2 + \gamma^3\partial_3\right) f \\ &= \int_A^B d\tau \left(\frac{\partial f}{\partial \tau} + \gamma_0\gamma^1\frac{\partial f}{\partial x^1} + \gamma_0\gamma^2\frac{\partial f}{\partial x^2} + \gamma_0\gamma^3\frac{\partial f}{\partial x^3}\right).\end{aligned} \tag{6.141}$$

Just because the path does not have any x^1, x^2, x^3 component dependencies does not mean that these last three partials are neccessarily zero. For example $f = f(x(\tau)) = \left(x^0\right)^2\gamma_0 + x^1\gamma_1$ will have a non-zero contribution from the ∂_1 operator. In that particular case, we can easily integrate f, but we have to know the specifics of the function to do the integral. However, if we had a differential operator that did not include any component off the integration path, we would ahve a perfect differential. That is, if we were to replace the gradient with the projection of the gradient onto the tangent space, we would have a perfect differential. We see that the function of the dot product in eq. (6.138) has the same effect, as it rejects any component of the gradient that does not lie on the tangent space.

Definition 6.14: Vector derivative.

Given a spacetime manifold parameterized by $x = x(u^0, \cdots u^{N-1})$, with tangent vectors $\mathbf{x}_\mu = \partial x/\partial u^\mu$, and reciprocal vectors $\mathbf{x}^\mu \in \text{span}\{\mathbf{x}_\nu\}$, such that $\mathbf{x}^\mu \cdot \mathbf{x}_\nu = \delta^\mu{}_\nu$, the *vector derivative* is defined as

$$\partial = \sum_{\mu=0}^{N-1} \mathbf{x}^\mu \frac{\partial}{\partial u^\mu}.$$

Observe that if this is a full parameterization of the space ($N = 4$), then the vector derivative is identical to the gradient. The vector derivative is the projection of the gradient onto the tangent space at the point of evaluation.

Furthermore, we designate $\overleftrightarrow{\partial}$ as the vector derivative allowed to act bidirectionally, as follows

$$R \overleftrightarrow{\partial} S = R\mathbf{x}^\mu \frac{\partial S}{\partial u^\mu} + \frac{\partial R}{\partial u^\mu}\mathbf{x}^\mu S,$$

where R, S are multivectors, and summation convention is implied. In this bidirectional action, the vector factors of the vector derivative must stay in place (as they do not neccessarily commute with R, S), but the derivative operators apply in a chain rule like fashion to both functions.

Noting that $\mathbf{x}_u \cdot \nabla = \mathbf{x}_u \cdot \partial$, we may rewrite the scalar line integral identity eq. (6.138) as

$$\int_A^B dx \cdot \partial f = f(B) - f(A). \tag{6.142}$$

However, as our example hinted at, the fundamental theorem for line integrals has a multivector generalization that does not rely on a dot product to do the tangent space filtering, and is more powerful. That generalization has the following form.

Theorem 6.14: Fundamental theorem for line integrals.

Given multivector functions F, G, and a single parameter curve $x(u)$ with line element $d^1\mathbf{x} = \mathbf{x}_u du$, then

$$\int_A^B F d^1\mathbf{x} \overleftrightarrow{\partial} G = F(B)G(B) - F(A)G(A).$$

Proof. Writing out the integrand explicitly, we find

$$\int_A^B F d^1\mathbf{x} \overleftrightarrow{\partial} G = \int_A^B F d\alpha\, \mathbf{x}_\alpha \mathbf{x}^\alpha \frac{\overleftrightarrow{\partial}}{\partial \alpha} G \tag{6.143}$$

However for a single parameter curve, we have $\mathbf{x}^\alpha = 1/\mathbf{x}_\alpha$, so we are left with

$$\begin{aligned} \int_A^B F d^1\mathbf{x} \overleftrightarrow{\partial} G &= \int_A^B d\alpha \frac{\partial(FG)}{\partial \alpha} \\ &= FG|_B - FG|_A. \end{aligned} \tag{6.144}$$

□

6.6 RELATIVISTIC MULTIVECTOR SURFACE INTEGRALS.

We've now covered line integrals and the fundamental theorem for line integrals, so it's now time to move on to surface integrals.

Definition 6.15: Surface integral.

Given a two variable parameterization $x = x(u, v)$, we write $d^2\mathbf{x} = \mathbf{x}_u \wedge \mathbf{x}_v du dv$, and call

$$\int F d^2\mathbf{x}\, G,$$

a *surface integral*, where F, G are arbitrary multivector functions.

Like our multivector line integral, this is intrinsically multivector valued, with a product of F with arbitrary grades, a bivector $d^2\mathbf{x}$, and G, also potentially with arbitrary grades. Let's consider an example.

Exercise 6.17 Surface area integral example.

Given the hyperbolic surface parameterization $x(\rho, \alpha) = \rho\gamma_0 e^{-\hat{\mathbf{v}}\alpha}$, where $\hat{\mathbf{v}} = \gamma_{20}$ evaluate the indefinite integral

$$\int \gamma_1 e^{\gamma_{21}\alpha} d^2\mathbf{x}\, \gamma_2. \tag{6.145}$$

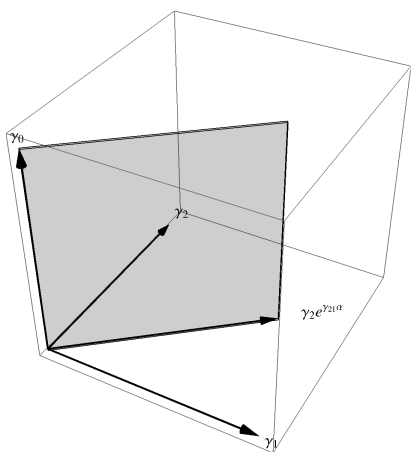


Figure 6.7: Spacetime plane.

6.6.1 *Fundamental theorem for surfaces.*

For line integrals we saw that $d\mathbf{x} \cdot \nabla = \langle d\mathbf{x}\partial \rangle$, and obtained the fundamental theorem for multivector line integrals by omitting the grade selection and using the multivector operator $d\mathbf{x}\partial$ in the integrand directly. We have the same situation for surface integrals. In particular, we know that the $\mathbb{R}^3$ Stokes theorem can be expressed in terms of $d^2\mathbf{x} \cdot \boldsymbol{\nabla}$

Exercise 6.18 GA form of 3D Stokes' theorem integrand.

Given an $\mathbb{R}^3$ vector field $\mathbf{f}$, show that

$$\int dA\hat{\mathbf{n}} \cdot (\boldsymbol{\nabla} \times \mathbf{f}) = -\int \left(d^2\mathbf{x} \cdot \boldsymbol{\nabla}\right) \cdot \mathbf{f}. \tag{6.148}$$

The moral of the story is that the conventional dual form of the $\mathbb{R}^3$ Stokes' theorem can be written directly by projecting the gradient onto the surface area element. Geometrically, this projection operation has a rotational effect as well, since for bivector B, and vector x, the bivector-vector dot product $B \cdot x$ is the component of x that lies in the plane $B \wedge x = 0$, but also rotated 90 degrees.

For multivector integration, we do not want an integral operator that includes such dot products. In the line integral case, we were able to achieve the same projective operation by using vector derivative instead of a dot product, and can do the same for the surface integral case. In particular

Lemma 6.2: Projection of gradient onto the tangent space.

Given a curvilinear representation of the gradient with respect to parameters u^0, u^1, u^2, u^3

$$\nabla = \sum_\mu \mathbf{x}^\mu \frac{\partial}{\partial u^\mu},$$

the surface projection onto the tangent space associated with any two of those parameters, satisfies

$$d^2\mathbf{x} \cdot \nabla = \left\langle d^2\mathbf{x}\partial \right\rangle_1.$$

Proof. Without loss of generality, we may pick u^0, u^1 as the parameters associated with the tangent space. The area element for the surface is

$$d^2\mathbf{x} = \mathbf{x}_0 \wedge \mathbf{x}_1 \, du^0 du^1. \tag{6.150}$$

Dotting this with the gradient gives

$$\begin{aligned} d^2\mathbf{x} \cdot \nabla &= du^0 du^1 \, (\mathbf{x}_0 \wedge \mathbf{x}_1) \cdot \mathbf{x}^\mu \frac{\partial}{\partial u^\mu} \\ &= du^0 du^1 \, (\mathbf{x}_0 \, (\mathbf{x}_1 \cdot \mathbf{x}^\mu) - \mathbf{x}_1 \, (\mathbf{x}_0 \cdot \mathbf{x}^\mu)) \frac{\partial}{\partial u^\mu} \\ &= du^0 du^1 \left(\mathbf{x}_0 \frac{\partial}{\partial u^1} - \mathbf{x}_0 \frac{\partial}{\partial u^1} \right). \end{aligned} \tag{6.151}$$

On the other hand, the vector derivative for this surface is

$$\partial = \mathbf{x}^0 \frac{\partial}{\partial u^0} + \mathbf{x}^1 \frac{\partial}{\partial u^1}, \tag{6.152}$$

so

$$\begin{aligned} \left\langle d^2\mathbf{x}\partial \right\rangle_1 &= du^0 du^1 \, (\mathbf{x}_0 \wedge \mathbf{x}_1) \cdot \left(\mathbf{x}^0 \frac{\partial}{\partial u^0} + \mathbf{x}^1 \frac{\partial}{\partial u^1} \right) \\ &= du^0 du^1 \left(\mathbf{x}_0 \frac{\partial}{\partial u^1} - \mathbf{x}_1 \frac{\partial}{\partial u^0} \right). \end{aligned} \tag{6.153}$$

□

We now want to formulate the geometric algebra form of the fundamental theorem for surface integrals.

Theorem 6.15: Fundamental theorem for surface integrals.

Given multivector functions F, G, and surface area element $d^2\mathbf{x} = (\mathbf{x}_u \wedge \mathbf{x}_v) \, dudv$, associated with a two parameter curve $x(u, v)$, then

$$\int_S F d^2\mathbf{x} \overleftrightarrow{\partial} G = \int_{\partial S} F d^1\mathbf{x} G,$$

where S is the integration surface, and ∂S designates its boundary, and the line integral on the RHS is really short hand for

$$\int \left(F(-d\mathbf{x}_v)G \right)\Big|_{\Delta u} + \int \left(F(d\mathbf{x}_u)G \right)\Big|_{\Delta v},$$

which is a line integral that traverses the boundary of the surface with the opposite orientation to the circulation of the area element.

Proof. The vector derivative for this surface is

$$\partial = \mathbf{x}^u \frac{\partial}{\partial u} + \mathbf{x}^v \frac{\partial}{\partial v}, \tag{6.154}$$

so

$$F d^2\mathbf{x} \overleftrightarrow{\partial} G = \frac{\partial}{\partial u}\left(F d^2\mathbf{x}\, \mathbf{x}^u G\right) + \frac{\partial}{\partial v}\left(F d^2\mathbf{x}\, \mathbf{x}^v G\right), \tag{6.155}$$

where $d^2\mathbf{x}\,\mathbf{x}^u$ is held constant with respect to u, and $d^2\mathbf{x}\,\mathbf{x}^v$ is held constant with respect to v (since the partials of the vector derivative act on F, G, but not on the area element, nor on the reciprocal vectors of $\overleftrightarrow{\partial}$ itself.) Note that

$$d^2\mathbf{x} \wedge \mathbf{x}^u = du dv\, \left(\mathbf{x}_u \wedge \mathbf{x}_v\right) \wedge \mathbf{x}^u = 0, \tag{6.156}$$

since $\mathbf{x}^u \in \text{span}\{\mathbf{x}_u\, \mathbf{x}_v\}$, so

$$\begin{aligned} d^2\mathbf{x}\,\mathbf{x}^u &= d^2\mathbf{x} \cdot \mathbf{x}^u + d^2\mathbf{x} \wedge \mathbf{x}^u \\ &= d^2\mathbf{x} \cdot \mathbf{x}^u \\ &= du dv\, (\mathbf{x}_u \wedge \mathbf{x}_v) \cdot \mathbf{x}^u \\ &= -du dv\, \mathbf{x}_v. \end{aligned} \tag{6.157}$$

Similarly

$$\begin{aligned} d^2\mathbf{x}\,\mathbf{x}^v &= d^2\mathbf{x} \cdot \mathbf{x}^v \\ &= du dv\, (\mathbf{x}_u \wedge \mathbf{x}_v) \cdot \mathbf{x}^v \\ &= du dv\, \mathbf{x}_u. \end{aligned} \tag{6.158}$$

This leaves us with

$$F d^2\mathbf{x} \overleftrightarrow{\partial} G = -du dv \frac{\partial}{\partial u}(F\mathbf{x}_v G) + du dv \frac{\partial}{\partial v}(F\mathbf{x}_u G), \tag{6.159}$$

where $\mathbf{x}_v, \mathbf{x}_u$ are held constant with respect to u, v respectively. Fortuitously, this constant condition can be dropped, since the antisymmetry of the wedge in the area element results in perfect cancellation. If these line elements are not held constant then

$$\begin{aligned} \frac{\partial}{\partial u}(F\mathbf{x}_v G) - \frac{\partial}{\partial v}(F\mathbf{x}_u G) = F\left(\frac{\partial \mathbf{x}_u}{\partial v} - \frac{\partial \mathbf{x}_v}{\partial u}\right) G + \left(\frac{\partial F}{\partial u}\mathbf{x}_v G + F\mathbf{x}_v \frac{\partial G}{\partial u}\right) \\ + \left(\frac{\partial F}{\partial v}\mathbf{x}_u G + F\mathbf{x}_u \frac{\partial G}{\partial v}\right), \end{aligned} \tag{6.160}$$

but the mixed partial contribution is zero

$$\frac{\partial \mathbf{x}_u}{\partial v} - \frac{\partial \mathbf{x}_v}{\partial u} = \frac{\partial}{\partial v}\frac{\partial x}{\partial u} - \frac{\partial}{\partial u}\frac{\partial x}{\partial v} = 0, \tag{6.161}$$

by equality of mixed partials. We have two perfect differentials, and can evaluate each of these integrals

$$\begin{aligned}\int F d^2\mathbf{x} \overset{\leftrightarrow}{\partial} G &= -\int dudv \frac{\partial}{\partial u}(F\mathbf{x}_v G) + \int dudv \frac{\partial}{\partial v}(F\mathbf{x}_u G) \\ &= -\int dv\, (F\mathbf{x}_v G)|_{\Delta u} + \int du\, (F\mathbf{x}_u G)|_{\Delta v} \\ &= \int (F(-d\mathbf{x}_v)G)|_{\Delta u} + \int (F(d\mathbf{x}_u)G)|_{\Delta v}.\end{aligned} \tag{6.162}$$

We use the shorthand $d^1\mathbf{x} = d\mathbf{x}_u - d\mathbf{x}_v$ to write

$$\int_S F d^2\mathbf{x} \overset{\leftrightarrow}{\partial} G = \int_{\partial S} F d^1\mathbf{x} G, \tag{6.163}$$

with the understanding that this is really instructions to evaluate the line integrals in the last step of eq. (6.162). □

Exercise 6.19 Integration in the t,y plane.

Let $x(t, y) = ct\gamma_0 + y\gamma_2$. Write out both sides of the fundamental theorem explicitly.

Exercise 6.20 A cylindrical hyperbolic surface.

Generalizing the example surface integral from exercise 6.17, let

$$x(\rho, \alpha) = \rho e^{-\hat{\mathbf{v}}\alpha/2} x(0, 1) e^{\hat{\mathbf{v}}\alpha/2}, \tag{6.170}$$

where ρ is a scalar, and $\hat{\mathbf{v}} = \cos\theta_k \gamma_{k0}$ is a unit spatial bivector, and $\cos\theta_k$ are direction cosines of that vector. This is a composite transformation, where the α variation boosts the $x(0, 1)$ four-vector, and the ρ parameter contracts or increases the magnitude of this vector, resulting in x spanning a hyperbolic region of spacetime.

Compute the tangent and reciprocal basis, the area element for the surface, and explicitly state both sides of the fundamental theorem.

Exercise 6.21 Non-orthogonal tangent space example.

Let $x(u, v) = ua + vb$, where u, v are scalar parameters, and a, b are non-null and non-colinear constant four-vectors. Write out the fundamental theorem for surfaces with respect to this parameterization.

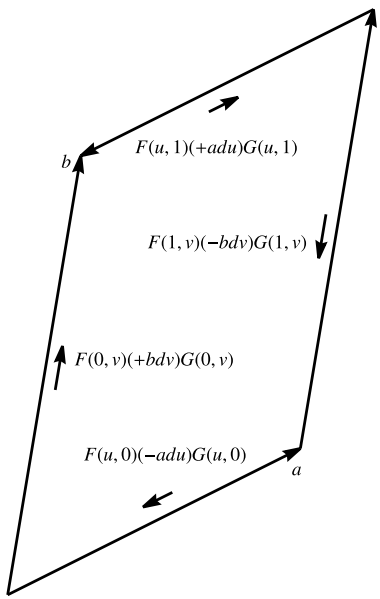


Figure 6.8: Line integral orientation.

7 RELATIVISTIC ACTION.

7.1 IN THIS CHAPTER.

This chapter will cover

- derivation of the relativistic form of the Euler-Lagrange equations from the covariant form of the action,
- relationship of the STA form of the Euler-Lagrange equations to their tensor equivalents,
- derivation of the Lorentz force equation from the STA Lorentz force Lagrangian,
- relationship of the STA Lorentz force equation to its equivalent in the tensor formalism,
- relationship of the STA Lorentz force equation to the traditional vector form.

Note that some of the prerequisite ideas and auxiliary details are presented as problems with solutions, all of which the reader is encouraged to try before looking at the solutions.

7.2 EULER-LAGRANGE EQUATIONS.

I'll start at ground zero, with the derivation of the relativistic form of the Euler-Lagrange equations from the action. A relativistic action for a single particle system has the form

$$S = \int d\tau L(x, \dot{x}), \tag{7.1}$$

where x is the spacetime coordinate, $\dot{x} = dx/d\tau$ is the four-velocity, and τ is proper time.

Theorem 7.1: Relativistic Euler-Lagrange equations.

Let $x \to x + \delta x$ be any variation of the Lagrangian four-vector coordinates, where $\delta x = 0$ at the boundaries of the action integral. The variation of the action is

$$\delta S = \int d\tau \delta x \cdot \delta L(x, \dot{x}),$$

where

$$\delta L = \nabla L - \frac{d}{d\tau}(\nabla_v L),$$

where $\nabla = \gamma^\mu \partial_\mu$ (per definition 6.5), and where we construct a similar velocity-gradient with respect to the proper-time derivatives of the coordinates $\nabla_v = \gamma^\mu \partial / \partial \dot{x}^\mu$.

The action is extremized when $\delta S = 0$, or when $\delta L = 0$. This latter condition is called the Euler-Lagrange equations.

Proof. Let $\epsilon = \delta x$, and expand the Lagrangian in Taylor series to first order

$$\begin{aligned} S &\to S + \delta S \\ &= \int d\tau L(x + \epsilon, \dot{x} + \dot{\epsilon}) \\ &= \int d\tau \left(L(x, \dot{x}) + \epsilon \cdot \nabla L + \dot{\epsilon} \cdot \nabla_v L \right). \end{aligned} \tag{7.2}$$

Subtracting off S and integrating by parts, leaves

$$\delta S = \int d\tau \epsilon \cdot \left(\nabla L - \frac{d}{d\tau} \nabla_v L \right) + \int d\tau \frac{d}{d\tau}(\nabla_v L) \cdot \epsilon. \tag{7.3}$$

The boundary integral

$$\int d\tau \frac{d}{d\tau}(\nabla_v L) \cdot \epsilon = (\nabla_v L) \cdot \epsilon|_{\Delta\tau} = 0, \tag{7.4}$$

is zero since the variation ϵ is required to vanish on the boundaries. So, if $\delta S = 0$, we must have

$$0 = \int d\tau \epsilon \cdot \left(\nabla L - \frac{d}{d\tau} \nabla_v L \right), \tag{7.5}$$

for all variations ϵ. Clearly, this requires that

$$\delta L = \nabla L - \frac{d}{d\tau}(\nabla_v L) = 0, \tag{7.6}$$

or

$$\nabla L = \frac{d}{d\tau}(\nabla_v L), \tag{7.7}$$

which is the coordinate free statement of the Euler-Lagrange equations. □

Exercise 7.1 Coordinate form of the Euler-Lagrange equations.

Working in coordinates, use the action argument show that the Euler-Lagrange equations have the form

$$\frac{\partial L}{\partial x^\mu} = \frac{d}{d\tau}\frac{\partial L}{\partial \dot{x}^\mu}$$

Observe that this is identical to the statement of theorem 7.1 after contraction with γ^μ.

7.3 LORENTZ FORCE EQUATION.

Theorem 7.2: Lorentz force.

The relativistic Lagrangian for a charged particle is

$$L = \frac{1}{2}mv^2 + qA \cdot v/c.$$

Application of the Euler-Lagrange equations to this Lagrangian yields the Lorentz-force equation

$$\frac{dp}{d\tau} = qF \cdot v/c,$$

where $p = mv$ is the proper momentum, F is the Faraday bivector $F = \nabla \wedge A$, and c is the speed of light.

Proof. To make life easier, let's take advantage of the linearity of the Lagrangian, and break it into the free particle Lagrangian $L_0 = (1/2)mv^2$ and a potential term $L_1 = qA \cdot v/c$. For the free particle case we have

$$\begin{aligned}\delta L_0 &= \nabla L_0 - \frac{d}{d\tau}(\nabla_v L_0) \\ &= -\frac{d}{d\tau}(mv) \\ &= -\frac{dp}{d\tau}.\end{aligned} \tag{7.11}$$

For the potential contribution we have

$$\begin{aligned}\delta L_1 &= \nabla L_1 - \frac{d}{d\tau}(\nabla_v L_1) \\ &= \frac{q}{c}\left(\nabla(A \cdot v) - \frac{d}{d\tau}\left(\nabla_v(A \cdot v)\right)\right) \\ &= \frac{q}{c}\left(\nabla(A \cdot v) - \frac{dA}{d\tau}\right).\end{aligned} \tag{7.12}$$

The proper time derivative can be evaluated using the chain rule

$$\begin{aligned}\frac{dA}{d\tau} &= \frac{\partial x^\mu}{\partial \tau}\partial_\mu A \\ &= (v \cdot \nabla)A.\end{aligned} \tag{7.13}$$

Putting all the pieces back together we have

$$\begin{aligned}0 &= \delta L \\ &= -\frac{dp}{d\tau} + \frac{q}{c}\left(\nabla(A \cdot v) - (v \cdot \nabla)A\right) \\ &= -\frac{dp}{d\tau} + \frac{q}{c}\left(\nabla \wedge A\right) \cdot v.\end{aligned} \tag{7.14}$$

□

Exercise 7.2 Gradient of a squared position vector.

Show that

$$\nabla(a \cdot x) = a,$$

and

$$\nabla x^2 = 2x.$$

It should be clear that the same ideas can be used for the velocity gradient, where we obtain $\nabla_v(v^2) = 2v$, and $\nabla_v(A \cdot v) = A$, as used in the derivation above.

It is desirable to put this relativistic Lorentz force equation into the usual vector and tensor forms for comparison.

Theorem 7.3: Tensor form of the Lorentz force equation.

The tensor form of the Lorentz force equation is

$$\frac{dp^\mu}{d\tau} = \frac{q}{c} F^{\mu\nu} v_\nu,$$

where the antisymmetric Faraday tensor is defined as $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

Proof. We have only to dot both sides with γ^μ. On the left we have

$$\gamma^\mu \cdot \frac{dp}{d\tau} = \frac{dp^\mu}{d\tau}. \tag{7.17}$$

On the right, we have

$$\begin{aligned}
\gamma^\mu \cdot \left(\frac{q}{c} F \cdot v\right) &= \frac{q}{c}((\nabla \wedge A) \cdot v) \cdot \gamma^\mu \\
&= \frac{q}{c}(\nabla(A \cdot v) - (v \cdot \nabla)A) \cdot \gamma^\mu \\
&= \frac{q}{c}\left((\partial^\mu A^\nu) v_\nu - v_\nu \partial^\nu A^\mu\right) \\
&= \frac{q}{c} F^{\mu\nu} v_\nu.
\end{aligned} \tag{7.18}$$

□

Exercise 7.3 Tensor expansion of F.

An alternate way to demonstrate theorem 7.3 is to first expand $F = \nabla \wedge A$ in terms of coordinates, an expansion that can be expressed in terms of a second rank tensor antisymmetric tensor $F^{\mu\nu}$. Find that expansion, and re-evaluate the dot products of eq. (7.18) using that.

Exercise 7.4 Lorentz force direct tensor derivation.

Instead of using the geometric algebra form of the Lorentz force equation as a stepping stone, we may derive the tensor form from the Lagrangian directly, provided the Lagrangian is put into tensor form

$$L = \frac{1}{2} m v^\mu v_\mu + q A^\mu v_\mu / c.$$

Evaluate the Euler-Lagrange equations in coordinate form and compare to theorem 7.3.

Theorem 7.4: Vector Lorentz force equation.

Relative to a fixed observer's frame, the Lorentz force equation of theorem 7.2 splits into a spatial rate of change of momentum, and (timelike component) rate of change of energy, as follows

$$\begin{aligned}
\frac{d(\gamma m \mathbf{v})}{dt} &= q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \\
\frac{d(\gamma m c^2)}{dt} &= q \mathbf{v} \cdot \mathbf{E},
\end{aligned}$$

where $F = \mathbf{E} + Ic\mathbf{B}$, $\gamma = 1/\sqrt{1 - \mathbf{v}^2/c^2}$.

Proof. The first step is to eliminate the proper time dependencies in the Lorentz force equation. Consider first the coordinate representation of an arbitrary position four-vector x

$$x = ct\gamma_0 + x^k \gamma_k. \tag{7.27}$$

The corresponding four-vector velocity is

$$v = \frac{dx}{d\tau} = c\frac{dt}{d\tau}\gamma_0 + \frac{dt}{d\tau}\frac{dx^k}{dt}\gamma_k. \tag{7.28}$$

By construction, $v^2 = c^2$ is a Lorentz invariant quantity (this is one of the relativistic postulates), so the LHS of eq. (7.28) must have the same square. That is

$$c^2 = \left(\frac{dt}{d\tau}\right)^2 \left(c^2 - \mathbf{v}^2\right), \tag{7.29}$$

where $\mathbf{v} = v \wedge \gamma_0$. This shows that we may make the identification

$$\gamma = \frac{dt}{d\tau} = \frac{1}{1 - \mathbf{v}^2/c^2}, \tag{7.30}$$

and

$$\frac{d}{d\tau} = \frac{dt}{d\tau}\frac{d}{dt} = \gamma\frac{d}{dt}. \tag{7.31}$$

We may now factor the four-velocity v into its spacetime split

$$v = \gamma\left(c + \mathbf{v}\right)\gamma_0. \tag{7.32}$$

In particular the LHS of the Lorentz force equation can be rewritten as

$$\frac{dp}{d\tau} = \gamma\frac{d}{dt}\left(\gamma\left(c + \mathbf{v}\right)\right)\gamma_0, \tag{7.33}$$

and the RHS of the Lorentz force equation can be rewritten as

$$\frac{q}{c}F \cdot v = \frac{\gamma q}{c}F \cdot \left((c + \mathbf{v})\gamma_0\right). \tag{7.34}$$

Equating timelike and spacelike components leaves us

$$\frac{d(m\gamma c)}{dt} = \frac{q}{c}\left(F \cdot \left((c + \mathbf{v})\gamma_0\right)\right) \cdot \gamma_0, \tag{7.35a}$$

$$\frac{d(m\gamma \mathbf{v})}{dt} = \frac{q}{c}\left(F \cdot \left((c + \mathbf{v})\gamma_0\right)\right) \wedge \gamma_0, \tag{7.35b}$$

Evaluating these products requires some care, but is an essentially manual process. The reader is encouraged to do so once, but the end result may also be obtained easily using software (see lorentzForce.nb in [13]). One finds

$$F = \mathbf{E} + Ic\mathbf{B} = E^1\gamma_{10} + E^2\gamma_{20} + E^3\gamma_{30} - cB^1\gamma_{23} - cB^2\gamma_{31} - cB^3\gamma_{12}, \tag{7.36a}$$

$$\frac{q}{c}\left(F \cdot \left((c + \mathbf{v})\gamma_0\right)\right) \cdot \gamma_0 = \frac{q}{c}\mathbf{E} \cdot \mathbf{v}, \tag{7.36b}$$

$$\frac{q}{c}\left(F \cdot \left((c + \mathbf{v})\gamma_0\right)\right) \wedge \gamma_0 = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right). \tag{7.36c}$$

□

Exercise 7.5 Algebraic spacetime split of the Lorentz force equation.

Derive the results of eq. (7.36) algebraically.

Exercise 7.6 Spacetime split of the Lorentz force tensor equation.

Show that theorem 7.4 also follows from the tensor form of the Lorentz force equation (theorem 7.3) provided we identify

$$F^{k0} = E^k, \tag{7.43a}$$

and

$$F^{rs} = -\epsilon^{rst} B^t. \tag{7.43b}$$

Also verify that the identification eq. (7.43) is consistent with the geometric algebra Faraday bivector $F = \mathbf{E} + Ic\mathbf{B}$, and the associated coordinate expansion of the field $F = (1/2)(\gamma_\mu \wedge \gamma_\nu)F^{\mu\nu}$.

Exercise 7.7 Lorentz force gauge transformation.

Show that the gauge transformation $A \to A + \nabla\psi$ applied to the Lorentz force Lagrangian

$$L = \frac{1}{2}mv^2 + qA \cdot v/c, \tag{7.44}$$

does not change the equations of motion.

7.4 SOLUTIONS.

Answer for Exercise 6.6

Let $x = x^\mu \gamma_\mu$, so that

$$\begin{aligned} x \cdot \hat{\mathbf{v}} &= \left\langle x^\mu \gamma_\mu \cos\theta^b \gamma_{b0} \right\rangle_1 \\ &= x^\mu \cos\theta^b \left\langle \gamma_\mu \gamma_{b0} \right\rangle_1 . \end{aligned} \tag{6.38}$$

The $\mu = 0$ component of this grade selection is

$$\langle \gamma_0 \gamma_{b0} \rangle_1 = -\gamma_b, \tag{6.39}$$

and for $\mu = a \neq 0$, we have

$$\langle \gamma_a \gamma_{b0} \rangle_1 = -\delta_{ab}\gamma_0, \tag{6.40}$$

so we have

$$\begin{aligned} x \cdot \hat{\mathbf{v}} &= x^0 \cos\theta^b(-\gamma_b) + x^a \cos\theta^b(-\delta_{ab}\gamma_0) \\ &= -x^0\hat{\mathbf{v}}\gamma_0 - x^b \cos\theta^b\gamma_0 \\ &= -\left(x^0\hat{\mathbf{v}} + \mathbf{x}\cdot\hat{\mathbf{v}}\right)\gamma_0, \end{aligned} \tag{6.41}$$

where $\mathbf{x} = x \wedge \gamma_0$ is the spatial portion of the four vector x relative to the stationary observer frame. Since $\hat{\mathbf{v}}$ anticommutes with γ_0, the component of x in the spacetime plane $\hat{\mathbf{v}}$ is

$$(x \cdot \hat{\mathbf{v}})\,\hat{\mathbf{v}} = \left(x^0 + (\mathbf{x}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}}\right)\gamma_0, \tag{6.42}$$

as expected.

For the rejection term, we have

$$x \wedge \hat{\mathbf{v}} = x^\mu \cos\theta^s \left\langle \gamma_\mu \gamma_{s0} \right\rangle_3. \tag{6.43}$$

The $\mu = 0$ term clearly contributes nothing, leaving us with:

$$\begin{aligned} (x \wedge \hat{\mathbf{v}})\,\hat{\mathbf{v}} &= (x \wedge \hat{\mathbf{v}}) \cdot \hat{\mathbf{v}} \\ &= x^r \cos\theta^s \cos\theta^t\,((\gamma_r \wedge \gamma_s)\,\gamma_0) \cdot (\gamma_{t0}) \\ &= x^r \cos\theta^s \cos\theta^t \langle (\gamma_r \wedge \gamma_s)\,\gamma_0\gamma_{t0}\rangle_1 \\ &= -x^r \cos\theta^s \cos\theta^t\,(\gamma_r \wedge \gamma_s) \cdot \gamma_t \\ &= -x^r \cos\theta^s \cos\theta^t\,(-\gamma_r\delta_{st} + \gamma_s\delta_{rt}) \\ &= x^r \cos\theta^t \cos\theta^t\gamma_r - x^t \cos\theta^s \cos\theta^t\gamma_s \\ &= \mathbf{x}\gamma_0 - (\mathbf{x}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}}\gamma_0 \\ &= (\mathbf{x} \wedge \hat{\mathbf{v}})\,\hat{\mathbf{v}}\gamma_0, \end{aligned} \tag{6.44}$$

as expected. Is there a clever way to demonstrate this without resorting to coordinates?

Answer for Exercise 6.7

This problem is left for the reader, as I don't feel like typing out my solution.

The first part of this problem can be done in the tedious coordinate approach used above, but hopefully there is a better way.

For the last (commutation) part of the problem, here is a hint. Let $x \wedge i = ni$, where $n \cdot i = 0$. The result then follows easily.

Answer for Exercise 6.8

The frame vectors are all easy to compute

$$\begin{aligned}
\mathbf{x}_0 &= \frac{\partial x}{\partial t} = c\gamma_0 \\
\mathbf{x}_1 &= \frac{\partial x}{\partial \rho} = \gamma_1 e^{-i\theta} \\
\mathbf{x}_2 &= \frac{\partial x}{\partial \theta} = -\rho\gamma_1\gamma_1\gamma_2 e^{-i\theta} = \rho\gamma_2 e^{-i\theta} \\
\mathbf{x}_3 &= \frac{\partial x}{\partial z} = \gamma_3.
\end{aligned} \tag{6.57}$$

The $\mathbf{x}_1$ vector is radial, $\mathbf{x}^2$ is perpendicular to that tangent to the same unit circle, as plotted in fig. 6.4. All of these particular frame vectors happen to be mutually perpendicular, something that will not generally be true for a more arbitrary parameterization.

To compute the reciprocal frame vectors, we must express our parameters in terms of x^μ coordinates, and use implicit integration techniques to deal with the coupling of the rotational terms. First observe that

$$\begin{aligned}
\gamma_1 e^{-i\theta} &= \gamma_1 (\cos\theta - \gamma_1\gamma_2 \sin\theta) \\
&= \gamma_1 \cos\theta + \gamma_2 \sin\theta,
\end{aligned} \tag{6.58}$$

so $x = x^\mu \gamma_\mu = ct\gamma_0 + \gamma_1 \rho e^{-i\theta} + z\gamma_3$ is equivalent to the following set of coordinate equations

$$\begin{aligned}
x^0 &= ct \\
x^1 &= \rho\cos\theta \\
x^2 &= \rho\sin\theta \\
x^3 &= z.
\end{aligned} \tag{6.59}$$

We can easily evaluate the t, z gradients

$$\begin{aligned}
\nabla t &= \frac{\gamma^1}{c} \\
\nabla z &= \gamma^3,
\end{aligned} \tag{6.60}$$

but the ρ, θ gradients are not as easy. First writing

$$\rho^2 = \left(x^1\right)^2 + \left(x^2\right)^2, \tag{6.61}$$

we find

$$\begin{aligned}
2\rho\nabla\rho &= 2\left(x^1\nabla x^1 + x^2\nabla x^2\right) \\
&= 2\rho\left(\cos\theta\gamma^1 + \sin\theta\gamma^2\right) \\
&= 2\rho\gamma^1\left(\cos\theta + \gamma_1\gamma^2\sin\theta\right) \\
&= 2\rho\gamma^1 e^{-i\theta},
\end{aligned} \tag{6.62}$$

so

$$\nabla\rho = \gamma^1 e^{-i\theta}. \tag{6.63}$$

For the θ gradient, we can write

$$\tan\theta = \frac{x^2}{x^1}, \tag{6.64}$$

so

$$\begin{aligned}
\frac{1}{\cos^2\theta}\nabla\theta &= \frac{\gamma^2}{x^1} - x^2\frac{\gamma^1}{(x^1)^2} \\
&= \frac{1}{(x^1)^2}\left(\gamma^2 x^1 - \gamma^1 x^2\right) \\
&= \frac{\rho}{\rho^2\cos^2\theta}\left(\gamma^2\cos\theta - \gamma^1\sin\theta\right) \\
&= \frac{1}{\rho\cos^2\theta}\gamma^2\left(\cos\theta - \gamma_2\gamma^1\sin\theta\right) \\
&= \frac{\gamma^2 e^{-i\theta}}{\rho\cos^2\theta},
\end{aligned} \tag{6.65}$$

or

$$\nabla\theta = \frac{1}{\rho}\gamma^2 e^{-i\theta}. \tag{6.66}$$

In summary,

$$\begin{aligned}
\mathbf{x}^0 &= \frac{\gamma^0}{c} \\
\mathbf{x}^1 &= \gamma^1 e^{-i\theta} \\
\mathbf{x}^2 &= \frac{1}{\rho}\gamma^2 e^{-i\theta} \\
\mathbf{x}^3 &= \gamma^3.
\end{aligned} \tag{6.67}$$

Answer for Exercise 6.9

We expanded $\mathbf{x}_1$ explicitly in eq. (6.58). Doing the same for $\mathbf{x}_2$, we have

$$\begin{aligned}\mathbf{x}_2 &= \rho\gamma_2 e^{i\theta} \\ &= \rho\gamma_2 \left(\cos\theta - \gamma_1\gamma_2 \sin\theta\right) \\ &= \rho\left(\gamma_2 \cos\theta - \gamma_1 \sin\theta\right).\end{aligned} \tag{6.72}$$

Reading off the coordinates of our frame vectors, we have

$$X = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & C & -\rho S & 0 \\ 0 & S & \rho C & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6.73}$$

where $C = \cos\theta$ and $S = \sin\theta$. We want

$$Y = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & -C & -S & 0 \\ 0 & \rho S & -\rho C & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{c} & 0 & 0 & 0 \\ 0 & -C & \frac{S}{\rho} & 0 \\ 0 & -S & -\frac{C}{\rho} & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{6.74}$$

We can read off the coordinates of the reciprocal frame vectors

$$\begin{aligned}\mathbf{x}^0 &= \frac{1}{c}\gamma_0 \\ \mathbf{x}^1 &= -\cos\theta\gamma_1 - \sin\theta\gamma_2 \\ \mathbf{x}^2 &= \frac{1}{\rho}\left(\sin\theta\gamma_1 + \cos\theta\gamma_2\right) \\ \mathbf{x}^3 &= -\gamma_3.\end{aligned} \tag{6.75}$$

Factoring out γ^1 from the $\mathbf{x}^1$ terms, we find

$$\begin{aligned}\mathbf{x}^1 &= -\cos\theta\gamma_1 - \sin\theta\gamma_2 \\ &= \gamma^1\left(\cos\theta - \gamma_1\gamma_2 \sin\theta\right) \\ &= \gamma^1 e^{-i\theta}.\end{aligned} \tag{6.76}$$

Similarly for $\mathbf{x}^2$,

$$\begin{aligned}\mathbf{x}^2 &= \frac{1}{\rho}(\sin\theta\gamma_1 - \cos\theta\gamma_2) \\ &= \frac{\gamma^2}{\rho}(\sin\theta\gamma_2\gamma_1 + \cos\theta) \\ &= \frac{\gamma^2}{\rho}e^{-i\theta}.\end{aligned} \tag{6.77}$$

This matches eq. (6.67), as expected, but required only algebraic work to compute.

Answer for Exercise 6.10

To characterize the vectors, we square them

$$a^2 = b^2 = \gamma_0^2 + \gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1 - 3 = -2, \tag{6.94}$$

so a, b are both spacelike vectors. The tangent space is clearly just span $\{a, b\}$ = span $\{e, f\}$ where

$$\begin{aligned}e &= \gamma_0 + \gamma_2 \\ f &= \gamma_1 + \gamma_3.\end{aligned} \tag{6.95}$$

Observe that $a = e + f, b = e - f$, and e is lightlike ($e^2 = 0$), whereas f is spacelike ($f^2 = -2$), and $e \cdot f = 0$, so $ef = -fe$. The bivector for the tangent plane is

$$\begin{aligned}\langle ab\rangle_2 &= \langle (e+f)(e-f)\rangle_2 \\ &= \left\langle e^2 - f^2 - 2ef\right\rangle_2 \\ &= -2ef,\end{aligned} \tag{6.96}$$

where

$$ef = \gamma_{01} + \gamma_{21} + \gamma_{23} + \gamma_{03}. \tag{6.97}$$

Because e is lightlike (zero square), and $ef = -fe$, the bivector ef squares to zero

$$(ef)^2 = -e^2 f^2 = 0, \tag{6.98}$$

which shows that the parameterization is degenerate.

This parameterization can also be expressed as

$$\begin{aligned} x(u, v) &= u(e + f) + v(e - f) \\ &= (u + v)e + (u - v)f, \end{aligned} \tag{6.99}$$

a linear combination of a lightlike and spacelike vector. Intuitively, we expect that a physically meaningful spacetime surface involves linear combinations spacelike vectors, or combinations of a timelike vector with spacelike vectors. This beastie is something entirely different.

Answer for Exercise 6.11

Here we have just a single tangent space direction (a line in spacetime) with tangent vector

$$\begin{aligned} \mathbf{x}_u &= a\frac{\partial f}{\partial u} \\ &= af_u, \end{aligned} \tag{6.100}$$

so we see that the tangent space vectors are just rescaled values of the direction vector a. This is a simple enough parameterization that we can compute the reciprocal frame vector explicitly using the gradient. We expect that $\mathbf{x}^u = 1/\mathbf{x}_u$, and find

$$\frac{1}{a} \cdot x = f(u), \tag{6.101}$$

but for constant a, we know that $\nabla a \cdot x = a$, so taking gradients of both sides we find

$$\frac{1}{a} = \nabla f = \frac{\partial f}{\partial u}\nabla u, \tag{6.102}$$

so the reciprocal vector is

$$\mathbf{x}^u = \nabla u = \frac{1}{af_u}, \tag{6.103}$$

as expected.

Answer for Exercise 6.12

The frame vectors are easy to compute, as they are just

$$\begin{aligned} \mathbf{x}_u &= \frac{\partial x}{\partial u} = a \\ \mathbf{x}_v &= \frac{\partial x}{\partial v} = b. \end{aligned} \tag{6.104}$$

This is an example of a parametric equation that we can easily invert, as we have

$$\begin{aligned} x \wedge a &= -v\left(a \wedge b\right) \\ x \wedge b &= u\left(a \wedge b\right), \end{aligned} \tag{6.105}$$

so

$$\begin{aligned} u &= \frac{1}{a \wedge b} \cdot \left(x \wedge b\right) \\ &= \frac{1}{\left(a \wedge b\right)^2} \left(a \wedge b\right) \cdot \left(x \wedge b\right) \\ &= \frac{\left(b \cdot x\right)\left(a \cdot b\right) - \left(a \cdot x\right)\left(b \cdot b\right)}{\left(a \wedge b\right)^2} \end{aligned} \tag{6.106}$$

$$\begin{aligned} v &= -\frac{1}{a \wedge b} \cdot \left(x \wedge a\right) \\ &= -\frac{1}{\left(a \wedge b\right)^2} \left(a \wedge b\right) \cdot \left(x \wedge a\right) \\ &= -\frac{\left(b \cdot x\right)\left(a \cdot a\right) - \left(a \cdot x\right)\left(a \cdot b\right)}{\left(a \wedge b\right)^2} \end{aligned} \tag{6.107}$$

Recall that $\nabla\left(a \cdot x\right) = a$, if a is a constant, so our gradients are just

$$\begin{aligned} \nabla u &= \frac{b\left(a \cdot b\right) - a\left(b \cdot b\right)}{\left(a \wedge b\right)^2} \\ &= b \cdot \frac{1}{a \wedge b}, \end{aligned} \tag{6.108}$$

and

$$\begin{aligned} \nabla v &= -\frac{b\left(a \cdot a\right) - a\left(a \cdot b\right)}{\left(a \wedge b\right)^2} \\ &= -a \cdot \frac{1}{a \wedge b}. \end{aligned} \tag{6.109}$$

Expressed in terms of the frame vectors, this is just

$$\begin{aligned} \mathbf{x}^u &= \mathbf{x}_v \cdot \frac{1}{\mathbf{x}_u \wedge \mathbf{x}_v} \\ \mathbf{x}^v &= -\mathbf{x}_u \cdot \frac{1}{\mathbf{x}_u \wedge \mathbf{x}_v}, \end{aligned} \tag{6.110}$$

so we were able to show, for this special two parameter linear case, that the explicit evaluation of the gradients has the exact structure that we intuited that the reciprocals must have, provided they are constrained to the spacetime plane $a \wedge b$. It is interesting to observe how this structure falls out of the linear system solution so directly. Also note that these reciprocals are not defined at the origin of the (u, v) parameter space.

Answer for Exercise 6.13

$$\begin{aligned}
\mathbf{x}_u &= \frac{\partial x}{\partial u} = 2ua \\
\mathbf{x}_v &= \frac{\partial x}{\partial v} = 2vb.
\end{aligned} \tag{6.111}$$

Our tangent space is still the $a \wedge b$ plane (as is the surface itself), but the spacing of the cells starts getting wider in proportion to u, v. Utilizing the work from the previous problem, we have

$$\begin{aligned}
2u\nabla u &= b \cdot \frac{1}{a \wedge b} \\
2v\nabla v &= -a \cdot \frac{1}{a \wedge b}.
\end{aligned} \tag{6.112}$$

A bit of rearrangement, shows that we recover eq. (6.110) easily. This is a second demonstration that the gradient and the algebraic formulations for the reciprocals match, at least for these special cases of linear non-coupled parameterizations.

Answer for Exercise 6.14

The tangent space vectors are

$$\mathbf{x}_\rho = \frac{x}{\rho}, \tag{6.113}$$

and

$$\begin{aligned}
\mathbf{x}_\theta &= -\frac{i}{2}x + x\frac{i}{2} \\
&= x \cdot i.
\end{aligned} \tag{6.114}$$

Recall that $x \cdot i$ lies perpendicular to x (in the plane i), as illustrated in fig. 6.1. This means that $\mathbf{x}_\rho$ and $\mathbf{x}_\theta$ are orthogonal, so we can find the reciprocal vectors by just inverting them

$$\begin{aligned}\mathbf{x}^\rho &= \frac{\rho}{x} \\ \mathbf{x}^\theta &= \frac{1}{x \cdot i}.\end{aligned} \tag{6.115}$$

Answer for Exercise 6.15

The tangent vector for the curve is

$$\mathbf{x}_\theta = -\gamma_1\gamma_1\gamma_2 e^{-i\theta} = \gamma_2 e^{-i\theta}, \tag{6.130}$$

with reciprocal vector $\mathbf{x}^\theta = e^{i\theta}\gamma^2$. The differential element is $d^1\mathbf{x} = \gamma_2 e^{-i\theta} d\theta$, so the integrand is

$$\begin{aligned}\int_0^{\pi/4} \left(\mathbf{x}^\theta + \gamma_3 + \gamma_1\gamma_0\right) d^1\mathbf{x}\,\gamma_0 &= \int_0^{\pi/4} \left(e^{i\theta}\gamma^2 + \gamma_3 + \gamma_1\gamma_0\right)\gamma_2 e^{-i\theta} d\theta\,\gamma_0 \\ &= \frac{\pi}{4}\gamma_0 + \left(\gamma_{32} + \gamma_{102}\right)\frac{1}{-i}\left(e^{-i\pi/4} - 1\right)\gamma_0 \\ &= \frac{\pi}{4}\gamma_0 + \frac{1}{\sqrt{2}}\left(\gamma_{32} + \gamma_{102}\right)\gamma_{120}\left(1 - \gamma_{12}\right) \\ &= \frac{\pi}{4}\gamma_0 + \frac{1}{\sqrt{2}}\left(\gamma_{310} + 1\right)\left(1 - \gamma_{12}\right).\end{aligned} \tag{6.131}$$

Observe how care is required not to reorder any terms. This particular end result is a multivector with scalar, vector, bivector, and trivector grades, but no pseudoscalar component. The grades in the end result depend on both the function in the integrand and on the path. For example, had we integrated all the way around the circle, the end result would have been the vector $2\pi\gamma_0$ (i.e. a γ_0 weighted unit circle circumference), as all the other grades would have been killed by the complex exponential integrated over a full period.

Answer for Exercise 6.16

Observe that $\hat{\mathbf{v}}$ and γ_0 anticommute, so we may write our boost as a one sided exponential

$$x(\alpha) = \gamma_0 e^{-\hat{\mathbf{v}}\alpha} = e^{\hat{\mathbf{v}}\alpha}\gamma_0 = \left(\cosh\alpha + \hat{\mathbf{v}}\sinh\alpha\right)\gamma_0. \tag{6.133}$$

The tangent vector is just

$$\mathbf{x}_\alpha = \frac{\partial x}{\partial \alpha} = e^{\hat{\mathbf{v}}\alpha}\hat{\mathbf{v}}\gamma_0. \tag{6.134}$$

Let's get a bit of intuition about the nature of this vector. It's square is

$$\begin{aligned}\mathbf{x}_\alpha^2 &= e^{\hat{\mathbf{v}}\alpha}\hat{\mathbf{v}}\gamma_0 e^{\hat{\mathbf{v}}\alpha}\hat{\mathbf{v}}\gamma_0 \\ &= -e^{\hat{\mathbf{v}}\alpha}\hat{\mathbf{v}}e^{-\hat{\mathbf{v}}\alpha}\hat{\mathbf{v}}(\gamma_0)^2 \\ &= -1,\end{aligned} \tag{6.135}$$

so we see that the tangent vector is a spacelike unit vector. As the vector representing points on the curve is necessarily timelike (due to Lorentz invariance), these two must be orthogonal at all points. Let's confirm this algebraically

$$\begin{aligned}x \cdot \mathbf{x}_\alpha &= \left\langle e^{\hat{\mathbf{v}}\alpha}\gamma_0 e^{\hat{\mathbf{v}}\alpha}\hat{\mathbf{v}}\gamma_0\right\rangle \\ &= \left\langle e^{-\hat{\mathbf{v}}\alpha}e^{\hat{\mathbf{v}}\alpha}\hat{\mathbf{v}}(\gamma_0)^2\right\rangle \\ &= \langle\hat{\mathbf{v}}\rangle \\ &= 0.\end{aligned} \tag{6.136}$$

Here we used $e^{\hat{\mathbf{v}}\alpha}\gamma_0 = \gamma_0 e^{-\hat{\mathbf{v}}\alpha}$, and $\langle AB\rangle = \langle BA\rangle$. Geometrically, we have the curious fact that the direction vectors to points on the curve are perpendicular (with respect to our relativistic dot product) to the tangent vectors on the curve, as illustrated in fig. 6.6.

Answer for Exercise 6.17

We have $\mathbf{x}_\rho = \gamma_0 e^{-\hat{\mathbf{v}}\alpha}$ and $\mathbf{x}_\alpha = \rho\gamma_2 e^{-\hat{\mathbf{v}}\alpha}$, so

$$\begin{aligned}d^2\mathbf{x} &= (\mathbf{x}_\rho \wedge \mathbf{x}_\alpha)d\rho d\alpha \\ &= \left\langle \gamma_0 e^{-\hat{\mathbf{v}}\alpha}\rho\gamma_2 e^{-\hat{\mathbf{v}}\alpha}\right\rangle_2 d\rho d\alpha \\ &= \rho\gamma_{02}d\rho d\alpha,\end{aligned} \tag{6.146}$$

so the integral is

$$\begin{aligned}\int \rho\gamma_1 e^{\gamma_{21}\alpha}\gamma_{022}d\rho d\alpha &= -\frac{1}{2}\rho^2\int \gamma_1 e^{\gamma_{21}\alpha}\gamma_0 d\alpha \\ &= \frac{\gamma_{01}}{2}\rho^2\int e^{\gamma_{21}\alpha}d\alpha \\ &= \frac{\gamma_{01}}{2}\rho^2\gamma^{12}e^{\gamma_{21}\alpha} \\ &= \frac{\rho^2\gamma_{20}}{2}e^{\gamma_{21}\alpha}.\end{aligned} \tag{6.147}$$

Because F and G were both vectors, the resulting integral could only have been a multivector with grades 0,2,4. As it happens, there were no scalar nor pseudoscalar grades in the end result, and we ended up with the space-time plane between γ_0, and $\gamma_2 e^{\gamma_{21}\alpha}$, which are rotations of γ_2 in the x,y plane. This is illustrated in fig. 6.7 (omitting scale and sign factors.)

Answer for Exercise 6.18

Let $d^2\mathbf{x} = I\hat{\mathbf{n}}dA$, implicitly fixing the relative orientation of the bivector area element compared to the chosen surface normal direction.

$$\begin{aligned}
\int \left(d^2\mathbf{x}\cdot\boldsymbol{\nabla}\right)\cdot\mathbf{f} &= \int dA\langle I\hat{\mathbf{n}}\boldsymbol{\nabla}\rangle_1\cdot\mathbf{f} \\
&= \int dA\left(I\left(\hat{\mathbf{n}}\wedge\boldsymbol{\nabla}\right)\right)\cdot\mathbf{f} \\
&= \int dA\left\langle I^2\left(\hat{\mathbf{n}}\times\boldsymbol{\nabla}\right)\mathbf{f}\right\rangle \\
&= -\int dA\left(\hat{\mathbf{n}}\times\boldsymbol{\nabla}\right)\cdot\mathbf{f} \\
&= -\int dA\hat{\mathbf{n}}\cdot\left(\boldsymbol{\nabla}\times\mathbf{f}\right).
\end{aligned} \tag{6.149}$$

Answer for Exercise 6.19

Let's designate the tangent basis vectors as

$$\mathbf{x}_0 = \frac{\partial x}{\partial t} = c\gamma_0, \tag{6.164}$$

and

$$\mathbf{x}_2 = \frac{\partial x}{\partial y} = \gamma_2, \tag{6.165}$$

so the vector derivative is

$$\partial = \frac{1}{c}\gamma^0\frac{\partial}{\partial t} + \gamma^2\frac{\partial}{\partial y}, \tag{6.166}$$

and the area element is

$$d^2\mathbf{x} = c\gamma_0\gamma_2. \tag{6.167}$$

The fundamental theorem of surface integrals is just a statement that

$$\int_{t_0}^{t_1} cdt \int_{y_0}^{y_1} dy F\gamma_0\gamma_2 \left(\frac{1}{c}\gamma^0\frac{\partial}{\partial t} + \gamma^2\frac{\partial}{\partial y}\right) G = \int F\left(c\gamma_0 dt - \gamma_2 dy\right) G, \tag{6.168}$$

where the RHS, when stated explicitly, really means

$$\begin{aligned}\int F\left(c\gamma_0 dt - \gamma_2 dy\right) G = &\int_{t_0}^{t_1} cdt \left(F(t,y_1)\gamma_0 G(t,y_1) - F(t,y_0)\gamma_0 G(t,y_0)\right) \\ &- \int_{y_0}^{y_1} dy \left(F(t_1,y)\gamma_2 G(t_1,y)\right. \\ &\left. - F(t_0,y)\gamma_0 G(t_0,y)\right).\end{aligned} \tag{6.169}$$

In this particular case, since $\mathbf{x}_0 = c\gamma_0, \mathbf{x}_2 = \gamma_2$ are both constant functions that depend on neither t nor y, it is easy to derive the full expansion of eq. (6.169) directly from the LHS of eq. (6.168).

Answer for Exercise 6.20

For the tangent basis vectors we have

$$\mathbf{x}_\rho = \frac{\partial x}{\partial \rho} = e^{-\hat{\mathbf{v}}\alpha/2} x(0,1) e^{\hat{\mathbf{v}}\alpha/2} = \frac{x}{\rho}, \tag{6.171}$$

and

$$\mathbf{x}_\alpha = \frac{\partial x}{\partial \alpha} = \left(-\hat{\mathbf{v}}/2\right) x + x\left(\hat{\mathbf{v}}/2\right) = x\cdot\hat{\mathbf{v}}. \tag{6.172}$$

These vectors $\mathbf{x}_\rho, \mathbf{x}_\alpha$ are orthogonal, as $x\cdot\hat{\mathbf{v}}$ is the projection of x onto the spacetime plane $x\wedge\hat{\mathbf{v}} = 0$, but rotated so that $x\cdot\left(x\cdot\hat{\mathbf{v}}\right) = 0$. Because of this orthogonality, the vector derivative for this tangent space is

$$\partial = \frac{1}{x\cdot\hat{\mathbf{v}}}\frac{\partial}{\partial\alpha} + \frac{\rho}{x}\frac{\partial}{\partial\rho}. \tag{6.173}$$

The area element is

$$\begin{aligned} d^2\mathbf{x} &= d\rho d\alpha \frac{x}{\rho}\wedge\left(x\cdot\hat{\mathbf{v}}\right) \\ &= \frac{1}{\rho} d\rho d\alpha\, x\left(x\cdot\hat{\mathbf{v}}\right). \end{aligned} \tag{6.174}$$

The full statement of the fundamental theorem for this surface is

$$\int_S d\rho d\alpha\, F\left(\frac{1}{\rho}x\,(x\cdot\hat{\mathbf{v}})\right)\left(\frac{1}{x\cdot\hat{\mathbf{v}}}\frac{\partial}{\partial\alpha}+\frac{\rho}{x}\frac{\partial}{\partial\rho}\right)G=\int_{\partial S}F\left(d\rho\frac{x}{\rho}-d\alpha\,(x\cdot\hat{\mathbf{v}})\right)G. \tag{6.175}$$

As in the previous example, due to the orthogonality of the tangent basis vectors, it's easy to show find the RHS directly from the LHS.

Answer for Exercise 6.21

The tangent basis vectors are just $\mathbf{x}_u = a, \mathbf{x}_v = b$, with reciprocals

$$\mathbf{x}^u = \mathbf{x}_v\cdot\frac{1}{\mathbf{x}_u\wedge\mathbf{x}_v} = b\cdot\frac{1}{a\wedge b}, \tag{6.176}$$

and

$$\mathbf{x}^v = -\mathbf{x}_u\cdot\frac{1}{\mathbf{x}_u\wedge\mathbf{x}_v} = -a\cdot\frac{1}{a\wedge b}. \tag{6.177}$$

The fundamental theorem, with respect to this surface, when written out explicitly takes the form

$$\int F\,dudv\,(a\wedge b)\,\frac{1}{a\wedge b}\cdot\left(a\frac{\partial}{\partial u}-b\frac{\partial}{\partial v}\right)G=\int F\,(adu-bdv)\,G \tag{6.178}$$

This is a good example to illustrate the geometry of the line integral circulation. Suppose that we are integrating over $u\in[0,1], v\in[0,1]$. In this case, the line integral really means

$$\begin{aligned}\int F\,(adu-bdv)\,G = &+\int F(u,1)(+adu)G(u,1)+\int F(u,0)(-adu)G(u,0)\\ &+\int F(1,v)(-bdv)G(1,v)\\ &+\int F(0,v)(+bdv)G(0,v),\end{aligned} \tag{6.179}$$

which is a path around the spacetime parallelogram spanned by u, v, as illustrated in fig. 6.8, which illustrates the orientation of the bivector area element with the arrows around the exterior of the parallelogram: $0\to a\to a+b\to b\to 0$.

Answer for Exercise 7.1

In terms of coordinates, the first order Taylor expansion of the action is

$$\begin{aligned} S &\to S + \delta S \\ &= \int d\tau L(x^\alpha + \epsilon^\alpha, \dot{x}^\alpha + \dot{\epsilon}^\alpha) \\ &= \int d\tau \left(L(x^\alpha, \dot{x}^\alpha) + \epsilon^\mu \frac{\partial L}{\partial x^\mu} + \dot{\epsilon}^\mu \frac{\partial L}{\partial \dot{x}^\mu} \right). \end{aligned} \tag{7.8}$$

As before, we integrate by parts to separate out a pure boundary term

$$\delta S = \int d\tau \epsilon^\mu \left(\frac{\partial L}{\partial x^\mu} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{x}^\mu} \right) + \int d\tau \frac{d}{d\tau} \left(\epsilon^\mu \frac{\partial L}{\partial \dot{x}^\mu} \right). \tag{7.9}$$

The boundary term is killed since $\epsilon^\mu = 0$ at the end points of the action integral. We conclude that extremization of the action ($\delta S = 0$, for all ϵ^μ) requires

$$\frac{\partial L}{\partial x^\mu} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{x}^\mu} = 0. \tag{7.10}$$

Answer for Exercise 7.2

The first identity follows easily by expansion in coordinates

$$\begin{aligned} \nabla(a \cdot x) &= \gamma^\mu \partial_\mu a_\alpha x^\alpha \\ &= \gamma^\mu a_\alpha \delta^\alpha_\mu \\ &= \gamma^\mu a_\mu \\ &= a. \end{aligned} \tag{7.15}$$

The second identity follows by linearity of the gradient

$$\begin{aligned} \nabla x^2 &= \nabla(x \cdot x) \\ &= \left(\nabla(x \cdot a)\right)\Big|_{a=x} + \left(\nabla(b \cdot x)\right)\Big|_{b=x} \\ &= a|_{a=x} + b|_{b=x} \\ &= 2x. \end{aligned} \tag{7.16}$$

Answer for Exercise 7.3

$$\begin{aligned} F &= \nabla \wedge A \\ &= (\gamma_\mu \partial^\mu) \wedge (\gamma_\nu A^\nu) \\ &= (\gamma_\mu \wedge \gamma_\nu)\, \partial^\mu A^\nu. \end{aligned} \tag{7.19}$$

To this we can use the usual tensor trick (add self to self, change indexes, and divide by two), to give

$$\begin{aligned} F &= \frac{1}{2}\left((\gamma_\mu \wedge \gamma_\nu)\, \partial^\mu A^\nu + (\gamma_\nu \wedge \gamma_\mu)\, \partial^\nu A^\mu\right) \\ &= \frac{1}{2}(\gamma_\mu \wedge \gamma_\nu)\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right), \end{aligned} \tag{7.20}$$

which is just

$$F = \frac{1}{2}(\gamma_\mu \wedge \gamma_\nu)\, F^{\mu\nu}. \tag{7.21}$$

Now, let's expand $(F \cdot v) \cdot \gamma^\mu$ to compare to the earlier expansion in terms of ∇ and A.

$$\begin{aligned} (F \cdot v) \cdot \gamma^\mu &= \frac{1}{2} F^{\alpha\nu}\left((\gamma_\alpha \wedge \gamma_\nu) \cdot \left(\gamma^\beta v_\beta\right)\right) \cdot \gamma^\mu \\ &= \frac{1}{2} F^{\alpha\nu} v_\beta \left(\delta_\nu{}^\beta \gamma_\alpha{}^\mu - \delta_\alpha{}^\beta \gamma_\nu{}^\mu\right) \\ &= \frac{1}{2}\left(F^{\mu\beta} v_\beta - F^{\beta\mu} v_\beta\right) \\ &= F^{\mu\nu} v_\nu. \end{aligned} \tag{7.22}$$

This alternate expansion illustrates some of the connectivity between the geometric algebra approach and the traditional tensor formalism.

Answer for Exercise 7.4

Let $\delta_\mu L = \gamma_\mu \cdot \delta L$, so that we can write the Euler-Lagrange equations as

$$0 = \delta_\mu L = \frac{\partial L}{\partial x^\mu} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{x}^\mu}. \tag{7.23}$$

Operating on the kinetic term of the Lagrangian, we have

$$\delta_\mu L_0 = -\frac{d}{d\tau} m v_\mu. \tag{7.24}$$

For the potential term

$$\begin{aligned}
\delta_\mu L_1 &= \frac{q}{c}\left(v_\nu \frac{\partial A^\nu}{\partial x^\mu} - \frac{d}{d\tau}A_\mu\right) \\
&= \frac{q}{c}\left(v_\nu \frac{\partial A^\nu}{\partial x^\mu} - \frac{dx_\alpha}{d\tau}\frac{\partial A_\mu}{\partial x_\alpha}\right) \\
&= \frac{q}{c}v^\nu\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) \\
&= \frac{q}{c}v^\nu F_{\mu\nu}.
\end{aligned} \tag{7.25}$$

Putting the pieces together gives

$$\frac{d}{d\tau}(mv_\mu) = \frac{q}{c}v^\nu F_{\mu\nu}, \tag{7.26}$$

which is identical[1] to the tensor form that we found by expanding the geometric algebra form of Maxwell's equation in coordinates.

Answer for Exercise 7.5

First calculate the field velocity product in terms of electric and magnetic components. In this new frame of reference write the proper velocity of the charged particle as $v = \gamma_\mu \dot{x}^\mu$

$$\begin{aligned}
F\cdot v &= (\mathbf{E} + Ic\mathbf{B})\cdot v \\
&= (E^i\gamma_{i0} - \epsilon_{ijk}cB^k\gamma_{ij})\cdot\gamma_\mu\dot{x}^\mu \\
&= E^i\dot{x}^0\gamma_{i0}\cdot\gamma_0 + E^i\dot{x}^j\gamma_{i0}\cdot\gamma_j - \epsilon_{ijk}cB^k\dot{x}^m\gamma_{ij}\cdot\gamma_m.
\end{aligned} \tag{7.37}$$

We apply a γ_0 wedge to determine this observer dependent expression of the force.

$$\begin{aligned}
\gamma^{-1}(F\cdot v)\wedge\gamma_0 &= \left(E^i\dot{x}^0(\gamma_{i0}\cdot\gamma_0) + E^i\dot{x}^j(\gamma_{i0}\cdot\gamma_j) - \epsilon_{ijk}cB^k\dot{x}^m\gamma_{ij}\cdot\gamma_m\right)\wedge\gamma_0 \\
&= E^i\dot{x}^0\gamma_{i0} - \epsilon_{ijk}cB^k\dot{x}^m(\gamma_i)^2(\gamma_i\delta_{jm} - \gamma_j\delta_{im})\wedge\gamma_0 \\
&= \left(E^i\dot{x}^0\gamma_{i0} + \epsilon_{ijk}cB^k\left(\dot{x}^j\gamma_{i0} - \dot{x}^i\gamma_{j0}\right)\right),
\end{aligned} \tag{7.38}$$

where $\gamma = dt/d\tau$. This wedge application has discarded the timelike components of the force equation with respect to this observer rest frame. Introduce the basis $\{\mathbf{e}_i = \gamma_i\wedge\gamma_0\}$ for this observers' Euclidean space. These

1 Some minor index raising and lowering gymnastics are required.

spacetime bivectors square to unity, and thus behave in every respect like Euclidean space vector basis vectors. Writing $\mathbf{E} = E^i \mathbf{e}_i$, $\mathbf{B} = B^i \mathbf{e}_i$, and $\mathbf{v} = \mathbf{e}_i dx^i/dt$ we have

$$\gamma^{-1}(F \cdot v) \wedge \gamma_0 = c\left(\mathbf{E} + \epsilon_{ijk} B^k \left(\frac{dx^j}{dt}\mathbf{e}_i - \frac{dx^i}{dt}\mathbf{e}_j\right)\right). \tag{7.39}$$

This inner antisymmetric sum is just the cross product. This can be observed by expanding the determinant

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \\ &= \mathbf{e}_1(a_2 b_3 - a_3 b_2) + \mathbf{e}_2(a_3 b_1 - a_1 b_3) + \mathbf{e}_3(a_1 b_2 - a_2 b_1) \\ &= \mathbf{e}_i a_j b_k \epsilon_{ijk}.\end{aligned} \tag{7.40}$$

This leaves

$$q(F \cdot v/c) \wedge \gamma_0 = \gamma q \left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right). \tag{7.41}$$

Plugging back into eq. (7.35b) gives

$$\frac{d}{dt}\left(m\gamma \mathbf{v}\right) = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right). \tag{7.42}$$

Answer for Exercise 7.7

The gauge transformed Lagrangian is

$$L = \frac{1}{2}mv^2 + qA \cdot v/c + \frac{qv}{c} \cdot \nabla\phi. \tag{7.45}$$

Observe that

$$\begin{aligned} v \cdot \nabla\phi &= \frac{dx^\mu}{d\tau}\frac{\partial\phi}{\partial x^\mu} \\ &= \frac{d\phi}{d\tau}, \end{aligned} \tag{7.46}$$

which means that the action is transformed to

$$\begin{aligned} S &\to S + \frac{q}{c}\int d\tau \frac{d\phi}{d\tau} \\ &= S + \frac{q}{c}\phi|_{\Delta\tau}. \end{aligned} \tag{7.47}$$

As the action is evaluated over the fixed interval $\Delta\tau$, the gauge transformation only changes the action by a constant, so the equations of motion are unchanged.

8

NOETHER'S THEOREM.

8.0.1 *Noether's theorem.*

Also covered in [2] is Noether's theorem in multivector form. This is used to calculate the conserved quantity the Hamiltonian for Lagrangian's with no time dependence. Lets try something similar for the scalar variable case, after which the multivector case may make more sense.

At its heart Noether's theorem appears to describe change of variables in Lagrangians.

Given a Lagrangian dependent on generalized coordinates q^i, and their first order derivatives, as well as the path parameter λ.

$$\begin{aligned} L &= L(q^i, \dot{q}^i, \lambda) \\ q^i &= q^i(r^i(\lambda), \alpha). \end{aligned} \tag{8.1}$$

One example of such a change of variables would be the Galilean transformation $q^i = x^i(t) + vt$, with $\lambda = t$.

Application of the chain rule shows how to calculate the first order change of the Lagrangian with respect to the new parameter α.

$$\frac{dL}{d\alpha} = \frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial \alpha} + \frac{\partial L}{\partial \dot{q}^i}\frac{\partial \dot{q}^i}{\partial \alpha}. \tag{8.2}$$

If q^i, and $\dot{q}^i$ satisfy the Euler-Lagrange equations eq. (4.7), then this can be written

$$\frac{dL}{d\alpha} = \left(\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{q}^i}\right)\frac{\partial q^i}{\partial \alpha} + \frac{\partial L}{\partial \dot{q}^i}\frac{\partial \dot{q}^i}{\partial \alpha}. \tag{8.3}$$

If one additionally has

$$\frac{\partial^2 q^i}{(\partial \alpha)^2} = \frac{\partial^2 \dot{q}^i}{(\partial \alpha)^2} = 0, \tag{8.4}$$

so that $\partial q^i/\partial\alpha$, and $\partial \dot{q}^i/\partial\alpha$ are dependent only on λ, then eq. (8.3) can be written as a total derivative

$$\frac{dL}{d\alpha} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{q}^i}\frac{\partial q^i}{\partial \alpha}\right). \tag{8.5}$$

If there is an α dependence in these derivatives a weaker total derivative statement is still possible, by evaluating the Lagrangian derivative and $\partial q^i/\partial\alpha$ at some specific constant value of α. This is

$$\left.\frac{dL}{d\alpha}\right|_{\alpha=\alpha_0} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{q}^i}\left.\frac{\partial q^i}{\partial\alpha}\right|_{\alpha=\alpha_0}\right). \tag{8.6}$$

8.0.1.1 *Hamiltonian.*

Hmm, the above equations do not much like the Noether's equation in [2]. However, in this form, we can get at the Hamiltonian statement without any trouble. Let us do that first, then return to Noether's

Of particular interest is when the change of variables for the generalized coordinates is dependent on the parameter $\alpha = \lambda$. Given this type of transformation we can write eq. (8.5) as

$$\frac{dL}{d\lambda} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{q}^i}\frac{\partial q^i}{\partial\lambda}\right). \tag{8.7}$$

For this to be valid in this $\alpha = \lambda$ case, note that the Lagrangian itself may not be explicitly dependent on the parameter λ. Such a dependence would mean that eq. (8.2) would require an additional $\partial L/\partial\lambda$ term.

The difference of the eq. (8.7) terms is called the Hamiltonian H

$$\frac{dH}{d\lambda} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - \frac{dL}{d\lambda}\right) = 0. \tag{8.8}$$

Up to a constant

$$H = \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - \frac{dL}{d\lambda}, \tag{8.9}$$

the Hamiltonian is a conserved quantity when the Lagrangian has no explicit λ dependence.

8.0.1.2 *Noether's take II.*

Noether's theorem is about conserved quantities under symmetry transformations. Let us revisit the attempt at derivation once more cutting down the complexity even further, considering a transformation of a single generalized coordinate and the corresponding change to the Lagrangian under such a transformation.

Write

$$\begin{aligned} q \to q' &= f(q,\alpha) \\ L(q,\dot{q},\lambda) \to L' &= L(q',\dot{q}',\lambda) = L(f,\dot{f},\lambda). \end{aligned} \tag{8.10}$$

Now as before consider the derivative

$$\frac{dL'}{d\alpha} = \frac{\partial L}{\partial f}\frac{\partial f}{\partial \alpha} + \frac{\partial L}{\partial \dot{f}}\frac{\partial \dot{f}}{\partial \alpha}. \tag{8.11}$$

In terms of the transformed coordinates the Euler-Lagrange equations require

$$\frac{\partial L}{\partial f} = \frac{d}{d\lambda}\frac{\partial L}{\partial \dot{f}}, \tag{8.12}$$

and back-substitution into eq. (8.11) gives

$$\frac{dL'}{d\alpha} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{f}}\right)\frac{\partial f}{\partial \alpha} + \frac{\partial L}{\partial \dot{f}}\frac{\partial \dot{f}}{\partial \alpha}. \tag{8.13}$$

This can be written as a total derivative if

$$\begin{aligned} \frac{\partial \dot{f}}{\partial \alpha} &= \frac{d}{d\lambda}\frac{\partial f}{\partial \alpha} \\ \frac{\partial}{\partial \alpha}\frac{df}{d\lambda} &= \frac{\partial^2 f}{\partial q \partial \alpha}\dot{q} + \frac{\partial^2 f}{(\partial \alpha)^2}\dot{\alpha} \\ \frac{\partial}{\partial \alpha}\left(\frac{\partial f}{\partial q}\dot{q} + \frac{\partial f}{\partial \alpha}\dot{\alpha}\right) &= \\ \frac{\partial^2 f}{\partial \alpha \partial q}\dot{q} + \frac{\partial^2 f}{(\partial \alpha)^2}\dot{\alpha} + \frac{\partial f}{\partial \alpha}\frac{\partial \dot{\alpha}}{\partial \alpha} &= \end{aligned} \tag{8.14}$$

Thus given a constraint of sufficient continuity

$$\frac{\partial^2 f}{\partial \alpha \partial q} = \frac{\partial^2 f}{\partial q \partial \alpha}, \tag{8.15}$$

and also that $\dot{\alpha}$ is not a function of α

$$\frac{\partial \dot{\alpha}}{\partial \alpha} = 0, \tag{8.16}$$

we have from eq. (8.13)

$$\frac{dL'}{d\alpha} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{f}}\frac{\partial f}{\partial \alpha}\right). \tag{8.17}$$

This is

$$\frac{dL'}{d\alpha} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{q}'}\frac{\partial q'}{\partial \alpha}\right). \tag{8.18}$$

The details of generalizing this to multiple variables are almost the same, but does not really add anything to the understanding. This generalization is included as an appendix below for completeness, but the end result is

$$\frac{dL'}{d\alpha} = \frac{d}{d\lambda}\left(\sum_i \frac{\partial L}{\partial \dot{q'}^i}\frac{\partial q^{i'}}{\partial \alpha}\right). \tag{8.19}$$

In words, when the transformed Lagrangian is symmetric (not a function of α) under coordinate transformation then this inner quantity, a generalized momentum velocity product, is constant (conserved)

$$\sum_i \frac{\partial L}{\partial \dot{q'}^i}\frac{\partial q^{i'}}{\partial \alpha} = \text{constant}. \tag{8.20}$$

Transformations that leave the Lagrangian unchanged have this associated conserved quantity, which dimensionally, assuming a time parametrization, has units of energy (mv^2).

FIXME: The $\partial\dot{\alpha}/\partial\alpha = 0$ requirement is what is removed by evaluation at $\alpha = \alpha_0$. This statement seems somewhat handwaving like. Firm it up with an example and concrete justification.

Note that it still does not quite match the multivector result from [2], equation 12.10

$$\left.\frac{dL'}{d\alpha}\right|_{\alpha=0} = \frac{d}{dt}\sum_{i=1}^{n}\left(\frac{\partial X_i'}{\partial \alpha} * \partial_{\dot{X}_i} L\right). \tag{8.21}$$

I believe there is a missing prime there, and it should read

$$\left.\frac{dL'}{d\alpha}\right|_{\alpha=0} = \frac{d}{dt}\sum_{i=1}^{n}\left(\frac{\partial X_i'}{\partial \alpha} * \partial_{\dot{X}_i'} L\right). \tag{8.22}$$

8.1 VECTOR FORMULATION OF EULER-LAGRANGE EQUATIONS.

8.1.1 *Simple case. Unforced purely kinetic Lagrangian.*

Before considering multivector Lagrangians, a step back to the simplest vector Lagrangian is in order

$$L = \frac{1}{2}m\dot{\mathbf{x}}\cdot\dot{\mathbf{x}}. \tag{8.23}$$

Writing $\mathbf{x}(\lambda) = \bar{\mathbf{x}} + \epsilon\mathbf{n}$, and using the variational technique directly the equation of motion for this unforced path should follow directly in vector form

$$S = \int d\lambda \frac{1}{2} m \dot{\bar{\mathbf{x}}}^2 + \int m d\lambda \epsilon \dot{\bar{\mathbf{x}}} \cdot \dot{\mathbf{n}} + \int d\lambda \frac{1}{2} m \epsilon^2 \dot{\bar{\mathbf{n}}}^2. \tag{8.24}$$

Integration by parts operating directly on the vector function we have

$$\begin{aligned} \frac{dS}{d\epsilon}\bigg|_{\epsilon=0} &= m\dot{\bar{\mathbf{x}}} \cdot \mathbf{n}\big|_{\partial\lambda} - \int m d\lambda \ddot{\bar{\mathbf{x}}} \cdot \mathbf{n} \\ &= -\int m d\lambda \ddot{\bar{\mathbf{x}}} \cdot \mathbf{n}. \end{aligned} \tag{8.25}$$

Introducing shorthand $\delta S/\delta\mathbf{x}$, for a vector functional derivative, we have

$$\frac{dS}{d\epsilon}\bigg|_{\epsilon=0} = \int d\lambda \mathbf{n} \cdot \frac{\delta S}{\delta \mathbf{x}}, \tag{8.26}$$

where the extremal condition is

$$\frac{\delta S}{\delta \mathbf{x}} = -m\ddot{\bar{\mathbf{x}}} = 0. \tag{8.27}$$

Here the expected and desired Euler Lagrange equation for the Lagrangian (constant velocity in some direction dependent on initial conditions) is arrived at directly in vector form without dropping down to coordinates and reassembling them to get back the vector expression.

8.1.2 *Position and velocity gradients in the configuration space.*

Having tackled the simplest case, to generalize this we need a construct to do first order Taylor series expansion in the neighborhood of a vector position. The (multivector) gradient is the obvious candidate operator to do the job. Before going down that road consider the scalar Lagrangian case once more, where we will see that it is natural to define position and velocity gradients in the configuration space. It will also be observed that the chain rule essentially motivates the initially somewhat odd seeming reciprocal basis used to express the gradient when operating in a non-orthonormal frame.

In eq. (4.3), the linear differential increment in the neighborhood of the optimal solution had the form

$$\Delta L = +\sum_i (\bar{q}^i + n^i) \frac{\partial L}{\partial q^i}\bigg|_{q^i=\bar{q}^i} + \sum_i (\dot{\bar{q}}^i + \dot{n}^i) \frac{\partial L}{\partial \dot{q}^i}\bigg|_{q^i=\bar{q}^i}. \tag{8.28}$$

If one defines a configuration space position and velocity gradients respectively as

$$\begin{aligned}\nabla_{\mathbf{q}} &= \left(\frac{\partial}{\partial q^1}, \frac{\partial}{\partial q^2}, \cdots, \frac{\partial}{\partial q^n}\right) = f_k \frac{\partial}{\partial q^k} \\ \nabla_{\dot{\mathbf{q}}} &= \left(\frac{\partial}{\partial \dot{q}^1}, \frac{\partial}{\partial \dot{q}^2}, \cdots, \frac{\partial}{\partial \dot{q}^n}\right) = f_k \frac{\partial}{\partial \dot{q}^k}.\end{aligned} \tag{8.29}$$

and forms a configuration space vector with respect to some linearly independent, but not necessarily orthonormal, basis

$$\mathbf{q} = q^i e_i. \tag{8.30}$$

then the chain rule dictates the relationship between the configuration vector basis and the basis with which the gradient must be expressed. In particular, if we wish to write eq. (8.28) in terms of the configuration space gradients

$$\Delta L = (\bar{\mathbf{q}} + \mathbf{n}) \cdot \nabla_{\mathbf{q}} L\Big|_{\mathbf{q}=\bar{\mathbf{q}}} + (\dot{\bar{\mathbf{q}}} + \dot{\mathbf{n}}) \cdot \nabla_{\dot{\mathbf{q}}} L\Big|_{\dot{\mathbf{q}}=\dot{\bar{\mathbf{q}}}}. \tag{8.31}$$

Then we must have a reciprocal relationship between the basis vector for the configuration space vectors e_i, and the corresponding vectors from which the gradient was formed

$$\begin{aligned} & e_i \cdot f_j = \delta_{ij} \\ & \Longrightarrow \\ & f_j = e^j. \end{aligned} \tag{8.32}$$

This gives us the position and velocity gradients in the configuration space

$$\begin{aligned}\nabla_{\mathbf{q}} &= e^k \frac{\partial}{\partial q^k} \\ \nabla_{\dot{\mathbf{q}}} &= e^k \frac{\partial}{\partial \dot{q}^k}.\end{aligned} \tag{8.33}$$

Note also that the size of this configuration space does not have to be the same space as the problem. With this definitions completion of the integration by parts yields the Euler-Lagrange equations in a hybrid configuration space vector form

$$\nabla_{\mathbf{q}} L = \frac{d}{d\lambda} \nabla_{\dot{\mathbf{q}}} L. \tag{8.34}$$

When the configuration space equals the geometrical space being operated in (ie: generalized coordinates are regular old coordinates), this provides a nice explanation for why we must have the funny pairing of upper index coordinates in the partials of the gradient and reciprocal frame vectors multiplying all these partials. Contrast to a text like [2] where the gradient (and spacetime gradient) are defined in this fashion instead, and one gradually sees that this does in fact work out.

That said, the negative side of this vector notation is that it obscures somewhat the Euler-Lagrange equations, which are not terribly complicated to begin with. However, since this appears to be the form of the multivector form of the Euler-Lagrange equations it is likely worthwhile to see how this also expresses the simpler familiar scalar case too.

8.2 EXAMPLE APPLICATIONS OF NOETHER'S THEOREM.

Linear translation and rotational translation appear to be the usual first example applications. [26] does this, as does the wikipedia article. Reading about those without actually working through it myself never made complete sense (esp. want to do the angular momentum example).

Noether's theorem is not really required to see that in the case of unforced motion eq. (8.23), translation of coordinates $\mathbf{x} \to \mathbf{x} + \mathbf{a}$ will not change the equation of motion. This is the conservation of linear momentum result so familiar from high school physics.

8.2.1 *Angular momentum in a radial potential.*

The conservation of angular momentum case is more interesting.

Suppose that one has a radial potential applied to a point particle

$$L = \frac{1}{2} m \dot{\mathbf{x}}^2 - \phi(|\mathbf{x}|^k). \tag{8.35}$$

and apply a rotational transformation to the coordinates

$$\mathbf{x} \to \exp(i\theta/2)\mathbf{x}\exp(-i\theta/2). \tag{8.36}$$

Provided that this is a fixed rotation with i, and θ constant (not functions of time), the transformed squared velocity is:

$$\begin{aligned}
\dot{\mathbf{x}}' \cdot \dot{\mathbf{x}}' &= \langle \exp(i\theta/2)\dot{\mathbf{x}}\exp(-i\theta/2)\exp(i\theta/2)\dot{\mathbf{x}}\exp(-i\theta/2)\rangle \\
&= \langle \exp(i\theta/2)\dot{\mathbf{x}}\dot{\mathbf{x}}\exp(-i\theta/2)\rangle \\
&= \dot{\mathbf{x}}^2\langle \exp(i\theta/2)\exp(-i\theta/2)\rangle \\
&= \dot{\mathbf{x}}^2.
\end{aligned} \tag{8.37}$$

Since $|\mathbf{x}'| = |\mathbf{x}|$ the transformed Lagrangian is unchanged by any rotation of coordinates.

Noether's equation eq. (8.19) takes the form

$$\frac{\partial L'}{\partial \theta} = \frac{d}{dt}\left(\frac{\partial \mathbf{x}'}{\partial \theta}\cdot \boldsymbol{\nabla}_{\mathbf{v}'} L\right). \tag{8.38}$$

Here the configuration space gradient is used to express the chain rule terms, picking the $\mathbb{R}^3$ standard basis vectors to express that gradient.

The velocity term can be expanded as

$$\begin{aligned}
\frac{\partial \mathbf{x}'}{\partial \theta} &= \frac{\partial}{\partial \theta}\left(\exp(i\theta/2)\mathbf{x}\exp(-i\theta/2)\right) \\
&= \frac{1}{2}(i\mathbf{x}' - \mathbf{x}'i) \\
&= i\cdot \mathbf{x}'.
\end{aligned} \tag{8.39}$$

The transformed conjugate momentum is

$$\boldsymbol{\nabla}_{\mathbf{v}'}\frac{1}{2}m\mathbf{v}'^2 = m\mathbf{v}' = \mathbf{p}'. \tag{8.40}$$

so the conserved quantity is

$$(i\cdot \mathbf{x}')\cdot \mathbf{p}' = \text{constant}. \tag{8.41}$$

Temporarily expressing the bivector for the rotational plane in terms of a dual relationship, $i = I\mathbf{n}$, where $\mathbf{n}$ is a unit normal to the plane we have

$$\begin{aligned}
(i\cdot \mathbf{x}')\cdot \mathbf{p}' &= ((I\mathbf{n})\cdot \mathbf{x}')\cdot \mathbf{p}' \\
&= \frac{1}{2}(I\mathbf{n}\mathbf{x}' - \mathbf{x}'I\mathbf{n})\cdot \mathbf{p}' \\
&= \frac{1}{2}\langle I(\mathbf{n}\mathbf{x}' - \mathbf{x}'\mathbf{n})\mathbf{p}'\rangle \\
&= \frac{1}{2}\langle I\mathbf{n}\mathbf{x}'\mathbf{p}\rangle - \langle I\mathbf{n}\mathbf{p}'\mathbf{x}'\rangle \\
&= \frac{1}{2}\left(\langle i(\mathbf{x}'\wedge \mathbf{p}')\rangle - \langle i(\mathbf{p}'\wedge \mathbf{x}')\rangle\right) \\
&= i\cdot(\mathbf{x}'\wedge \mathbf{p}').
\end{aligned} \tag{8.42}$$

Since i is a constant bivector we have angular momentum (dropping primes), by virtue of Lagrangian transformational symmetry and Noether's theorem the angular momentum

$$\mathbf{x} \wedge \mathbf{p} = \text{constant}, \tag{8.43}$$

is a constant of motion for a point particle Lagrangian in a radial potential field.

This is typically expressed in terms of the dual relationship using cross products

$$\mathbf{x} \times \mathbf{p} = \text{constant}. \tag{8.44}$$

Also observe the time derivative of the angular momentum in eq. (8.43)

$$\begin{aligned} \frac{d}{dt}(\mathbf{x} \wedge \mathbf{p}) &= \mathbf{p}/m \wedge \mathbf{p} + \mathbf{x} \wedge \dot{\mathbf{p}} \\ &= \mathbf{x} \wedge \dot{\mathbf{p}} \\ &= 0. \end{aligned} \tag{8.45}$$

Which says that the torque on a particle in a radial potential is zero. This finally supplies the rational for texts like [18], which while implicitly talking about motion in a (radial) gravitational potential, says something to the effect of "in the absence of external torques the angular momentum is conserved"!

What other more general non-radial potentials, if any, allow for this conservation statement? I had guess that something like the Lorentz force with velocity dependence in the potential will explicitly not have this conservation of angular momentum. [26] and [4] both cover Lagrangian transformation, and specifically cover this angular momentum issue, but blundering through it myself as done here was required to really see where it was coming from and to apply the idea.

8.2.2 *Hamiltonian.*

Consider a general kinetic form and a possibly velocity dependent potential

$$L = K - \phi = \frac{1}{2} \sum_{ij} g_{ij} \dot{q}^i \dot{q}^j - \phi, \tag{8.46}$$

and form the Hamiltonian. First calculate

$$\frac{\partial L}{\partial \dot{q}^i} = m \sum_j g_{ij}\dot{q}^j - \frac{\partial \phi}{\partial \dot{q}^i}$$

$$\Longrightarrow$$

$$\begin{aligned}\sum_i \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} &= m \sum_{ij} g_{ij}\dot{q}^i\dot{q}^j - \sum_i \dot{q}^i \frac{\partial \phi}{\partial \dot{q}^i} \\ &= 2K - \sum_i \dot{q}^i \frac{\partial \phi}{\partial \dot{q}^i}\end{aligned} \tag{8.47}$$

.

So, the Hamiltonian is

$$H = K - \sum_i \dot{q}^i \frac{\partial \phi}{\partial \dot{q}^i} + \phi. \tag{8.48}$$

For the less general case where $\mathbf{v}^2 = g_{ij}\dot{q}^i\dot{q}^j$, this is

$$H = K - \mathbf{v} \cdot \boldsymbol{\nabla}_{\mathbf{v}}\phi + \phi, \tag{8.49}$$

a conserved quantity with respect to the time derivative.

Similarly, for squared proper velocity $v^2 = g_{ij}\dot{q}^i\dot{q}^j$, and derivatives with respect to proper time

$$H = K - v \cdot \nabla_v \phi + \phi, \tag{8.50}$$

is conserved with respect to proper time.

As an example, consider the Lorentz force Lagrangian. For proper velocity v, four potential A, and positive time metric signature $(\gamma_0)^2 = 1$, the Lorentz force Lagrangian is

$$L = \frac{1}{2}mv \cdot v + qA \cdot v/c. \tag{8.51}$$

We therefore have

$$0 = \frac{d}{d\tau}\left(\frac{1}{2}mv^2 + v \cdot \nabla_v(qA \cdot v/c) - qA \cdot v/c\right). \tag{8.52}$$

Or

$$\frac{1}{2}mv^2 + v \cdot \nabla_v(qA \cdot v/c) - qA \cdot v/c = \kappa. \tag{8.53}$$

Where κ is some constant. Since $\nabla_v A^\mu = 0$, we have $\nabla_v A \cdot v = A$, and

$$\begin{aligned}\kappa &= \frac{1}{2}mv^2 + v \cdot (qA/c) - qA \cdot v/c \\ &= \frac{1}{2}mv^2.\end{aligned} \tag{8.54}$$

At a glance this does not look terribly interesting, since by definition of proper time we already know that $\frac{1}{2}mv^2 = \frac{1}{2}mc^2$ is a constant.

However, suppose that one did not assume proper time to start with, and instead considered an arbitrarily parametrized coordinate worldline and their corresponding solutions

$$\begin{aligned}x &= x(\lambda) \\ L &= \frac{1}{2}m\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda} + qA \cdot \frac{dx}{d\lambda}/c \\ \frac{\partial L}{\partial \lambda} &= \frac{d}{d\lambda}\frac{\partial L}{\partial \lambda}.\end{aligned} \tag{8.55}$$

The Hamiltonian conservation with respect to this parametrization then implies

$$\frac{d}{d\lambda}\left(\frac{1}{2}m\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}\right) = 0. \tag{8.56}$$

So that, independent of the parametrization, the quantity $\frac{1}{2}m\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}$ is a constant. This then follows as a consequence of Noether's theorem instead of by definition. Proper time then becomes that particular worldline parametrization $\lambda = \tau$ such that $\frac{1}{2}m\frac{dx}{d\tau} \cdot \frac{dx}{d\tau} = \frac{1}{2}mc^2$.

8.2.3 *Covariant Lorentz force Lagrangian.*

The Hamiltonian was used above to extract v^2 invariance from the Lorentz force Lagrangian under changes of proper time. The next obvious Noether's application is for a Lorentz transformation of the interaction Lagrangian. This was interesting enough seeming in its own right to treat separately and has been moved to E.

8.3 APPENDIX.

8.3.1 *Noether's equation derivation, multivariable case.*

Employing a couple judicious regular expressions starting from the text for the single variable treatment, plus some minor summation sign addition does the job.

$$\begin{aligned} q^i \to q^{i\prime} &= f^i(q^i, \alpha) \\ L(q^i, \dot{q}^i, \lambda) \to L' &= L(q^{i\prime}, \dot{q}'^i, \lambda) = L(f^i, \dot{f}^i, \lambda). \end{aligned} \tag{8.57}$$

Now as before consider the derivative

$$\frac{dL'}{d\alpha} = \sum_i \frac{\partial L}{\partial f^i}\frac{\partial f^i}{\partial \alpha} + \frac{\partial L}{\partial \dot{f}^i}\frac{\partial \dot{f}^i}{\partial \alpha}. \tag{8.58}$$

In terms of the transformed coordinates the Euler-Lagrange equations require

$$\frac{\partial L}{\partial f^i} = \frac{d}{d\lambda}\frac{\partial L}{\partial \dot{f}^i}. \tag{8.59}$$

and backsubstitution into eq. (8.58) gives

$$\frac{dL'}{d\alpha} = \sum_i \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{f}^i}\right)\frac{\partial f^i}{\partial \alpha} + \frac{\partial L}{\partial \dot{f}^i}\frac{\partial \dot{f}^i}{\partial \alpha}. \tag{8.60}$$

This can be written as a total derivative if

$$\begin{aligned} \frac{\partial \dot{f}^i}{\partial \alpha} &= \frac{d}{d\lambda}\frac{\partial f^i}{\partial \alpha} \\ \frac{\partial}{\partial \alpha}\frac{df}{d\lambda} &= \sum_j \frac{\partial^2 f^i}{\partial q^j \partial \alpha}\dot{q}^j + \frac{\partial^2 f^i}{(\partial \alpha)^2}\dot{\alpha} \\ \frac{\partial}{\partial \alpha}\left(\sum_j \frac{\partial f^i}{\partial q^j}\dot{q}^j + \frac{\partial f^i}{\partial \alpha}\dot{\alpha}\right) &= \\ \sum_j \frac{\partial^2 f^i}{\partial \alpha \partial q^j}\dot{q}^j + \frac{\partial^2 f^i}{(\partial \alpha)^2}\dot{\alpha} + \frac{\partial f^i}{\partial \alpha}\frac{\partial \dot{\alpha}}{\partial \alpha} &= \end{aligned} \tag{8.61}$$

Thus given constraints of sufficient continuity

$$\frac{\partial^2 f^i}{\partial \alpha \partial q^j} = \frac{\partial^2 f^i}{\partial q^j \partial \alpha}. \tag{8.62}$$

and also that $\dot{\alpha}$ is not a function of α

$$\frac{\partial \dot{\alpha}}{\partial \alpha} = 0. \tag{8.63}$$

we have from eq. (8.60)

$$\frac{dL'}{d\alpha} = \frac{d}{d\lambda}\left(\sum_i \frac{\partial L}{\partial \dot{f}^i}\frac{\partial f^i}{\partial \alpha}\right). \tag{8.64}$$

QED.

9 HAMILTONIAN MECHANICS.

9.1 MOTIVATION.

I have now seen Hamiltonian's used, mostly in a Quantum context, and think that I understand at least some of the math associated with the Hamiltonian and the Hamiltonian principle. I have, however, not used either of these enough that it seems natural to do so.

Here I attempt to summarize for myself what I know about Hamiltonian's, and work through a number of examples. Some of the examples considered will be ones already treated with the Lagrangian formalism 4.2.

Some notation will be invented along the way as reasonable, since I had like to try to also relate the usual coordinate representation of the Hamiltonian, the Hamiltonian principle, and the Poisson bracket, with the bivector representation of the 2N complex configuration space introduced in [2]. (NOT YET DONE).

9.2 HAMILTONIAN AS A CONSERVED QUANTITY.

Starting with the Lagrangian formalism the Hamiltonian can be found as a conserved quantity associated with time translation when the Lagrangian has no explicit time dependence. This follows directly by considering the time derivative of the Lagrangian $L = L(q^i, \dot{q}^i)$.

$$\begin{aligned}
\frac{dL}{dt} &= \frac{\partial L}{\partial q^i}\frac{dq^i}{dt} + \frac{\partial L}{\partial \dot{q}^i}\frac{d\dot{q}^i}{dt} \\
&= \dot{q}^i \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} + \frac{\partial L}{\partial \dot{q}^i}\frac{d\dot{q}^i}{dt} \\
&= \frac{d}{dt}\left(\dot{q}^i \frac{\partial L}{\partial \dot{q}^i}\right).
\end{aligned} \tag{9.1}$$

We can therefore form the difference

$$\frac{d}{dt}\left(\dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - L\right) = 0, \tag{9.2}$$

and find that this quantity, labeled H, is a constant of motion for the system

$$H \equiv \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - L = \text{constant}. \tag{9.3}$$

We will see later that this constant is sometimes the total energy of the system.

The $\dot{q}^i$ partials of the Lagrangian are called the canonical momentum conjugate to q^i. Quite a mouthful, so just canonical momenta seems like a good compromise. We will write (reserving $p^i = mq^i$ for the non-canonical momenta)

$$P_i \equiv \frac{\partial L}{\partial \dot{q}^i}. \tag{9.4}$$

and note that these are the coordinates of a sort of velocity gradient of the Lagrangian. We have seen these canonical momenta in velocity gradient form previously where it was noted that we could write the Euler-Lagrange equations in vector form in an orthonormal reciprocal frame space as

$$\nabla L = \frac{d}{dt} \nabla_v L. \tag{9.5}$$

where $\nabla_v = e^i \partial L / \partial \dot{x}^i = e^i P_i$, $\nabla = e^i \partial / \partial x^i$, and $x = e_i x^i$.

9.3 SOME SYNTACTIC SUGAR. IN VECTOR FORM.

Following Jackson [7] (section 12.1, relativistic Lorentz force Hamiltonian), this can be written in vector form if the velocity gradient, the vector sum of the momenta conjugate to the q^i's is given its own symbol **P**. He writes

$$H = \mathbf{v} \cdot \mathbf{P} - L. \tag{9.6}$$

This makes most sense when working in orthonormal coordinates, but can be generalized. Suppose we introduce a pair of reciprocal frame basis for the generalized position and velocity coordinates, writing as vectors in configuration space

$$\begin{aligned} q &= e_i q^i \\ v &= f_i \dot{q}^i. \end{aligned} \tag{9.7}$$

Following [2] (who use this for their bivector complexification of the configuration space), we have the freedom to impose orthonormal constraints on this configuration space basis

$$\begin{aligned} e^i \cdot e_j &= \delta^i{}_j \\ f^i \cdot f_j &= \delta^i{}_j \\ e^i \cdot f_j &= \delta^i{}_j. \end{aligned} \tag{9.8}$$

We can now define configuration space position and velocity gradients

$$\begin{aligned} \nabla &\equiv e^i \frac{\partial}{\partial q^i} \\ \nabla_v &\equiv f^i \frac{\partial}{\partial \dot{q}^i}, \end{aligned} \tag{9.9}$$

so the conjugate momenta in vector form is now

$$P \equiv \nabla_v L = f^i \frac{\partial L}{\partial \dot{q}^i}. \tag{9.10}$$

Our Hamiltonian takes the form

$$H = v \cdot P - L. \tag{9.11}$$

9.4 THE HAMILTONIAN PRINCIPLE.

We want to take partials of eq. (9.3) with respect to P_i and q^i. In terms of the canonical momenta we want to differentiate

$$H \equiv \dot{q}^i P_i - L(q^i, \dot{q}^i, t), \tag{9.12}$$

for the P_i partial we have

$$\frac{\partial H}{\partial P_i} = \dot{q}^i, \tag{9.13}$$

and for the q^i partial

$$\begin{aligned} \frac{\partial H}{\partial q^i} &= -\frac{\partial L}{\partial q^i} \\ &= -\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}. \end{aligned} \tag{9.14}$$

These two results taken together form what I believe is called the Hamiltonian principle

$$\begin{aligned}
\frac{\partial H}{\partial P_i} &= \dot{q}^i \\
\frac{\partial H}{\partial q^i} &= -\dot{P}_i \\
P_i &= \frac{\partial L}{\partial \dot{q}^i}.
\end{aligned} \tag{9.15}$$

A set of 2N first order equations equivalent to the second order Euler-Lagrange equations. These appear to follow straight from the definitions. Given that I am curious why the more complex method of derivation is chosen in [4]. There the total differential of the Hamiltonian is computed

$$\begin{aligned}
dH &= \dot{q}^i dP_i + d\dot{q}^i P_i - dq^i \frac{\partial L}{\partial q^i} - d\dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - dt \frac{\partial L}{\partial t} \\
&= \dot{q}^i dP_i + d\dot{q}^i \left(P_i - \frac{\partial L}{\partial \dot{q}^i} \right) - dq^i \frac{\partial L}{\partial q^i} - dt \frac{\partial L}{\partial t} \\
&= \dot{q}^i dP_i - dq^i \underbrace{\boxed{\frac{\partial L}{\partial q^i}}}_{= dP_i/dt} - dt \frac{\partial L}{\partial t}.
\end{aligned} \tag{9.16}$$

A term by term comparison to the total differential written out explicitly

$$dH = \frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial P_i} dP_i + \frac{\partial H}{\partial t} dt, \tag{9.17}$$

allows the Hamiltonian equations to be picked off.

$$\begin{aligned}
\frac{\partial H}{\partial P_i} &= \dot{q}^i \\
\frac{\partial H}{\partial q^i} &= -\dot{P}_i \\
\frac{\partial H}{\partial t} &= -\frac{\partial L}{\partial t}.
\end{aligned} \tag{9.18}$$

I guess that is not that much more complicated and it does yield a relation between the Hamiltonian and Lagrangian time derivatives.

9.5 EXAMPLES.

Now, that is just about the most abstract way we can start things off is not it? Getting some initial feel for this constant of motion can be had by considering a sequence of Lagrangians, starting with the very simplest.

9.5.1 *Force free motion.*

Our very simplest Lagrangian is that of one dimensional purely kinetic motion

$$L = \frac{1}{2}mv^2 = \frac{1}{2}m\dot{x}^2. \tag{9.19}$$

Our Hamiltonian is in this case just

$$H = \dot{x}m\dot{x} - \frac{1}{2}m\dot{x} = \frac{1}{2}mv^2. \tag{9.20}$$

The Hamiltonian is just the kinetic energy. The canonical momentum in this case is also equal to the momentum, so eliminating v to apply the Hamiltonian equations we have

$$H = \frac{1}{2m}p^2. \tag{9.21}$$

We have then

$$\begin{aligned} \frac{\partial H}{\partial p} &= \frac{p}{m} = \dot{x} \\ \frac{\partial H}{\partial x} &= 0 = -\dot{p}. \end{aligned} \tag{9.22}$$

Just for fun we can put this simple linear system in matrix form

$$\frac{d}{dt}\begin{bmatrix} p \\ x \end{bmatrix} = \frac{1}{m}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} p \\ x \end{bmatrix}. \tag{9.23}$$

A linear system of this form $y' = Ay$ can be solved by exponentiation with solution

$$y = e^{At}y_0. \tag{9.24}$$

In this case our matrix is nilpotent degree 2 so we can exponentiate only requiring up to the first order power

$$e^{At} = I + At. \tag{9.25}$$

Specifically

$$\begin{bmatrix} p \\ x \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{t}{m} & 1 \end{bmatrix} \begin{bmatrix} p_0 \\ x_0 \end{bmatrix}. \tag{9.26}$$

Written out in full this is just

$$\begin{aligned} p &= p_0 \\ x &= \frac{p_0}{m} t + x_0. \end{aligned} \tag{9.27}$$

Since the canonical momentum is the regular momentum $p = mv$ in this case, we have the usual constant rate change of position $x = v_0 t + x_0$ that we could have gotten in many easier ways. I had hazard a guess that any single variable Lagrangian that is at most quadratic in position or velocity will yield a linear system.

The generalization of this Hamiltonian to three dimensions is straightforward, and we get

$$H = \frac{1}{m} \mathbf{p}^2. \tag{9.28}$$

$$\frac{d}{dt} \begin{bmatrix} p_x \\ x \\ p_y \\ y \\ p_z \\ z \end{bmatrix} = \frac{1}{m} \begin{bmatrix} 0 & 0 & & & & \\ 1 & 0 & & & & \\ & & 0 & 0 & & \\ & & 1 & 0 & & \\ & & & & 0 & 0 \\ & & & & 1 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ x \\ p_y \\ y \\ p_z \\ z \end{bmatrix}. \tag{9.29}$$

Since there is no coupling (nilpotent matrices down the diagonal) between the coordinates this can be treated as three independent sets of equations of the form eq. (9.23), and we have

$$\begin{aligned} p_i(t) &= p_i(0) \\ x_i(t) &= \frac{p_i(0)}{m} t + x_i(0). \end{aligned} \tag{9.30}$$

Or just

$$\begin{aligned} \mathbf{p}(t) &= \mathbf{p}(0) \\ \mathbf{x}(t) &= \frac{\mathbf{p}(0)}{m} t + \mathbf{x}(0). \end{aligned} \tag{9.31}$$

9.5.2 *Linear potential (surface gravitation).*

For the gravitational force $F = -mg\hat{\mathbf{z}} = -\boldsymbol{\nabla}\phi$, we have $\phi = mgz$, and a Lagrangian of

$$L = \frac{1}{2}m\mathbf{v}^2 - \phi = \frac{1}{2}m\mathbf{v}^2 - mgz. \tag{9.32}$$

Without velocity dependence the canonical momentum is the momentum $m\mathbf{v}$, and our Hamiltonian is

$$H = \frac{1}{2m}\mathbf{p}^2 + mgz. \tag{9.33}$$

The Hamiltonian equations are

$$\begin{aligned}
\frac{\partial H}{\partial p_i} &= \dot{x}_i = \frac{1}{m}p_i \\
\sigma_i\frac{\partial H}{\partial x_i} &= -\sigma_i\dot{p}_i = \begin{bmatrix} 0 \\ 0 \\ mg \end{bmatrix}.
\end{aligned} \tag{9.34}$$

In matrix form we have

$$\frac{d}{dt}\begin{bmatrix} p_x \\ x \\ p_y \\ y \\ p_z \\ z \end{bmatrix} = \frac{1}{m}\begin{bmatrix} 0 & 0 & & & & \\ 1 & 0 & & & & \\ & & 0 & 0 & & \\ & & 1 & 0 & & \\ & & & & 0 & 0 \\ & & & & 1 & 0 \end{bmatrix}\begin{bmatrix} p_x \\ x \\ p_y \\ y \\ p_z \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -mg \\ 0 \end{bmatrix}. \tag{9.35}$$

So our problem is now reduced to solving a linear system of the form

$$y' = Ay + b. \tag{9.36}$$

That extra little term b throws a wrench into things and I am no longer sure how to integrate by inspection. What can be noted is that we really only have to consider the z components since we have solved the problem for the x and y coordinates in the force free case. That leaves

$$\frac{d}{dt}\begin{bmatrix} p_z \\ z \end{bmatrix} = \frac{1}{m}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} p_z \\ z \end{bmatrix} + \begin{bmatrix} -mg \\ 0 \end{bmatrix}. \tag{9.37}$$

Is there any reason that we have to solve in matrix form? Except for a coolness factor, not really, and we can integrate each equation directly. For the momentum equation we have

$$p_z = -mgt + p_z(0). \tag{9.38}$$

This can be substituted into the position equation for

$$\dot{z} = \frac{1}{m}(p_z(0) - mgt). \tag{9.39}$$

Direct integration is now possible for the final solution

$$\begin{aligned} z &= \frac{1}{m}(p_z(0)t - mgt^2/2) + z_0 \\ &= \frac{p_z(0)}{m}t - \frac{g}{2}t^2 + z_0. \end{aligned} \tag{9.40}$$

Again something that we could have gotten in many easier ways. Using the result we see that the solution to eq. (9.37) in matrix form, again with $A = \frac{1}{m}\left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$, is

$$\begin{bmatrix} p_z \\ z \end{bmatrix} = e^{At} \begin{bmatrix} p_z(0) \\ z(0) \end{bmatrix} - mg \begin{bmatrix} t \\ \frac{1}{2m}t^2 \end{bmatrix}. \tag{9.41}$$

I thought if I wrote this out how to solve eq. (9.36) may be more obvious, but that path is still unclear. If A were invertible, which it is not, then writing $b = Ac$ would allow for a change of variables. Does this matter for consideration of a physical problem. Not really, so I will fight the urge to play with the math for a while and perhaps revisit this later separately.

9.5.3 *Harmonic oscillator (spring potential).*

Like the free particle, the harmonic oscillator is very tractable in a phase space representation. For a restoring force $F = -kx\hat{\mathbf{x}} = -\boldsymbol{\nabla}\phi$, we have $\phi = kx^2/2$, and a Lagrangian of

$$L = \frac{1}{2}m\mathbf{v}^2 - \frac{1}{2}k\mathbf{x}^2. \tag{9.42}$$

Our Hamiltonian is again just the total energy

$$H = \frac{1}{2m}\mathbf{p}^2 + \frac{1}{2}k\mathbf{x}^2. \tag{9.43}$$

Evaluating the Hamiltonian equations we have

$$\begin{aligned}\frac{\partial H}{\partial p_i} &= \dot{x}_i = p_i/m \\ \frac{\partial H}{\partial x_i} &= -\dot{p}_i = kx_i.\end{aligned} \tag{9.44}$$

Considering just the x dimension (the others have the free particle behavior), our matrix phase space representation is

$$\frac{d}{dt}\begin{bmatrix} p \\ x \end{bmatrix} = \begin{bmatrix} 0 & -k \\ 1/m & 0 \end{bmatrix}\begin{bmatrix} p \\ x \end{bmatrix}. \tag{9.45}$$

So with

$$A = \begin{bmatrix} 0 & -k \\ 1/m & 0 \end{bmatrix}. \tag{9.46}$$

Our solution is

$$\begin{bmatrix} p \\ x \end{bmatrix} = e^{At}\begin{bmatrix} p_0 \\ x_0 \end{bmatrix}. \tag{9.47}$$

The stateful nature of the phase space solution is interesting. Provided we can make a simultaneous measurement of position and momentum, this initial state determines a next position and momentum state at a new time $t = t_0 + \Delta t_1$, and we have a trajectory through phase space of discrete transitions from one state to another

$$\begin{bmatrix} p \\ x \end{bmatrix}_{i+1} = e^{A\Delta t_{i+1}}\begin{bmatrix} p \\ x \end{bmatrix}_i. \tag{9.48}$$

Or

$$\begin{bmatrix} p \\ x \end{bmatrix}_{i+1} = e^{A\Delta t_{i+1}} e^{A\Delta t_i} \cdots e^{A\Delta t_1}\begin{bmatrix} p \\ x \end{bmatrix}_0. \tag{9.49}$$

As for solving the system, we require again the exponential of our matrix. This matrix being antisymmetric, has complex eigenvalues and again cannot be exponentiated easily by diagonalization. However, this antisymmetric matrix is very much like the complex imaginary and its square is

a negative scalar multiple of identity, so we can proceed directly forming the power series

$$A^2 = \begin{bmatrix} 0 & -k \\ 1/m & 0 \end{bmatrix} \begin{bmatrix} 0 & -k \\ 1/m & 0 \end{bmatrix} = -\frac{k}{m} I. \tag{9.50}$$

The first few powers are

$$\begin{aligned} A^2 &= -\frac{k}{m} I \\ A^3 &= -\frac{k}{m} A \\ A^4 &= \left(\frac{k}{m}\right)^2 I \\ A^5 &= \left(\frac{k}{m}\right)^2 A. \end{aligned} \tag{9.51}$$

So exponentiating we can collect cosine and sine terms

$$\begin{aligned} e^{At} &= I\left(1 - \frac{k}{m}\frac{t^2}{2!} + \left(\frac{k}{m}\right)^2 \frac{t^4}{4!} + \cdots\right) \\ &\quad + A\sqrt{\frac{m}{k}}\left(\sqrt{\frac{k}{m}} - \left(\sqrt{\frac{k}{m}}\right)^3 \frac{t^3}{3!} + \left(\sqrt{\frac{k}{m}}\right)^5 \frac{t^5}{5!}\right) \\ &= I\cos\left(\sqrt{\frac{k}{m}}t\right) + \sqrt{\frac{m}{k}} A \sin\left(\sqrt{\frac{k}{m}}t\right). \end{aligned} \tag{9.52}$$

As a check it is readily verified that this satisfies the desired $d(e^{At})/dt = Ae^{At}$ property.

The full solution in phase space representation is therefore

$$\begin{bmatrix} p \\ x \end{bmatrix} = \begin{bmatrix} p_0 \\ x_0 \end{bmatrix} \cos\left(\sqrt{\frac{k}{m}}t\right) + \sqrt{\frac{m}{k}} \begin{bmatrix} -kx_0 \\ p_0/m \end{bmatrix} \sin\left(\sqrt{\frac{k}{m}}t\right). \tag{9.53}$$

Written out separately this is clearer

$$\begin{aligned} p &= p_0 \cos\left(\sqrt{\frac{k}{m}}t\right) - \sqrt{\frac{m}{k}} k x_0 \sin\left(\sqrt{\frac{k}{m}}t\right) \\ x &= x_0 \cos\left(\sqrt{\frac{k}{m}}t\right) + \sqrt{\frac{m}{k}} \frac{p_0}{m} \sin\left(\sqrt{\frac{k}{m}}t\right). \end{aligned} \tag{9.54}$$

One can readily verify that $m\dot{x} = p$, and $m\ddot{x} = -kx$ as expected.

Let us pause before leaving the harmonic oscillator to see if eq. (9.54) seems to make sense. Consider the position solution. With only initial position and no initial velocity p_0/m we have oscillation that has no dependence on the mass or spring constant. This is the unmoving mass about to be let go at the end of a spring case, and since we have no damping force the magnitude of the oscillation is exactly the initial position of the mass. If the instantaneous velocity is measured at position zero, it makes sense in this case that the oscillation amplitude does depend on both the mass and the spring constant. The stronger the spring (k), the bigger the oscillation, and the smaller the mass, the bigger the oscillation.

It is definitely no easier to work with the phase space formulation than just solving the second order system directly. The fact that we have a linear system to solve, at least in this particular case is kind of nice. Perhaps this methodology can be helpful considering linear approximation solutions in a neighborhood of some phase space point for more complex non-linear systems.

9.5.4 *Harmonic oscillator (change of variables.)*

It was pointed out to me by Lut that the following rather strange looking change of variables has nice properties

$$\begin{aligned} P &= x\sqrt{\frac{k}{2}} + \frac{p}{\sqrt{2m}} \\ Q &= x\sqrt{\frac{k}{2}} - \frac{p}{\sqrt{2m}}. \end{aligned} \tag{9.55}$$

In particular the Hamiltonian is then just

$$H = P^2 + Q^2. \tag{9.56}$$

Part of this change of variables, which rotates in phase space, as well as scales, looks like just a way of putting the system into natural units. We do not however, need the rotation to do that. Suppose we write for just the scaling change of variables

$$\begin{aligned} p &= \sqrt{2m}P_s \\ x &= \sqrt{\frac{2}{k}}Q_s. \end{aligned} \tag{9.57}$$

or

$$\begin{bmatrix} p \\ x \end{bmatrix} = \begin{bmatrix} \sqrt{2m} & 0 \\ 0 & \sqrt{\frac{2}{k}} \end{bmatrix} \begin{bmatrix} P_s \\ Q_s \end{bmatrix}. \tag{9.58}$$

This also gives the Hamiltonian eq. (9.56), and the Hamiltonian equations are transformed to

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} P_s \\ Q_s \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2m} & 0 \\ 0 & \sqrt{\frac{k}{2}} \end{bmatrix} \begin{bmatrix} 0 & -k \\ 1/m & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2m} & 0 \\ 0 & \sqrt{\frac{2}{k}} \end{bmatrix} \begin{bmatrix} P_s \\ Q_s \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\sqrt{\frac{k}{m}} \\ \sqrt{\frac{k}{m}} & 0 \end{bmatrix} \begin{bmatrix} P_s \\ Q_s \end{bmatrix}. \end{aligned} \tag{9.59}$$

This first change of variables is nice since it groups the two factors k and m into a reciprocal pair. Since only the ratio is significant to the kinetics it is nice to have that explicit. Since $\sqrt{k/m}$ is in fact the angular frequency, we can define

$$\omega \equiv \sqrt{\frac{k}{m}}, \tag{9.60}$$

and our system is reduced to

$$\frac{d}{dt}\begin{bmatrix} P_s \\ Q_s \end{bmatrix} = \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} P_s \\ Q_s \end{bmatrix}. \tag{9.61}$$

Solution of this system now becomes particularly easy, especially if one notes that the matrix factor above can be expressed in terms of the y axis Pauli matrix σ_2. That is

$$\sigma_2 = i \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{9.62}$$

Inverting this, and labeling this matrix $\mathcal{I}$ we can write

$$\mathcal{I} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = -i\sigma_2. \tag{9.63}$$

Recalling that $\sigma_2^2 = I$, we then have $\mathcal{I}^2 = -I$, and see that this matrix behaves exactly like a unit imaginary. This reduces the Hamiltonian system to

$$\frac{d}{dt}\begin{bmatrix} P_s \\ Q_s \end{bmatrix} = \mathcal{I}\omega \begin{bmatrix} P_s \\ Q_s \end{bmatrix}. \tag{9.64}$$

We can now solve the system directly. Writing $\mathbf{z}_s = \left[\begin{smallmatrix} P_s \\ Q_s \end{smallmatrix}\right]$, this is just

$$\mathbf{z}_s(t) = e^{\mathcal{I}\omega t}\mathbf{z}_s(0) = \left(I\cos(\omega t) + \mathcal{I}\sin(\omega t)\right)\mathbf{z}_s(0). \tag{9.65}$$

With just the scaling giving both the simple Hamiltonian, and a simple solution, what is the advantage of the further change of variables that mixes (rotates in phase space by 45 degrees with a factor of two scaling) the momentum and position coordinates? That second transformation is

$$\begin{aligned} P &= Q_s + P_s \\ Q &= Q_s - P_s. \end{aligned} \tag{9.66}$$

Inverting this we have

$$\begin{bmatrix} P_s \\ Q_s \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} P \\ Q \end{bmatrix}. \tag{9.67}$$

The Hamiltonian after this change of variables is now

$$\frac{d}{dt}\begin{bmatrix} P \\ Q \end{bmatrix} = \frac{\omega}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} P \\ Q \end{bmatrix}. \tag{9.68}$$

But multiplying this out one finds that the equations of motion for the state space vector are unchanged by the rotation, and writing $\mathbf{z} = \left(\begin{smallmatrix} P \\ Q \end{smallmatrix}\right)$ for the state vector, the Hamiltonian equations are

$$\mathbf{z}' = \mathcal{I}\omega\mathbf{z}. \tag{9.69}$$

This is just as we had before the rotation-like mixing of position and momentum coordinates. Now it looks like the rotational change of coordinates is related to the raising and lowering operators in the Quantum treatment of the Harmonic oscillator, but it is not clear to me what the advantage is in the classical context? Perhaps the point is, that at least for the classical Harmonic oscillator, we are free to rotate the phase space vector

arbitrarily and not change the equations of motion. A restriction to the classical domain is required since squaring the results of this 45 degree rotation of the phase space vector requires commutation of the position and momentum coordinates in order for the cross terms to cancel out.

Is there a deeper meaning to this rotational freedom? It seems to me that one ought to be able to relate the rotation and the quantum ladder operators in a more natural way, but it is not clear to me exactly how.

9.5.5 *Force free system dependent on only differences.*

In gravitational and electrostatic problems are forces are all functions of only the difference in positions of the particles. Lets look at how the purely kinetic Lagrangian and Hamiltonian change when one or more of the vector positions is reexpressed in terms of a difference in position change of variables. In the force free case this is primarily a task of rewriting the Hamiltonian in terms of the conjugate momenta after such a change of variables.

The very simplest case is the two particle single dimensional Kinetic Lagrangian,

$$L = \frac{1}{2}m_1\dot{r}_1^2 + \frac{1}{2}m_2\dot{r}_2^2. \tag{9.70}$$

With a change of variables

$$\begin{aligned} x &= r_2 - r_1 \\ y &= r_2, \end{aligned} \tag{9.71}$$

and elimination of r_1, and r_2 we have

$$L = \frac{1}{2}m_1(\dot{y} - \dot{x})^2 + \frac{1}{2}m_2\dot{y}^2. \tag{9.72}$$

We now need the conjugate momenta in terms of $\dot{x}$ and $\dot{y}$. These are

$$\begin{aligned} P_x &= \frac{\partial L}{\partial \dot{x}} = -m_1(\dot{y} - \dot{x}) \\ P_y &= \frac{\partial L}{\partial \dot{y}} = m_1(\dot{y} - \dot{x}) + m_2\dot{y}. \end{aligned} \tag{9.73}$$

We must now rewrite the Lagrangian in terms of P_x and P_y, essentially requiring the inversion of this which amounts to the inversion of the two by two linear system of eq. (9.73). That is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} m_1 & -m_1 \\ -m_1 & (m_1 + m_2) \end{bmatrix}^{-1} \begin{bmatrix} P_x \\ P_y \end{bmatrix}. \tag{9.74}$$

This is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \frac{1}{m_1}\frac{1}{m_2}\begin{bmatrix} m_1 + m_2 & m_1 \\ m_1 & m_1 \end{bmatrix}\begin{bmatrix} P_x \\ P_y \end{bmatrix}. \tag{9.75}$$

Of these only $\dot{y}$ and $\dot{y} - \dot{x}$ are of interest and after a bit of manipulation we find

$$\begin{aligned} \dot{y} &= \frac{1}{m_2}(P_x + P_y) \\ \dot{x} &= \frac{1}{m_1}\frac{1}{m_2}((m_1 + m_2)P_x + m_1 P_y). \end{aligned} \tag{9.76}$$

From this we find the Lagrangian in terms of the conjugate momenta

$$L = \frac{1}{2m_1}{P_x}^2 + \frac{1}{2m_2}(P_x + P_y)^2. \tag{9.77}$$

A quick check shows that $P_x + P_y = m_2\dot{r}_2$, and $P_x = -m_1\dot{r}_1$, so we have agreement with the original Lagrangian. Generalizing to the three dimensional case is straightforward, and we have

$$L = \frac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \frac{1}{2}m_2\dot{\mathbf{r}}_2^2 - \phi(\mathbf{x}_1 - \mathbf{x}_2). \tag{9.78}$$

With

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_1 - \mathbf{x}_2 \\ \mathbf{y} &= \mathbf{x}_2. \end{aligned} \tag{9.79}$$

The 3D generalization of the above (following by adding indices then summing) becomes

$$\begin{aligned} \mathbf{P}_x &= \sigma_j\frac{\partial L}{\partial \dot{x}^j} = -m_1(\dot{\mathbf{y}} - \dot{\mathbf{x}}) \\ \mathbf{P}_y &= \sigma_j\frac{\partial L}{\partial \dot{y}^j} = m_1(\dot{\mathbf{y}} - \dot{\mathbf{x}}) + m_2\dot{\mathbf{y}}. \end{aligned} \tag{9.80}$$

$$\begin{aligned} L &= \frac{1}{2m_1}{\mathbf{P}_x}^2 + \frac{1}{2m_2}(\mathbf{P}_x + \mathbf{P}_y)^2 - \phi(\mathbf{x}) \\ H &= \frac{1}{2m_1}{\mathbf{P}_x}^2 + \frac{1}{2m_2}(\mathbf{P}_x + \mathbf{P}_y)^2 + \phi(\mathbf{x}). \end{aligned} \tag{9.81}$$

Finally, evaluation of the Hamiltonian equations we have

$$\begin{aligned}
\sigma_j \frac{\partial H}{\partial P_x^j} &= \dot{\mathbf{x}} \\
&= \sigma_j \left(\frac{1}{m_1} P_x^j + \frac{1}{m_2} (P_x^j + P_y^j) \right) \\
&= \frac{1}{m_1} \mathbf{P}_x + \frac{1}{m_2} (\mathbf{P}_x + \mathbf{P}_y)
\end{aligned} \tag{9.82}$$

$$\begin{aligned}
\sigma_j \frac{\partial H}{\partial P_y^j} &= \dot{\mathbf{y}} \\
&= \sigma_j \frac{1}{m_2} (P_x^j + P_y^j) \\
&= \frac{1}{m_2} (\mathbf{P}_x + \mathbf{P}_y)
\end{aligned} \tag{9.83}$$

$$\begin{aligned}
\sigma_j \frac{\partial H}{\partial x^j} &= -\dot{\mathbf{P}}_x \\
&= -\sigma_j \frac{\partial \phi}{\partial x^j} \\
&= -\boldsymbol{\nabla}_{\mathbf{x}} \phi(\mathbf{x})
\end{aligned} \tag{9.84}$$

$$\begin{aligned}
\sigma_j \frac{\partial H}{\partial y^j} &= -\dot{\mathbf{P}}_y \\
&= -\sigma_j \frac{\partial \phi}{\partial y^j} \\
&= 0.
\end{aligned} \tag{9.85}$$

Summarizing we have four first order equations

$$\begin{aligned}
\dot{\mathbf{x}} &= \left(\frac{1}{m_1} + \frac{1}{m_2} \right) \mathbf{P}_x + \frac{1}{m_2} \mathbf{P}_y \\
\dot{\mathbf{y}} &= \frac{1}{m_2} (\mathbf{P}_x + \mathbf{P}_y) \\
\dot{\mathbf{P}}_x &= \boldsymbol{\nabla}_{\mathbf{x}} \phi(\mathbf{x}) \\
\dot{\mathbf{P}}_y &= 0.
\end{aligned} \tag{9.86}$$

FIXME: what would we get if using the center of mass position as one of the variables. A parametrization with three vector variables should also still work, even if it includes additional redundancy.

9.5.6 *Gravitational potential.*

Next I had like to consider a two particle gravitational interaction. However, to start we need the Lagrangian, and what should the potential term be in a two particle gravitational Lagrangian? I had guess something with a $1/x$ form, but do we need one potential term for each mass or something interrelated? Whatever the Lagrangian is, we want it to produce the pair of force relationships

$$\begin{aligned}\text{Force on 2} &= -Gm_1m_2\frac{(\mathbf{r}_2-\mathbf{r}_1)}{|\mathbf{r}_2-\mathbf{r}_1|}\\ \text{Force on 1} &= Gm_1m_2\frac{(\mathbf{r}_2-\mathbf{r}_1)}{|\mathbf{r}_2-\mathbf{r}_1|}.\end{aligned} \tag{9.87}$$

Guessing that the Lagrangian has a single term for the interaction potential

$$\phi_{21} = \kappa\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}. \tag{9.88}$$

so that we have

$$L = \frac{1}{2}m\mathbf{v}_1{}^2 + \frac{1}{2}m\mathbf{v}_2{}^2 - \phi_{21}. \tag{9.89}$$

We can evaluate the Euler-Lagrange equations and see if the result is consistent with the Newtonian force laws of eq. (9.87). Suppose we write the coordinates of $\mathbf{r}_i$ as $x^k{}_i$. There are then six Euler-Lagrange equations

$$\begin{aligned}\frac{\partial L}{\partial x^j{}_i} &= \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^j_i}\\ -\frac{\partial \phi_{21}}{\partial x^j{}_i} &= m_i\ddot{x}^j_i.\end{aligned} \tag{9.90}$$

Evaluating the potential derivatives separately. Consider the $i = 2$ derivative

$$\begin{aligned}\frac{\partial \phi_{21}}{\partial x^j{}_2} &= \kappa\frac{\partial}{\partial x^j{}_2}\left(\sum_k (x^k{}_2 - x^k{}_1)^2\right)^{-1/2}\\ &= -\kappa\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|^3}\sum_k (x^k{}_2 - x^k{}_1)\frac{\partial}{\partial x^j{}_2}(x^k{}_2 - x^k{}_1)\\ &= -\kappa\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|^3}(x^j{}_2 - x^j{}_1).\end{aligned} \tag{9.91}$$

Therefore the final result of the Euler-Lagrange equations is

$$\begin{aligned} \kappa \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|^3}(x^j{}_2 - x^j{}_1) &= m_2 \ddot{x}_2^j \\ -\kappa \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|^3}(x^j{}_2 - x^j{}_1) &= m_1 \ddot{x}_1^j. \end{aligned} \tag{9.92}$$

which confirms the Lagrangian and potential guess and fixes the constant $\kappa = -Gm_1m_2$. With the sign fixed, our potential, Lagrangian, and Hamiltonian are respectively

$$\begin{aligned} \phi_{21} &= -\frac{Gm_2m_1}{|\mathbf{r}_2 - \mathbf{r}_1|} \\ L &= \frac{1}{2}m_1\mathbf{v}_1{}^2 + \frac{1}{2}m_2\mathbf{v}_2{}^2 - \phi_{21} \\ H &= \frac{1}{2m_1}\mathbf{p}_1{}^2 + \frac{1}{2m_2}\mathbf{p}_2{}^2 + \phi_{21}. \end{aligned} \tag{9.93}$$

There is however an undesirable asymmetry to this expression, in particular a formulation that extends to multiple particles seems desirable. Let us write instead a slight variation

$$\phi_{ij} = -\frac{Gm_im_j}{\left|\mathbf{r}_i - \mathbf{r}_j\right|}, \tag{9.94}$$

and form a scaled by two double summation over all pairs of potentials

$$L = \sum_i \frac{1}{2}m_i\mathbf{v}_i{}^2 - \frac{1}{2}\sum_{i\neq j}\phi_{ij}. \tag{9.95}$$

Having established what seems like an appropriate form for the Lagrangian, we can write the Hamiltonian for the multiparticle gravitational interaction by inspection

$$H = \sum_i \frac{1}{2m_i}\mathbf{p}_i{}^2 + \frac{1}{2}\sum_{i\neq j}\phi_{ij}. \tag{9.96}$$

This leaves us finally in position to evaluate the Hamiltonian equations, but the result of doing so is rudely nothing more than the Newtonian equations in coordinate form. We get, for the kth component of the ith particle

$$\frac{\partial H}{\partial p^k{}_i} = \dot{x}_i^k = \frac{1}{m_i}p^k{}_i, \tag{9.97a}$$

$$\frac{\partial H}{\partial x^k{}_i} = -\dot{p}^k_i = G \sum_{j \neq i} m_i m_j \frac{x^k{}_i - x^k{}_j}{\left|\mathbf{r}_i - \mathbf{r}_j\right|^3}. \tag{9.97b}$$

The state space vector for this system of equations is brutally ugly, and could be put into the following form for example

$$\mathbf{z} = \begin{bmatrix} p^1{}_1 \\ p^2{}_1 \\ p^3{}_1 \\ x^1{}_1 \\ x^2{}_1 \\ x^3{}_1 \\ p^1{}_2 \\ p^2{}_2 \\ p^3{}_2 \\ x^1{}_2 \\ \vdots \end{bmatrix}. \tag{9.98}$$

Where the Hamiltonian equations take the form of a non-linear function on such state space vectors We have a somewhat sparse equation of the form

$$\frac{d\mathbf{z}}{dt} = A(\mathbf{z}). \tag{9.99}$$

One thing that is possible in such a representation is calculating the first order approximate change in position and momentum moving from one time to a small time later

$$\mathbf{z}(t_0 + \Delta t) = \mathbf{z}(t_0) + A(\mathbf{z}(t_0))\Delta t. \tag{9.100}$$

One could conceivably calculate the trajectories in phase space using such increments, and if a small enough time increment is used this can be thought of as solving the gravitational system. I recall that Feynman did something like this in his lectures, but set up the problem in a more computationally efficient form (it definitely did not have the redundancy built into the Hamiltonian equations).

FIXME: should be able to solve this for an arbitrary Δt later time if this was extended to the higher order terms. Need something like the $e^{z\cdot\nabla}$ chain rule expansion. Think this through. Will be a little different since we are already starting with the first order contribution.

What does this system of equations look like with a reduction of order through center of mass change of variables?

9.5.7 *Pendulum.*

FIXME: picture. x-axis down, y-axis right. The bob speed for a stiff rod of length l is $(l\dot{\theta})^2$, and our potential is $mgh = mgl(1-\cos\theta)$. The Lagrangian is therefore

$$L = \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1-\cos\theta). \tag{9.101}$$

The constant mgl term can be dropped, and our canonical momentum conjugate to $\dot{\theta}$ is $p_\theta = ml^2\dot{\theta}$, so our Hamiltonian is

$$H = \frac{1}{2ml^2}{p_\theta}^2 - mgl\cos\theta. \tag{9.102}$$

We can now compute the Hamiltonian equations

$$\begin{aligned}\frac{\partial H}{\partial p_\theta} &= \dot{\theta} = \frac{1}{ml^2}p_\theta \\ \frac{\partial H}{\partial q} &= -\dot{p}_\theta = mgl\sin\theta.\end{aligned} \tag{9.103}$$

Only in the neighborhood of a particular angle can we write this in matrix form. Suppose we expand this around $\theta = \theta_0 + \alpha$. The sine is then

$$\sin\theta \approx \sin\theta_0 + \cos\theta_0\alpha. \tag{9.104}$$

The linear approximation of the Hamiltonian equations after a change of variables become

$$\frac{d}{dt}\begin{bmatrix} p_\theta \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 & -mgl\cos\theta_0 \\ 1/ml^2 & 0 \end{bmatrix}\begin{bmatrix} p_\theta \\ \alpha \end{bmatrix} + \begin{bmatrix} -mgl\sin\theta_0 \\ \dot{\theta}_0 \end{bmatrix}. \tag{9.105}$$

A change of variables that scales the factors in the matrix to have equal magnitude and equivalent dimensions is helpful. Writing

$$\begin{bmatrix} p_\theta \\ \alpha \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}\mathbf{z}. \tag{9.106}$$

one finds

$$\frac{d\mathbf{z}}{dt} = \begin{bmatrix} 0 & -mgl\cos\theta_0/a \\ a/ml^2 & 0 \end{bmatrix}\mathbf{z} + \frac{1}{a}\begin{bmatrix} -mgl\sin\theta_0 \\ \dot{\theta}_0 \end{bmatrix}. \tag{9.107}$$

To tidy this up, we want

$$\left|\frac{a}{ml^2}\right| = \left|\frac{mgl\cos\theta_0}{a}\right| \tag{9.108}$$

Or

$$a = ml^2\sqrt{\frac{g}{l}|\cos\theta_0|}. \tag{9.109}$$

The result of applying this scaling is quite different above and below the horizontal due to the sign difference in the cosine. Below the horizontal where $\theta_0 \in (-\pi/2, \pi/2)$ we get

$$\frac{d\mathbf{z}}{dt} = \sqrt{\frac{g}{l}\cos\theta_0}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\mathbf{z} + \frac{1}{ml^2\sqrt{\frac{g}{l}\cos\theta_0}}\begin{bmatrix} -mgl\sin\theta_0 \\ \dot{\theta}_0 \end{bmatrix}. \tag{9.110}$$

and above the horizontal where $\theta_0 \in (\pi/2, 3\pi/2)$ we get

$$\frac{d\mathbf{z}}{dt} = \sqrt{-\frac{g}{l}\cos\theta_0}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\mathbf{z} + \frac{1}{ml^2\sqrt{-\frac{g}{l}\cos\theta_0}}\begin{bmatrix} -mgl\sin\theta_0 \\ \dot{\theta}_0 \end{bmatrix}. \tag{9.111}$$

Since $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has the characteristics of an imaginary number (squaring to the negative of the identity) the homogeneous part of the solution for the change of the phase space vector in the vicinity of any initial angle in the lower half plane is trigonometric. Similarly the solutions are necessarily hyperbolic in the upper half plane since $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ squares to identity. And around $\pm\pi/2$ something totally different (return to this later). The problem is now reduced to solving a non-homogeneous first order matrix equation of the form

$$\mathbf{z}' = \Omega\mathbf{z} + \mathbf{b}. \tag{9.112}$$

But we have the good fortune of being able to easily exponentiate and invert this matrix Ω. The homogeneous problem

$$\mathbf{z}' = \Omega\mathbf{z}. \tag{9.113}$$

has the solution

$$\mathbf{z}_h(t) = e^{\Omega t}\mathbf{z}_{t=0}. \tag{9.114}$$

Assuming a specific solution $z = e^{\Omega t} f(t)$ for the non-homogeneous equation, one finds $z = \Omega^{-1}(e^{\Omega t} - I)\mathbf{b}$. The complete solution with both the homogeneous and non-homogeneous parts is thus

$$\mathbf{z}(t) = e^{\Omega t}\mathbf{z}_0 + \Omega^{-1}(e^{\Omega t} - I)\mathbf{b}. \tag{9.115}$$

Going back to the pendulum problem, lets write

$$\omega = \sqrt{\frac{g}{l}|\cos\theta_0|}. \tag{9.116}$$

Below the horizontal we have

$$\begin{aligned}
\Omega &= \omega\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\
\Omega^{-1} &= -\frac{1}{\omega}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\
e^{\Omega t} &= \cos(\omega t)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sin(\omega t)\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.
\end{aligned} \tag{9.117}$$

Whereas above the horizontal we have

$$\begin{aligned}
\Omega &= \omega\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
\Omega^{-1} &= \frac{1}{\omega}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
e^{\Omega t} &= \cosh(\omega t)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sinh(\omega t)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
\end{aligned} \tag{9.118}$$

In both cases we have

$$\begin{aligned}
\begin{bmatrix} p_\theta \\ \alpha \end{bmatrix} &= \begin{bmatrix} ml^2\omega & 0 \\ 0 & 1 \end{bmatrix}\mathbf{z} \\
\mathbf{b} &= \frac{1}{\omega}\begin{bmatrix} -\frac{g}{l}\sin\theta_0 \\ \frac{\dot{\theta}_0}{ml^2} \end{bmatrix}.
\end{aligned} \tag{9.119}$$

(where the real angle was $\theta = \theta_0 + \alpha$). Since in this case Ω^{-1} and $e^{\Omega t}$ commute, we have below the horizontal

$$\begin{aligned}\mathbf{z}(t) &= e^{\Omega t}(\mathbf{z}_0 - \Omega^{-1}\mathbf{b}) - \Omega^{-1}\mathbf{b} \\ &= \left(\cos(\omega t)\left[\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right] + \sin(\omega t)\left[\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right]\right)\left(\mathbf{z}_0 + \frac{1}{\omega}\left[\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right]\mathbf{b}\right) + \frac{1}{\omega}\left[\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right]\mathbf{b}.\end{aligned} \tag{9.120}$$

Expanding out the $\mathbf{b}$ terms and doing the same for above the horizontal we have respectively (below and above)

$$\begin{aligned}\mathbf{z}_{\text{low}}(t) &= \left(\cos(\omega t)\left[\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right] + \sin(\omega t)\left[\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right]\right)\left(\mathbf{z}_0 - \frac{1}{\omega^2}\begin{bmatrix}\frac{\dot\theta_0}{ml^2}\\ \frac{g}{l}\sin\theta_0\end{bmatrix}\right) - \frac{1}{\omega^2}\begin{bmatrix}\frac{\dot\theta_0}{ml^2}\\ \frac{g}{l}\sin\theta_0\end{bmatrix} \\ \mathbf{z}_{\text{high}}(t) &= \left(\cosh(\omega t)\left[\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right] + \sinh(\omega t)\left[\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right]\right)\left(\mathbf{z}_0 + \frac{1}{\omega^2}\begin{bmatrix}\frac{\dot\theta_0}{ml^2}\\ \frac{g}{l}\sin\theta_0\end{bmatrix}\right) + \frac{1}{\omega^2}\begin{bmatrix}\frac{\dot\theta_0}{ml^2}\\ \frac{g}{l}\sin\theta_0\end{bmatrix}.\end{aligned} \tag{9.121}$$

The only thing that is really left is re-insertion of the original momentum and position variables using the inverse relation

$$\mathbf{z} = \begin{bmatrix}1/(ml^2\omega) & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}p_\theta\\ \theta - \theta_0\end{bmatrix}. \tag{9.122}$$

Will that final insertion do anything more than make things messier? Observe that the $\mathbf{z}_0$ only has a momentum component when expressed back in terms of the total angle θ. Also recall that $p_\theta = ml^2\dot\theta$, so we have

$$\begin{aligned}\mathbf{z} &= \begin{bmatrix}\dot\theta/\omega\\ \theta - \theta_0\end{bmatrix} \\ \mathbf{z}_0 &= \begin{bmatrix}\dot\theta_{t=0}/\omega\\ 0\end{bmatrix}.\end{aligned} \tag{9.123}$$

If this is somehow mystically free of all math mistakes then we have the final solution

$$\begin{aligned}
\begin{bmatrix} \dot\theta(t)/\omega \\ \theta(t)-\theta_0 \end{bmatrix}_{\text{low}} &= \left(\cos(\omega t)\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] + \sin(\omega t)\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]\right)\left(\frac{\dot\theta(0)}{\omega}\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{1}{\omega^2}\begin{bmatrix} \frac{\dot\theta_0}{ml^2} \\ \frac{g}{l}\sin\theta_0 \end{bmatrix}\right) \\
&\quad - \frac{1}{\omega^2}\begin{bmatrix} \frac{\dot\theta_0}{ml^2} \\ \frac{g}{l}\sin\theta_0 \end{bmatrix} \\
\begin{bmatrix} \dot\theta(t)/\omega \\ \theta(t)-\theta_0 \end{bmatrix}_{\text{high}} &= \left(\cosh(\omega t)\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] + \sinh(\omega t)\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]\right)\left(\frac{\dot\theta(0)}{\omega}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{\omega^2}\begin{bmatrix} \frac{\dot\theta_0}{ml^2} \\ \frac{g}{l}\sin\theta_0 \end{bmatrix}\right) \\
&\quad + \frac{1}{\omega^2}\begin{bmatrix} \frac{\dot\theta_0}{ml^2} \\ \frac{g}{l}\sin\theta_0 \end{bmatrix}.
\end{aligned} \tag{9.124}$$

A qualification is required to call this a solution since it is only a solution is the restricted range where θ is close enough to θ_0 (in some imprecisely specified sense). One could conceivably apply this in a recursive fashion however, calculating the result for a small incremental change, yielding the new phase space point, and repeating at the new angle.

The question of what the form of the solution in the neighborhood of $\pm\pi/2$ has also been ignored. That is probably also worth considering but I do not feel like trying now.

9.5.8 *Spherical pendulum.*

For the spherical rigid pendulum of length l, we have for the distance above the lowest point

$$h = l(1+\cos\theta). \tag{9.125}$$

(measuring θ down from the North pole as conventional). The Lagrangian is therefore

$$L = \frac{1}{2}ml^2(\dot\theta^2 + \sin^2\theta\dot\phi^2) - mgl(1+\cos\theta). \tag{9.126}$$

We can drop the constant term, using the simpler Lagrangian

$$L = \frac{1}{2}ml^2(\dot\theta^2 + \sin^2\theta\dot\phi^2) - mgl\cos\theta. \tag{9.127}$$

To express the Hamiltonian we need first the conjugate momenta, which are

$$\begin{aligned} P_\theta &= \frac{\partial L}{\partial \dot{\theta}} = ml^2\dot{\theta} \\ P_\phi &= \frac{\partial L}{\partial \dot{\phi}} = ml^2 \sin^2\theta\dot{\phi}. \end{aligned} \tag{9.128}$$

We can now write the Hamiltonian

$$H = \frac{1}{2ml^2}\left({P_\theta}^2 + \frac{1}{\sin^2\theta}{P_\phi}^2\right) + mgl\cos\theta. \tag{9.129}$$

Before going further one sees that there is going to be trouble where $\sin\theta = 0$. Curiously, this is at the poles, the most dangling position and the upright. The south pole is the usual point where we solve the planar pendulum problem using the harmonic oscillator approximation, so it is somewhat curious that the energy of the system appears to go undefined at this point where the position is becoming more defined. It seems almost like a quantum uncertainty phenomena until one realizes that the momentum conjugate to ϕ is itself proportional to $\sin^2\theta$. By expressing the energy in terms of this P_ϕ momentum we have to avoid looking at the poles for a solution to the equations. If we go back to the Lagrangian and the Euler-Lagrange equations, this point becomes perfectly tractable since we are no longer dividing through by $\sin^2\theta$.

Examining the polar solutions is something to return to. For now, let us avoid that region. For regions where $\sin\theta$ is nicely non-zero, we get for the Hamiltonian equations

$$\begin{aligned} \frac{\partial H}{\partial P_\phi} &= \dot{\phi} = \frac{1}{ml^2\sin^2\theta}P_\phi \\ \frac{\partial H}{\partial P_\theta} &= \dot{\theta} = \frac{1}{ml^2}P_\theta \\ \frac{\partial H}{\partial \phi} &= -\dot{P}_\phi = 0 \\ \frac{\partial H}{\partial \theta} &= -\dot{P}_\theta = -\frac{\cos\theta}{ml^2\sin^3\theta}{P_\phi}^2 - mgl\sin\theta. \end{aligned} \tag{9.130}$$

These now expressing the dynamics of the system. The first two equations are just the definitions of the canonical momenta that we started with using the Lagrangian. Not surprisingly, but unfortunate, we have a non-linear system here like the planar rigid pendulum, so despite this being

one of the most simple systems it does not look terribly tractable. What would it take to linearize this system of equations?

Write the state space vector for the system as

$$\mathbf{x} = \begin{bmatrix} P_\theta \\ \theta \\ P_\phi \\ \phi \end{bmatrix}. \tag{9.131}$$

lets also suppose that we are interested in the change to the state vector in the neighborhood of an initial state

$$\mathbf{x} = \begin{bmatrix} P_\theta \\ \theta \\ P_\phi \\ \phi \end{bmatrix} = \begin{bmatrix} P_\theta \\ \theta \\ P_\phi \\ \phi \end{bmatrix}_0 + \mathbf{z}. \tag{9.132}$$

The Hamiltonian equations can then be written

$$\frac{d\mathbf{z}}{dt} = \begin{bmatrix} \frac{\cos\theta}{ml^2\sin^3\theta}{P_\phi}^2 + mgl\sin\theta \\ \frac{1}{ml^2}P_\theta \\ 0 \\ \frac{1}{ml^2\sin^2\theta}P_\phi \end{bmatrix}. \tag{9.133}$$

Getting away from the specifics of this particular system is temporarily helpful. We have a set of equations that we wish to calculate a linear approximation for

$$\frac{dz_\mu}{dt} = A_\mu(x_\nu) \approx A_\mu(\mathbf{x}_0) + \sum_\alpha \left.\frac{\partial A_\mu}{\partial x_\alpha}\right|_{\mathbf{x}_0} z_\alpha. \tag{9.134}$$

Our linear approximation is thus

$$\frac{d\mathbf{z}}{dt} \approx \begin{bmatrix} \frac{\cos\theta}{ml^2\sin^3\theta}{P_\phi}^2 + mgl\sin\theta \\ \frac{1}{ml^2}P_\theta \\ 0 \\ \frac{1}{ml^2\sin^2\theta}P_\phi \end{bmatrix}_0 + \begin{bmatrix} 0 & -\frac{{P_\phi}^2(1+2\cos^2\theta)}{ml^2\sin^4\theta} + mgl\cos\theta & \frac{2\cos\theta}{ml^2\sin^3\theta}P_\phi & 0 \\ \frac{1}{ml^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{-2P_\phi\cos\theta}{ml^2\sin^3\theta} & \frac{1}{ml^2\sin^2\theta} & 0 \end{bmatrix}_0 \mathbf{z}. \tag{9.135}$$

Now, this is what we get blindly trying to set up the linear approximation of the state space differential equation. We see that the cyclic coordinate ϕ leads to a bit of trouble since no explicit ϕ dependence in the Hamiltonian makes the resulting matrix factor non-invertible. It appears that we would be better explicitly utilizing this cyclic coordinate to note that P_ϕ = constant, and to omit this completely from the state vector. Our equations in raw form are now

$$\begin{aligned} \dot\theta &= \frac{1}{ml^2}P_\theta \\ \dot P_\theta &= \frac{\cos\theta}{ml^2\sin^3\theta}{P_\phi}^2 + mgl\sin\theta \\ \dot\phi &= \frac{1}{ml^2\sin^2\theta}P_\phi. \end{aligned} \tag{9.136}$$

We can treat the ϕ dependence later once we have solved for θ. That equation to later solve is just this last

$$\dot\phi = \frac{1}{ml^2\sin^2\theta}P_\phi. \tag{9.137}$$

This integrates directly, presuming $\theta = \theta(t)$ is known, and we have

$$\phi - \phi(0) = \frac{P_\phi}{ml^2}\int_0^t \frac{d\tau}{\sin^2\theta(\tau)}. \tag{9.138}$$

Now the state vector and its perturbation can be redefined omitting all but the θ dependence. Namely

$$\mathbf{x} = \begin{bmatrix} P_\theta \\ \theta \end{bmatrix}. \tag{9.139}$$

$$\mathbf{x} = \begin{bmatrix} P_\theta \\ \theta \end{bmatrix} = \begin{bmatrix} P_\theta \\ \theta \end{bmatrix}_0 + \mathbf{z}. \tag{9.140}$$

We can now write the remainder of this non-linear system as

$$\frac{d\mathbf{z}}{dt} = \begin{bmatrix} \frac{\cos\theta}{ml^2\sin^3\theta}{P_\phi}^2 + mgl\sin\theta \\ \frac{1}{ml^2}P_\theta \end{bmatrix}. \tag{9.141}$$

and make the linear approximation around $\mathbf{x}_0$ as

$$\frac{d\mathbf{z}}{dt} \approx \begin{bmatrix} \frac{\cos\theta}{ml^2\sin^3\theta}{P_\phi}^2 + mgl\sin\theta \\ \frac{1}{ml^2}P_\theta \end{bmatrix}_0 + \begin{bmatrix} 0 & -\frac{{P_\phi}^2(1+2\cos^2\theta)}{ml^2\sin^4\theta} + mgl\cos\theta \\ \frac{1}{ml^2} & 0 \end{bmatrix}_0 \mathbf{z}. \tag{9.142}$$

This now looks a lot more tractable, and is in fact exactly the same form now as the equation for the linearized planar pendulum. The only difference is the normalization required to switch to less messy dimensionless variables. The main effect of allowing the trajectory to have a non-planar component is a change in the angular frequency in the θ dependent motion. That frequency will no longer be $\sqrt{|\cos\theta_0|g/l}$, but also has a P_ϕ and other more complex trigonometric θ dependencies. It also appears that we can probably have hyperbolic or trigonometric solutions in the neighborhood of any point, regardless of whether it is a northern hemispherical point or a southern one. In the planar pendulum the unambiguous sign of the matrix terms led to hyperbolic only above the horizon, and trigonometric only below.

9.5.9 *Double and multiple pendulums, and general quadratic velocity dependence.*

In the following section I started off with the goal of treating two connected pendulums moving in a plane. Even setting up the Hamiltonian's for this turned out to be a bit messy, requiring a matrix inversion. Tackling the problem in the guise of using a more general quadratic form (which works for the two particle as well as N particle cases) seemed like it would actually be simpler than using the specifics from the angular velocity dependence of the specific pendulum problem. Once the Hamiltonian equations were found in this form, an attempt to do the first order

Taylor expansion as done for the single planar pendulum and the spherical pendulum was performed. This turned out to be a nasty mess and is seen to not be particularly illuminating. I did not know that is how it would turn out ahead of time since I had my fingers crossed for some sort of magic simplification once the final substitution were made. If such a simplification is possible, the procedure to do so is not obvious.

Although the Hamiltonian equations for a spherical pendulum have been considered previously, for the double pendulum case it seems prudent to avoid temptation, and to first see what happens with a simpler first step, a planar double pendulum.

Setting up coordinates x axis down, and y axis to the left with $i = \hat{\mathbf{x}}\hat{\mathbf{y}}$ we have for the position of the first mass m_1, at angle θ_1 and length l_1

$$z_1 = \hat{\mathbf{x}} l_1 e^{i\theta_1}. \tag{9.143}$$

If the second mass, dangling from this is at an angle θ_2 from the x axis, its position is

$$z_2 = z_1 + \hat{\mathbf{x}} l_2 e^{i\theta_2}. \tag{9.144}$$

We need the velocities, and their magnitudes. For z_1 this is

$$|\dot{z}_1|^2 = {l_1}^2 \dot{\theta}_1^2. \tag{9.145}$$

For the second mass

$$\dot{z}_2 = \hat{\mathbf{x}} i \left(l_1 \dot{\theta}_1 e^{i\theta_1} + l_2 \dot{\theta}_2 e^{i\theta_2} \right). \tag{9.146}$$

Taking conjugates and multiplying out we have

$$|\dot{z}_2|^2 = {l_1}^2 \dot{\theta}_1^2 + 2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2) + {l_2}^2 \dot{\theta}_2^2. \tag{9.147}$$

That is all that we need for the Kinetic terms in the Lagrangian. Now we need the height for the mgh terms. If we set the reference point at the lowest point for the double pendulum system, the height of the first particle is

$$h_1 = l_2 + l_1(1 - \cos\theta_1). \tag{9.148}$$

For the second particle, the distance from the horizontal is

$$d = l_1 \cos\theta_1 + l_2 \cos\theta_2. \tag{9.149}$$

So the total distance from the reference point is

$$h_2 = l_1(1 - \cos\theta_1) + l_2(1 - \cos\theta_2). \tag{9.150}$$

We now have the Lagrangian

$$\begin{aligned} L' = {} & \frac{1}{2}m_1 {l_1}^2\dot{\theta}_1^2 + \frac{1}{2}m_2\left({l_1}^2\dot{\theta}_1^2 + 2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2) + {l_2}^2\dot{\theta}_2^2\right) \\ & - m_1g(l_2 + l_1(1 - \cos\theta_1)) - m_2g(l_1(1 - \cos\theta_1) + l_2(1 - \cos\theta_2)). \end{aligned} \tag{9.151}$$

Dropping constant terms (effectively choosing a difference reference point for the potential) and rearranging a bit, also writing $M = m_1 + m_2$, we have the simpler Lagrangian

$$\begin{aligned} L = {} & \frac{1}{2}M{l_1}^2\dot{\theta}_1^2 + \frac{1}{2}m_2{l_2}^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2) \\ & + Ml_1g\cos\theta_1 + m_2l_2g\cos\theta_2. \end{aligned} \tag{9.152}$$

The conjugate momenta that we need for the Hamiltonian are

$$\begin{aligned} P_{\theta_1} &= M{l_1}^2\dot{\theta}_1 + m_2l_1l_2\dot{\theta}_2\cos(\theta_1 - \theta_2) \\ P_{\theta_2} &= m_2{l_2}^2\dot{\theta}_2 + m_2l_1l_2\dot{\theta}_1\cos(\theta_1 - \theta_2). \end{aligned} \tag{9.153}$$

Unlike any of the other simpler Hamiltonian systems considered so far, the coupling between the velocities means that we have a system of equations that we must first invert before we can even express the Hamiltonian in terms of the respective momenta.

That is

$$\begin{bmatrix} P_{\theta_1} \\ P_{\theta_2} \end{bmatrix} = \begin{bmatrix} M{l_1}^2 & m_2l_1l_2\cos(\theta_1 - \theta_2) \\ m_2l_1l_2\cos(\theta_1 - \theta_2) & m_2{l_2}^2 \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix}. \tag{9.154}$$

While this is easily invertible, doing so and attempting to substitute it back, results in an unholy mess (albeit perhaps one that can be simplified). Is there a better way? A possibly promising way is motivated by observing that this matrix, a function of the angular difference $\delta = \theta_1 - \theta_2$, looks like it is something like a moment of inertia tensor. If we call this $\mathcal{I}$, and write

$$\boldsymbol{\Theta} \equiv \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}. \tag{9.155}$$

Then the relation between the conjugate momenta in vector form

$$\mathbf{p} \equiv \begin{bmatrix} P_{\theta_1} \\ P_{\theta_2} \end{bmatrix}, \tag{9.156}$$

and the angular velocity vector can be written

$$\mathbf{p} = \mathcal{I}(\delta)\dot{\mathbf{\Theta}}. \tag{9.157}$$

Can we write the Lagrangian in terms of $\dot{\mathbf{\Theta}}$? The first Kinetic term is easy, just

$$\frac{1}{2}m_1 l^2 \dot{\theta}_1^2 = \frac{1}{2}m_1 \dot{\mathbf{\Theta}}^{\mathrm{T}} \begin{bmatrix} l_1^2 & 0 \\ 0 & 0 \end{bmatrix} \dot{\mathbf{\Theta}}. \tag{9.158}$$

For the second mass, going back to eq. (9.146), we can write

$$\dot{z}_2 = \hat{\mathbf{x}} i \left[l_1 e^{i\theta_1} l_2 e^{i\theta_2} \right] \dot{\mathbf{\Theta}}. \tag{9.159}$$

Writing $\mathbf{r}$ for this $1x2$ matrix, we can utilize the associative property for compatible sized matrices to rewrite the speed for the second particle in terms of a quadratic form

$$|\dot{z}_2|^2 = \left(\mathbf{r}\dot{\mathbf{\Theta}}\right)\left(\bar{\mathbf{r}}\dot{\mathbf{\Theta}}\right) = \dot{\mathbf{\Theta}}^{\mathrm{T}}\left(\mathbf{r}^{\mathrm{T}}\bar{\mathbf{r}}\right)\dot{\mathbf{\Theta}}. \tag{9.160}$$

The Lagrangian kinetic can all now be grouped into a single quadratic form

$$Q \equiv m_1 \begin{bmatrix} l_1 \\ 0 \end{bmatrix} \begin{bmatrix} l_1 & 0 \end{bmatrix} + m_2 \begin{bmatrix} l_1 e^{i\theta_1} \\ l_2 e^{i\theta_2} \end{bmatrix} \begin{bmatrix} l_1 e^{-i\theta_1} & l_2 e^{-i\theta_2} \end{bmatrix}. \tag{9.161}$$

$$L = \frac{1}{2}\dot{\mathbf{\Theta}}^{\mathrm{T}} Q \dot{\mathbf{\Theta}} + M l_1 g \cos\theta_1 + m_2 l_2 g \cos\theta_2. \tag{9.162}$$

It is also clear that this generalize easily to multiple connected pendulums, as follows

$$\begin{aligned} K &= \frac{1}{2}\dot{\mathbf{\Theta}}^{\mathrm{T}} \sum_k m_k Q_k \dot{\mathbf{\Theta}} \\ Q_k &= \left[l_r l_s e^{i(\theta_r - \theta_s)} \right]_{r,s \le k} \\ \phi &= -g \sum_k l_k \cos\theta_k \sum_{j=k}^{N} m_j \\ L &= K - \phi. \end{aligned} \tag{9.163}$$

In the expression for Q_k above, it is implied that the matrix is zero for any indices $r, s > k$, so it would perhaps be better to write explicitly

$$Q = \sum_k m_k Q_k = \left[\sum_{j=\max(r,s)}^{N} m_j l_r l_s e^{i(\theta_r - \theta_s)}\right]_{r,s}. \tag{9.164}$$

Returning to the problem, it is convenient and sufficient in many cases to only discuss the representative double pendulum case. For that we can calculate the conjugate momenta from eq. (9.162) directly

$$\begin{aligned} P_{\theta_1} &= \frac{\partial}{\partial \dot{\theta}_1} \frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} Q \dot{\boldsymbol{\Theta}} \\ &= \frac{\partial}{\partial \dot{\theta}_1} \frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} Q \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} Q \dot{\boldsymbol{\Theta}} \\ &= \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\frac{1}{2} (Q + Q^{\mathrm{T}}) \right) \dot{\boldsymbol{\Theta}}. \end{aligned} \tag{9.165}$$

Similarly the θ_2 conjugate momentum is

$$P_{\theta_2} = \begin{bmatrix} 0 & 1 \end{bmatrix} \left(\frac{1}{2} (Q + Q^{\mathrm{T}}) \right) \dot{\boldsymbol{\Theta}}. \tag{9.166}$$

Putting both together, it is straightforward to verify that this recovers eq. (9.154), which can now be written

$$\mathbf{p} = \frac{1}{2} (Q + Q^{\mathrm{T}}) \dot{\boldsymbol{\Theta}} = \mathcal{I} \dot{\boldsymbol{\Theta}}. \tag{9.167}$$

Observing that $\mathcal{I} = \mathcal{I}^{\mathrm{T}}$, and thus $(\mathcal{I}^{\mathrm{T}})^{-1} = \mathcal{I}^{-1}$, we now have everything required to express the Hamiltonian in terms of the conjugate momenta

$$H = \mathbf{p}^{\mathrm{T}} \left(\frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \right) \mathbf{p} - M g l_1 \cos\theta_1 - m_2 l_2 g \cos\theta_2. \tag{9.168}$$

This is now in a convenient form to calculate the first set of Hamiltonian equations.

$$
\begin{aligned}
\dot{\theta}_k &= \frac{\partial H}{\partial P_{\theta_k}} \\
&= \frac{\partial \mathbf{p}^{\mathrm{T}}}{\partial P_{\theta_k}} \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \mathbf{p} + \mathbf{p}^{\mathrm{T}} \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \frac{\partial \mathbf{p}^{\mathrm{T}}}{\partial P_{\theta_k}} \\
&= \left[\delta_{kj}\right]_j \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \mathbf{p} + \mathbf{p}^{\mathrm{T}} \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \left[\delta_{ik}\right]_i \\
&= \left[\delta_{kj}\right]_j \mathcal{I}^{-1} \overbrace{\boxed{\frac{1}{2}(Q + Q^{\mathrm{T}})}}^{\mathcal{I}} \mathcal{I}^{-1} \mathbf{p} \\
&= \left[\delta_{kj}\right]_j \mathcal{I}^{-1} \mathbf{p}.
\end{aligned}
\tag{9.169}
$$

So, when the velocity dependence is a quadratic form as identified in eq. (9.161), the first half of the Hamiltonian equations in vector form are just

$$
\dot{\mathbf{\Theta}} = \begin{bmatrix} \frac{\partial}{\partial P_{\theta_1}} & \cdots & \frac{\partial}{\partial P_{\theta_N}} \end{bmatrix}^{\mathrm{T}} H = \mathcal{I}^{-1} \mathbf{p}. \tag{9.170}
$$

This is exactly the relation we used in the first place to re-express the Lagrangian in terms of the conjugate momenta in preparation for this calculation. The remaining Hamiltonian equations are trickier, and what we now want to calculate. Without specific reference to the pendulum problem, lets do this calculation for the general Hamiltonian for a non-velocity dependent potential. That is

$$
H = \mathbf{p}^{\mathrm{T}} \left(\frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \right) \mathbf{p} + \phi(\mathbf{\Theta}). \tag{9.171}
$$

The remaining Hamiltonian equations are $\partial H/\partial\theta_a = -\dot{P}_{\theta_a}$, and the tricky part of evaluating this is going to all reside in the Kinetic term. Diving right in this is

$$\begin{aligned}
\frac{\partial K}{\partial\theta_a} &= \mathbf{p}^{\mathrm{T}}\left(\frac{1}{2}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_a}Q\mathcal{I}^{-1}\right)\mathbf{p} + \mathbf{p}^{\mathrm{T}}\left(\frac{1}{2}\mathcal{I}^{-1}\frac{\partial Q}{\partial\theta_a}\mathcal{I}^{-1}\right)\mathbf{p} \\
&\quad + \mathbf{p}^{\mathrm{T}}\left(\frac{1}{2}\mathcal{I}^{-1}Q\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_a}\right)\mathbf{p} \\
&= \mathcal{I} \\
&= \mathbf{p}^{\mathrm{T}}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_a}\boxed{\frac{1}{2}(Q+Q^{\mathrm{T}})}\mathcal{I}^{-1}\mathbf{p} + \mathbf{p}^{\mathrm{T}}\left(\frac{1}{2}\mathcal{I}^{-1}\frac{\partial Q}{\partial\theta_a}\mathcal{I}^{-1}\right)\mathbf{p} \\
&= \mathbf{p}^{\mathrm{T}}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_a}\mathbf{p} + \mathbf{p}^{\mathrm{T}}\left(\frac{1}{2}\mathcal{I}^{-1}\frac{\partial Q}{\partial\theta_a}\mathcal{I}^{-1}\right)\mathbf{p}.
\end{aligned} \tag{9.172}$$

For the two particle case we can expand this inverse easily enough, and then take derivatives to evaluate this, but this is messier and intractable for the general case. We can however, calculate the derivative of the identity matrix using the standard trick from rigid body mechanics

$$\begin{aligned}
0 &= \frac{\partial I}{\partial\theta_a} \\
&= \frac{\partial(\mathcal{I}\mathcal{I}^{-1})}{\partial\theta_a} \\
&= \frac{\partial\mathcal{I}}{\partial\theta_a}\mathcal{I}^{-1} + \mathcal{I}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_a}.
\end{aligned} \tag{9.173}$$

Thus the derivative of the inverse (moment of inertia?) matrix is

$$\begin{aligned}
\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_a} &= -\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_a}\mathcal{I}^{-1} \\
&= -\mathcal{I}^{-1}\frac{1}{2}\left(\frac{\partial Q}{\partial\theta_a} + \frac{\partial Q^{\mathrm{T}}}{\partial\theta_a}\right)\mathcal{I}^{-1}.
\end{aligned} \tag{9.174}$$

This gives us for the Hamiltonian equation

$$\frac{\partial H}{\partial\theta_a} = -\frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_a}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p} + \frac{\partial\phi}{\partial\theta_a}. \tag{9.175}$$

If we introduce a phase space position gradients

$$\nabla \equiv \begin{bmatrix} \frac{\partial}{\partial\theta_1} & \cdots & \frac{\partial}{\partial\theta_N} \end{bmatrix}^{\mathrm{T}}. \tag{9.176}$$

Then for the second half of the Hamiltonian equations we have the vector form

$$-\nabla H = \dot{\mathbf{p}} = \left[\tfrac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right]_r - \nabla\phi. \tag{9.177}$$

The complete set of Hamiltonian equations for eq. (9.171), in block matrix form, describing all the phase space change of the system is therefore

$$\frac{d}{dt}\begin{bmatrix}\mathbf{p}\\ \mathbf{\Theta}\end{bmatrix} = \begin{bmatrix}\left[\tfrac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right]_r - \nabla\phi\\ \mathcal{I}^{-1}\mathbf{p}\end{bmatrix} = \begin{bmatrix}\left[\tfrac{1}{2}\dot{\mathbf{\Theta}}\left(\frac{\partial Q}{\partial \theta_r}\right)^{\mathrm{T}}\dot{\mathbf{\Theta}}\right]_r - \nabla\phi\\ \dot{\mathbf{\Theta}}\end{bmatrix}. \tag{9.178}$$

This is a very general relation, much more so than required for the original two particle problem. We have the same non-linearity that prevents this from being easily solved. If we want a linear expansion around a phase space point to find an approximate first order solution, we can get that applying the chain rule, calculating all the $\partial/\partial\theta_k$, and $\partial/\partial P_{\theta_k}$ derivatives of the top N rows of this matrix.

If we write

$$\mathbf{z} \equiv \begin{bmatrix}\mathbf{p}\\ \mathbf{\Theta}\end{bmatrix} - \begin{bmatrix}\mathbf{p}\\ \mathbf{\Theta}\end{bmatrix}_{t=0}. \tag{9.179}$$

and the Hamiltonian equations eq. (9.178) as

$$\frac{d}{dt}\left[\mathbf{p}\mathbf{\Theta}\right] = A(\mathbf{p}, \mathbf{\Theta}). \tag{9.180}$$

Then the linearization, without simplifying or making explicit yet is

$$\dot{\mathbf{z}} \approx \begin{bmatrix}\left[\tfrac{1}{2}\dot{\mathbf{\Theta}}\left(\frac{\partial Q}{\partial \theta_r}\right)^{\mathrm{T}}\dot{\mathbf{\Theta}}\right]_r - \nabla\phi\\ \dot{\mathbf{\Theta}}\end{bmatrix}_{t=0} + \left[\frac{\partial A}{\partial P_{\theta_1}} \quad \cdots \quad \frac{\partial A}{\partial P_{\theta_N}} \quad \frac{\partial A}{\partial \theta_1} \quad \cdots \quad \frac{\partial A}{\partial \theta_N}\right]\Bigg|_{t=0}\mathbf{z}. \tag{9.181}$$

For brevity the constant term evaluated at $t = 0$ is expressed in terms of the original angular velocity vector from our Lagrangian. The task is now evaluating the derivatives in the first order term of this Taylor series. Let us do these one at a time and then reassemble all the results afterward.

So that we can discuss just the first order terms lets write Δ for the matrix of first order derivatives in our Taylor expansion, as in

$$f(\mathbf{p}, \mathbf{\Theta}) = f(\mathbf{p}, \mathbf{\Theta})|_0 + \Delta f|_0\mathbf{z} + \cdots \tag{9.182}$$

First, lets do the potential gradient.

$$\Delta(\nabla\phi) = \left[0\left[\frac{\partial^2\phi}{\partial\theta_r\partial\theta_c}\right]_{r,c}\right]. \tag{9.183}$$

Next in terms of complexity is the first order term of $\dot{\mathbf{\Theta}}$, for which we have

$$\Delta(\mathcal{I}^{-1}\mathbf{p}) = \left[\left[\mathcal{I}^{-1}\left[\delta_{rc}\right]_r\right]_c \quad \left[\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\mathbf{p}\right]_c\right]. \tag{9.184}$$

The δ over all rows r and columns c is the identity matrix and we are left with

$$\Delta(\mathcal{I}^{-1}\mathbf{p}) = \left[\mathcal{I}^{-1} \quad \left[\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\mathbf{p}\right]_c\right]. \tag{9.185}$$

Next, consider just the P_θ dependence in the elements of the row vector

$$\left[\frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right]_r. \tag{9.186}$$

We can take derivatives of this, and exploiting the fact that these elements are scalars, so they equal their transpose. Also noting that ${A^{-1}}^{\mathrm{T}} = {A^{\mathrm{T}}}^{-1}$, and $\mathcal{I} = \mathcal{I}^{\mathrm{T}}$, we have

$$\begin{aligned}
&\frac{\partial}{\partial P_{\theta_c}}\left(\frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right)\\
&\quad= \frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\left[\delta_{rc}\right]_r + \frac{1}{2}\left(\left[\delta_{rc}\right]_r\right)^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\\
&\quad= \mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial}{\partial\theta_r}\frac{1}{2}\left(Q+Q^T\right)\right)\mathcal{I}^{-1}\left[\delta_{rc}\right]_r\\
&\quad= \mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_r}\mathcal{I}^{-1}\left[\delta_{rc}\right]_r.
\end{aligned} \tag{9.187}$$

Since we also have $B'B^{-1} + B(B^{-1})' = 0$, for invertible matrixes B, this reduces to

$$\frac{\partial}{\partial P_{\theta_c}}\left(\frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right) = -\mathbf{p}^{\mathrm{T}}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_r}\left[\delta_{rc}\right]_r. \tag{9.188}$$

Forming the matrix over all rows r, and columns c, we get a trailing identity multiplying from the right, and are left with

$$\left[\frac{\partial}{\partial P_{\theta_c}}\left(\frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right)\right]_{r,c} = \left[-\mathbf{p}^{\mathrm{T}}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_r}\right]_r = \left[-\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\mathbf{p}\right]_c. \tag{9.189}$$

Okay, getting closer. The only thing left is to consider the remaining θ dependence of eq. (9.186), and now want the theta partials of the scalar matrix elements

$$\begin{aligned}
&\frac{\partial}{\partial\theta_c}\left(\frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right)\\
&\quad=\mathbf{p}^{\mathrm{T}}\left(\frac{\partial}{\partial\theta_c}\left(\frac{1}{2}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\right)\right)\mathbf{p}\\
&\quad=\mathbf{p}^{\mathrm{T}}\frac{1}{2}\mathcal{I}^{-1}\frac{\partial^2 Q^{\mathrm{T}}}{\partial\theta_c\partial\theta_r}\mathcal{I}^{-1}\mathbf{p}\\
&\qquad+\mathbf{p}^{\mathrm{T}}\frac{1}{2}\left(\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}+\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\right)\mathbf{p}\\
&\quad=\mathbf{p}^{\mathrm{T}}\frac{1}{2}\mathcal{I}^{-1}\frac{\partial^2 Q^{\mathrm{T}}}{\partial\theta_c\partial\theta_r}\mathcal{I}^{-1}\mathbf{p}+\mathbf{p}^{\mathrm{T}}\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\frac{\partial\mathcal{I}}{\partial\theta_r}\mathcal{I}^{-1}\mathbf{p}.
\end{aligned}\tag{9.190}$$

There is a slight asymmetry between the first and last terms here that can possibly be eliminated. Using $B^{-1\prime}=-B^{-1}B'B^{-1}$, we can factor out the $\mathcal{I}^{-1}\mathbf{p}=\dot{\mathbf{\Theta}}$ terms

$$\frac{\partial}{\partial\theta_c}\left(\frac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right)=\dot{\mathbf{\Theta}}^{\mathrm{T}}\left(\frac{1}{2}\frac{\partial^2 Q^{\mathrm{T}}}{\partial\theta_c\partial\theta_r}-\frac{\partial\mathcal{I}}{\partial\theta_c}\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_r}\right)\dot{\mathbf{\Theta}}.\tag{9.191}$$

Is this any better? Maybe a bit. Since we are forming the matrix over all r,c indices and can assume mixed partial commutation the transpose can be dropped leaving us with

$$\left[\frac{\partial}{\partial\theta_c}\left(\tfrac{1}{2}\mathbf{p}^{\mathrm{T}}\mathcal{I}^{-1}\left(\tfrac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\mathcal{I}^{-1}\mathbf{p}\right)\right]_{r,c}=\left[\dot{\mathbf{\Theta}}^{\mathrm{T}}\left(\tfrac{1}{2}\tfrac{\partial^2 Q}{\partial\theta_c\partial\theta_r}-\tfrac{\partial\mathcal{I}}{\partial\theta_c}\mathcal{I}^{-1}\tfrac{\partial\mathcal{I}}{\partial\theta_r}\right)\dot{\mathbf{\Theta}}\right]_{r,c}.\tag{9.192}$$

We can now assemble all these individual derivatives

$$\begin{aligned}
\dot{\mathbf{z}}\approx&\begin{bmatrix}\left[\tfrac{1}{2}\dot{\mathbf{\Theta}}\left(\tfrac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\dot{\mathbf{\Theta}}\right]_r-\nabla\phi\\ \dot{\mathbf{\Theta}}\end{bmatrix}_{t=0}\\
&+\begin{bmatrix}-\left[\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\mathbf{p}\right]_c & \left[\dot{\mathbf{\Theta}}^{\mathrm{T}}\left(\frac{1}{2}\frac{\partial^2 Q}{\partial\theta_c\partial\theta_r}-\frac{\partial\mathcal{I}}{\partial\theta_c}\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_r}\right)\dot{\mathbf{\Theta}}-\frac{\partial^2\phi}{\partial\theta_r\partial\theta_c}\right]_{r,c}\\ \mathcal{I}^{-1} & \left[\frac{\partial(\mathcal{I}^{-1})}{\partial\theta_c}\mathbf{p}\right]_c\end{bmatrix}_{t=0}\mathbf{z}.
\end{aligned}\tag{9.193}$$

We have both $\partial(\mathcal{I}^{-1})/\partial\theta_k$ and $\partial\mathcal{I}/\partial\theta_k$ derivatives above, which will complicate things when trying to evaluate this for any specific system. A final elimination of the derivatives of the inverse inertial matrix leaves us with

$$\begin{aligned}\dot{\mathbf{z}} \approx & \begin{bmatrix} \left[\frac{1}{2}\dot{\mathbf{\Theta}}\left(\frac{\partial Q}{\partial\theta_r}\right)^{\mathrm{T}}\dot{\mathbf{\Theta}}\right]_r - \nabla\phi \\ \dot{\mathbf{\Theta}} \end{bmatrix}_{t=0} \\ & + \begin{bmatrix} \left[\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_c}\dot{\mathbf{\Theta}}\right]_c & \left[\dot{\mathbf{\Theta}}^{\mathrm{T}}\left(\frac{1}{2}\frac{\partial^2 Q}{\partial\theta_c\partial\theta_r} - \frac{\partial\mathcal{I}}{\partial\theta_c}\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_r}\right)\dot{\mathbf{\Theta}} - \frac{\partial^2\phi}{\partial\theta_r\partial\theta_c}\right]_{r,c} \\ \mathcal{I}^{-1} & -\left[\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_c}\dot{\mathbf{\Theta}}\right]_c \end{bmatrix}_{t=0} \mathbf{z}.\end{aligned} \tag{9.194}$$

9.5.9.1 *Single pendulum verification.*

Having accumulated this unholy mess of abstraction, lets verify this first against the previous result obtained for the single planar pendulum. Then if that checks out, calculate these matrices explicitly for the double and multiple pendulum cases. For the single mass pendulum we have

$$\begin{aligned} Q &= \mathcal{I} = ml^2 \\ \phi &= -mgl\cos\theta. \end{aligned} \tag{9.195}$$

So all the θ partials except that of the potential are zero. For the potential we have

$$-\left.\frac{\partial^2\phi}{\partial^2\theta}\right|_0 = -mgl\cos\theta_0. \tag{9.196}$$

and for the angular gradient

$$-\nabla\phi|_0 = \left[-mgl\sin\theta_0\right]. \tag{9.197}$$

Putting these all together in this simplest application of eq. (9.194) we have for the linear approximation of a single point mass pendulum about some point in phase space at time zero:

$$\dot{\mathbf{z}} \approx \begin{bmatrix} -mgl\sin\theta_0 \\ \dot{\theta}_0 \end{bmatrix} + \begin{bmatrix} 0 & -mgl\cos\theta_0 \\ \frac{1}{ml^2} & 0 \end{bmatrix}\mathbf{z}. \tag{9.198}$$

Excellent. Have not gotten into too much trouble with the math so far. This is consistent with the previous results obtained considering the simple pendulum directly (it actually pointed out an error in the earlier pendulum treatment which is now fixed (I had dropped the $\dot{\theta}_0$ term)).

9.5.9.2 *Double pendulum explicitly.*

For the double pendulum, with $\delta = \theta_1 - \theta_2$, and $M = m_1 + m_2$, we have

$$Q = \begin{bmatrix} M{l_1}^2 & m_2 l_2 l_1 e^{i(\theta_2-\theta_1)} \\ m_2 l_1 l_2 e^{i(\theta_1-\theta_2)} & m_2 {l_2}^2 \end{bmatrix} = \begin{bmatrix} M{l_1}^2 & m_2 l_2 l_1 e^{-i\delta} \\ m_2 l_1 l_2 e^{i\delta} & m_2 {l_2}^2 \end{bmatrix}. \tag{9.199}$$

$$\begin{aligned} \frac{1}{2}\dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left(\frac{\partial Q}{\partial \theta_1}\right)^{\mathrm{T}}\dot{\boldsymbol{\Theta}} &= \frac{1}{2} m_2 l_1 l_2 i \dot{\boldsymbol{\Theta}}^{\mathrm{T}} \begin{bmatrix} 0 & -e^{-i\delta} \\ e^{i\delta} & 0 \end{bmatrix}^{\mathrm{T}} \dot{\boldsymbol{\Theta}} \\ &= \frac{1}{2} m_2 l_1 l_2 i \dot{\boldsymbol{\Theta}}^{\mathrm{T}} \begin{bmatrix} e^{i\delta}\dot{\theta}_2 \\ -e^{-i\delta}\dot{\theta}_1 \end{bmatrix} \\ &= \frac{1}{2} m_2 l_1 l_2 i \dot{\theta}_1 \dot{\theta}_2 (e^{i\delta} - e^{-i\delta}) \\ &= -m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin\delta. \end{aligned} \tag{9.200}$$

The θ_2 derivative is the same but inverted in sign, so we have most of the constant term calculated. We need the potential gradient to complete. Our potential was

$$\phi = -M l_1 g \cos\theta_1 - m_2 l_2 g \cos\theta_2. \tag{9.201}$$

So, the gradient is

$$\nabla\phi = \begin{bmatrix} M l_1 g \sin\theta_1 m_2 l_2 g \sin\theta_2 \end{bmatrix}. \tag{9.202}$$

Putting things back together we have for the linear approximation of the two pendulum system

$$\dot{\mathbf{z}} = \begin{bmatrix} m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) \begin{bmatrix} -1 \\ 1 \end{bmatrix} & -g \begin{bmatrix} M l_1 \sin\theta_1 \\ m_2 l_2 \sin\theta_2 \end{bmatrix} \\ \dot{\theta}_1 & \\ \dot{\theta}_2 & \end{bmatrix}_{t=0} + A\mathbf{z}. \tag{9.203}$$

Where A is still to be determined (from eq. (9.194)). One of the elements of A are the matrix of potential derivatives. These are

$$\begin{bmatrix} \frac{\partial \nabla\phi}{\partial\theta_1} & \frac{\partial \nabla\phi}{\partial\theta_2} \end{bmatrix} = \begin{bmatrix} M l_1 g \cos\theta_1 & 0 \\ 0 & m_2 l_2 g \cos\theta_2 \end{bmatrix}. \tag{9.204}$$

We also need the inertial matrix and its inverse. These are

$$\mathcal{I} = \begin{bmatrix} M{l_1}^2 & m_2 l_2 l_1 \cos\delta \\ m_2 l_1 l_2 \cos\delta & m_2 {l_2}^2 \end{bmatrix}. \tag{9.205}$$

$$\mathcal{I}^{-1} = \frac{1}{{l_1}^2 {l_2}^2 m_2 (M - m_2 \cos^2\delta)} \begin{bmatrix} m_2 {l_2}^2 & -m_2 l_2 l_1 \cos\delta \\ -m_2 l_1 l_2 \cos\delta & M{l_1}^2 \end{bmatrix}. \tag{9.206}$$

Since

$$\frac{\partial Q}{\partial \theta_1} = m_2 l_1 l_2 i \begin{bmatrix} 0 & -e^{-i\delta} \\ e^{i\delta} & 0 \end{bmatrix}. \tag{9.207}$$

We have

$$\begin{aligned} \frac{\partial}{\partial \theta_1}\frac{\partial Q}{\partial \theta_1} &= -m_2 l_1 l_2 \begin{bmatrix} 0 & e^{-i\delta} \\ e^{i\delta} & 0 \end{bmatrix} \\ \frac{\partial}{\partial \theta_2}\frac{\partial Q}{\partial \theta_1} &= m_2 l_1 l_2 \begin{bmatrix} 0 & e^{-i\delta} \\ e^{i\delta} & 0 \end{bmatrix} \\ \frac{\partial}{\partial \theta_1}\frac{\partial Q}{\partial \theta_2} &= m_2 l_1 l_2 \begin{bmatrix} 0 & e^{-i\delta} \\ e^{i\delta} & 0 \end{bmatrix} \\ \frac{\partial}{\partial \theta_2}\frac{\partial Q}{\partial \theta_2} &= -m_2 l_1 l_2 \begin{bmatrix} 0 & e^{-i\delta} \\ e^{i\delta} & 0 \end{bmatrix}. \end{aligned} \tag{9.208}$$

and the matrix of derivatives becomes

$$\frac{1}{2}\dot{\mathbf{\Theta}}^{\mathrm{T}} \frac{\partial}{\partial \theta_c}\frac{\partial Q}{\partial \theta_r}\dot{\mathbf{\Theta}} = m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2) \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{9.209}$$

For the remaining two types of terms in the matrix A we need $\mathcal{I}^{-1}\partial\mathcal{I}/\partial\theta_k$. The derivative of the inertial matrix is

$$\frac{\partial \mathcal{I}}{\partial \theta_k} = -m_2 l_1 l_2 (\delta_{k1} - \delta_{k2}) \begin{bmatrix} 0 & \sin\delta \\ \sin\delta & 0 \end{bmatrix}. \tag{9.210}$$

Computing the product

$$\begin{aligned}\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_k} &= \frac{-m_2 l_1 l_2(\delta_{k1}-\delta_{k2})}{{l_1}^2{l_2}^2 m_2(M-m_2\cos^2\delta)}\times \\ &\quad \begin{bmatrix} m_2{l_2}^2 & -m_2 l_2 l_1\cos\delta \\ -m_2 l_1 l_2\cos\delta & M{l_1}^2\end{bmatrix}\begin{bmatrix} 0 & \sin\delta \\ \sin\delta & 0\end{bmatrix} \\ &= \frac{-m_2 l_1 l_2(\delta_{k1}-\delta_{k2})\sin\delta}{{l_1}^2{l_2}^2 m_2(M-m_2\cos^2\delta)}\begin{bmatrix} -m_2 l_2 l_1\cos\delta & m_2{l_2}^2 \\ M{l_1}^2 & -m_2 l_1 l_2\cos\delta\end{bmatrix}.\end{aligned} \tag{9.211}$$

We want the matrix of $\mathcal{I}^{-1}\partial\mathcal{I}/\partial\theta_c\dot{\mathbf{\Theta}}$ over columns c, and this is

$$\begin{aligned}\left[\mathcal{I}^{-1}\partial\mathcal{I}/\partial\theta_c\dot{\mathbf{\Theta}}\right]_c &= \frac{m_2 l_1 l_2\sin\delta}{{l_1}^2{l_2}^2 m_2(M-m_2\cos^2\delta)}\times \\ &\quad \begin{bmatrix} m_2 l_2 l_1\cos\delta\dot\theta_1 - m_2{l_2}^2\dot\theta_2 & -m_2 l_2 l_1\cos\delta\dot\theta_1 + m_2{l_2}^2\dot\theta_2 \\ -M{l_1}^2\dot\theta_1 + m_2 l_1 l_2\cos\delta\dot\theta_2 & M{l_1}^2\dot\theta_1 - m_2 l_1 l_2\cos\delta\dot\theta_2\end{bmatrix}.\end{aligned} \tag{9.212}$$

Very messy. Perhaps it would be better not even bothering to expand this explicitly? The last term in the matrix A is probably no better. For that we want

$$\begin{aligned}-\frac{\partial\mathcal{I}}{\partial\theta_c}\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_r} &= \frac{-{m_2}^2{l_1}^2{l_2}^2(\delta_{c1}-\delta_{c2})(\delta_{r1}-\delta_{r2})\sin^2\delta}{{l_1}^2{l_2}^2 m_2(M-m_2\cos^2\delta)}\begin{bmatrix} 0 & 1 \\ 1 & 0\end{bmatrix}\times \\ &\quad \begin{bmatrix} -m_2 l_2 l_1\cos\delta & m_2{l_2}^2 \\ M{l_1}^2 & -m_2 l_1 l_2\cos\delta\end{bmatrix} \\ &= \frac{-{m_2}^2{l_1}^2{l_2}^2(\delta_{c1}-\delta_{c2})(\delta_{r1}-\delta_{r2})\sin^2\delta}{{l_1}^2{l_2}^2 m_2(M-m_2\cos^2\delta)}\begin{bmatrix} M{l_1}^2 & -m_2 l_1 l_2\cos\delta \\ -m_2 l_2 l_1\cos\delta & m_2{l_2}^2\end{bmatrix}.\end{aligned} \tag{9.213}$$

With a sandwich of this between $\dot{\mathbf{\Theta}}^{\mathrm{T}}$ and $\dot{\mathbf{\Theta}}$ we are almost there

$$\begin{aligned}-\dot{\mathbf{\Theta}}^{\mathrm{T}}\frac{\partial\mathcal{I}}{\partial\theta_c}\mathcal{I}^{-1}\frac{\partial\mathcal{I}}{\partial\theta_r}\dot{\mathbf{\Theta}} &= \frac{-{m_2}^2{l_1}^2{l_2}^2(\delta_{c1}-\delta_{c2})(\delta_{r1}-\delta_{r2})\sin^2\delta}{{l_1}^2{l_2}^2 m_2(M-m_2\cos^2\delta)} \\ &\quad \left(M{l_1}^2\dot\theta_1^2 - 2m_2 l_1 l_2\cos\delta\dot\theta_1\dot\theta_2 + +m_2{l_2}^2\dot\theta_2^2\right).\end{aligned} \tag{9.214}$$

We have a matrix of these scalars over r, c, and that is

$$\begin{aligned}\left[-\dot{\mathbf{\Theta}}^{\mathrm{T}}\tfrac{\partial\mathcal{I}}{\partial\theta_c}\mathcal{I}^{-1}\tfrac{\partial\mathcal{I}}{\partial\theta_r}\dot{\mathbf{\Theta}}\right]_{rc} &= \frac{{m_2}^2{l_1}^2{l_2}^2\sin^2\delta}{{l_1}^2{l_2}^2 m_2(M-m_2\cos^2\delta)}\times \\ &\quad \left(M{l_1}^2\dot\theta_1^2 - 2m_2 l_1 l_2\cos\delta\dot\theta_1\dot\theta_2 + m_2{l_2}^2\dot\theta_2^2\right)\left[\begin{smallmatrix} -1 & 1 \\ 1 & -1\end{smallmatrix}\right].\end{aligned}$$

(9.215)

Putting all the results for the matrix A together is going to make a disgusting mess, so lets summarize in block matrix form

$$A = \begin{bmatrix} B & C \\ \mathcal{I}^{-1} & -B \end{bmatrix}_{t=0}. \tag{9.216}$$

$$\begin{aligned} B = & \frac{m_2 l_1 l_2 \sin\delta}{{l_1}^2 {l_2}^2 m_2 (M - m_2 \cos^2\delta)} \times \\ & \begin{bmatrix} m_2 l_2 l_1 \cos\delta \dot{\theta}_1 - m_2 {l_2}^2 \dot{\theta}_2 & -m_2 l_2 l_1 \cos\delta \dot{\theta}_1 + m_2 {l_2}^2 \dot{\theta}_2 \\ -M {l_1}^2 \dot{\theta}_1 + m_2 l_1 l_2 \cos\delta \dot{\theta}_2 & M {l_1}^2 \dot{\theta}_1 - m_2 l_1 l_2 \cos\delta \dot{\theta}_2 \end{bmatrix}. \end{aligned} \tag{9.217}$$

$$\begin{aligned} C = & m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos\delta \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \\ & + \frac{{m_2}^2 {l_1}^2 {l_2}^2 \sin^2\delta}{{l_1}^2 {l_2}^2 m_2 (M - m_2 \cos^2\delta)} \times \\ & \left(M {l_1}^2 \dot{\theta}_1^2 - 2 m_2 l_1 l_2 \cos\delta \dot{\theta}_1 \dot{\theta}_2 + m_2 {l_2}^2 \dot{\theta}_2^2 \right) \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \\ & + \begin{bmatrix} M l_1 g \cos\theta_1 & 0 \\ 0 & m_2 l_2 g \cos\theta_2 \end{bmatrix} \end{aligned} \tag{9.218}$$

$$\mathcal{I}^{-1} = \frac{1}{{l_1}^2 {l_2}^2 m_2 (M - m_2 \cos^2\delta)} \begin{bmatrix} m_2 {l_2}^2 & -m_2 l_2 l_1 \cos\delta \\ -m_2 l_1 l_2 \cos\delta & M {l_1}^2 \end{bmatrix}. \tag{9.219}$$

$$b = \begin{bmatrix} m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) \begin{bmatrix} -1 \\ 1 \end{bmatrix} & -g \begin{bmatrix} M l_1 \sin\theta_1 \\ m_2 l_2 \sin\theta_2 \end{bmatrix} \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix}, \tag{9.220}$$

where these are all related by the first order matrix equation

$$\frac{d\mathbf{z}}{dt} = \mathbf{b}|_{t=0} + A|_{t=0}\mathbf{z}. \tag{9.221}$$

Wow, even to just write down the equations required to get a linear approximation of the two pendulum system is horrendously messy, and this is not even trying to solve it. Numerical and or symbolic computation is really called for here. If one elected to do this numerically, which looks pretty much mandatory since the analytic way did not turn out to be simple even for just the two pendulum system, then one is probably better off going all the way back to eq. (9.178) and just calculating the increment for the trajectory using a very small time increment, and do this repeatedly (i.e. do a zeroth order numerical procedure instead of the first order which turns out much more complicated).

9.5.10 *Non-covariant Lorentz force.*

In [7], the Lagrangian for a charged particle is given as (12.9) as

$$L = -mc^2\sqrt{1-\mathbf{u}^2/c^2} + \frac{e}{c}\mathbf{u}\cdot\mathbf{A} - e\Phi. \tag{9.222}$$

Let us work in detail from this to the Lorentz force law and the Hamiltonian and from the Hamiltonian again to the Lorentz force law using the Hamiltonian equations. We should get the same results in each case, and have enough details in doing so to render the text a bit more comprehensible.

9.5.10.1 *Canonical momenta.*

We need the conjugate momenta for both the Euler-Lagrange evaluation and the Hamiltonian, so lets get that first. The components of this are

$$\begin{aligned}\frac{\partial L}{\partial \dot{x}_i} &= -\frac{1}{2}mc^2\gamma(-2/c^2)\dot{x}_i + \frac{e}{c}A_i \\ &= m\gamma\dot{x}_i + \frac{e}{c}A_i.\end{aligned} \tag{9.223}$$

In vector form the canonical momenta are then

$$\mathbf{P} = \gamma m\mathbf{u} + \frac{e}{c}\mathbf{A}. \tag{9.224}$$

9.5.10.2 *Euler-Lagrange evaluation.*

Completing the Euler-Lagrange equation evaluation is the calculation of

$$\frac{d\mathbf{P}}{dt} = \boldsymbol{\nabla}L. \tag{9.225}$$

On the left hand side we have

$$\frac{d\mathbf{P}}{dt} = \frac{d(\gamma m\mathbf{u})}{dt} + \frac{e}{c}\frac{d\mathbf{A}}{dt}, \tag{9.226}$$

and on the right, with implied summation over repeated indices, we have

$$\boldsymbol{\nabla} L = \frac{e}{c}\mathbf{e}_k(\mathbf{u}\cdot\partial_k\mathbf{A}) - e\boldsymbol{\nabla}\Phi. \tag{9.227}$$

Putting things together we have

$$\begin{aligned}\frac{d(\gamma m\mathbf{u})}{dt} &= -e\left(\boldsymbol{\nabla}\Phi + \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t} + \frac{1}{c}\left(\frac{\partial\mathbf{A}}{\partial x_a}\frac{\partial x_a}{\partial t} - \mathbf{e}_k(\mathbf{u}\cdot\partial_k\mathbf{A})\right)\right)\\ &= -e\left(\boldsymbol{\nabla}\Phi + \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t} + \frac{1}{c}\mathbf{e}_b u_a\left(\frac{\partial A_b}{\partial x_a} - \frac{\partial A_a}{\partial x_b}\right)\right).\end{aligned} \tag{9.228}$$

With

$$\mathbf{E} = -\boldsymbol{\nabla}\Phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}, \tag{9.229}$$

the first two terms are recognizable as the electric field. To put some structure in the remainder start by writing

$$\frac{\partial A_b}{\partial x_a} - \frac{\partial A_a}{\partial x_b} = \epsilon^{fab}(\boldsymbol{\nabla}\times\mathbf{A})_f. \tag{9.230}$$

The remaining term, with $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$ is now

$$\begin{aligned}-\frac{e}{c}\mathbf{e}_b u_a\epsilon^{gab}B_g &= \frac{e}{c}\mathbf{e}_a u_b\epsilon^{abg}B_g\\ &= \frac{e}{c}\mathbf{u}\times\mathbf{B}.\end{aligned} \tag{9.231}$$

We are left with the momentum portion of the Lorentz force law as expected

$$\frac{d(\gamma m\mathbf{u})}{dt} = e\left(\mathbf{E} + \frac{1}{c}\mathbf{u}\times\mathbf{B}\right). \tag{9.232}$$

Observe that with a small velocity Taylor expansion of the Lagrangian we obtain the approximation

$$-mc^2\sqrt{1-\mathbf{u}^2/c^2} \approx -mc^2\left(1 - \frac{1}{2}\mathbf{u}^2/c^2\right) = \frac{1}{2}m\mathbf{u}^2. \tag{9.233}$$

If that is our starting place, we can only obtain the non-relativistic approximation of the momentum change by evaluating the Euler-Lagrange equations

$$\frac{d(m\mathbf{u})}{dt} = e\left(\mathbf{E} + \frac{1}{c}\mathbf{u}\times\mathbf{B}\right). \tag{9.234}$$

That was an exercise previously attempting working the Tong Lagrangian problem set [24].

9.5.10.3 *Hamiltonian.*

Having confirmed the by old fashioned Euler-Lagrange equation evaluation that our Lagrangian provides the desired equations of motion, let us now try it using the Hamiltonian approach. First we need the Hamiltonian, which is nothing more than

$$H = \mathbf{P}\cdot\mathbf{u} - L. \tag{9.235}$$

However, in the Lagrangian and the dot product we have velocity terms that we must eliminate in favor of the canonical momenta. The Hamiltonian remains valid in either form, but to apply the Hamiltonian equations we need $H = H(\mathbf{P},\mathbf{x})$, and not $H = H(\mathbf{u},\mathbf{P},\mathbf{x})$.

$$H = \mathbf{P}\cdot\mathbf{u} + mc^2\sqrt{1-\mathbf{u}^2/c^2} - \frac{e}{c}\mathbf{u}\cdot\mathbf{A} + e\Phi. \tag{9.236}$$

Or

$$H = \mathbf{u}\cdot\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right) + mc^2\sqrt{1-\mathbf{u}^2/c^2} + e\Phi. \tag{9.237}$$

We can rearrange eq. (9.224) for $\mathbf{u}$

$$\mathbf{u} = \frac{1}{m\gamma}\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right), \tag{9.238}$$

but γ also has a $\mathbf{u}$ dependence, so this is not complete. Squaring gets us closer

$$\mathbf{u}^2 = \frac{1-\mathbf{u}^2/c^2}{m^2}\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2, \tag{9.239}$$

and a bit of final rearrangement yields

$$\mathbf{u}^2 = \frac{(c\mathbf{P} - e\mathbf{A})^2}{m^2c^2 + \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2}. \tag{9.240}$$

Writing $\mathbf{p} = \mathbf{P} - e\mathbf{A}/c$, we can rearrange and find

$$\sqrt{1 - \mathbf{u}^2/c^2} = \frac{mc}{\sqrt{m^2c^2 + \mathbf{p}^2}}. \tag{9.241}$$

Also, taking roots of eq. (9.240) we must have the directions of $\mathbf{u}$ and $\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)$ differ only by a rotation. From eq. (9.238) we also know that these are colinear, and therefore have

$$\mathbf{u} = \frac{c\mathbf{P} - e\mathbf{A}}{\sqrt{m^2c^2 + \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2}}. \tag{9.242}$$

This and eq. (9.241) can now be substituted into eq. (9.237), for

$$H = \frac{c}{m^2c^2 + \mathbf{p}^2}\left(\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2 + m^2c^2\right) + e\Phi. \tag{9.243}$$

Dividing out the common factors we finally have the Hamiltonian in a tidy form

$$H = \sqrt{(c\mathbf{P} - e\mathbf{A})^2 + m^2c^4} + e\Phi. \tag{9.244}$$

9.5.10.4 *Hamiltonian equation evaluation.*

Let us now go through the exercise of evaluating the Hamiltonian equations. We want the starting point to be just the energy expression eq. (9.244), and the use of the Hamiltonian equations and none of what led up to that. If we were given only this Hamiltonian and the Hamiltonian principle

$$\begin{aligned} \frac{\partial H}{\partial P_k} &= u_k \\ \frac{\partial H}{\partial x_k} &= -\dot{P}_k, \end{aligned} \tag{9.245a}$$

How far can we go?

For the particle velocity we have no Φ dependence and get

$$u_k = \frac{c(cP_k - eA_k)}{\sqrt{(c\mathbf{P} - e\mathbf{A})^2 + m^2c^4}}. \tag{9.246}$$

This is eq. (9.242) in coordinate form, one of our stepping stones on the way to the Hamiltonian, and we recover it quickly with our first set of

derivatives. We have the gradient part $\dot{\mathbf{P}} = -\boldsymbol{\nabla} H$ of the Hamiltonian left to evaluate

$$\frac{d\mathbf{P}}{dt} = \frac{e(cP_k - eA_k)\boldsymbol{\nabla} A_k}{\sqrt{(c\mathbf{P} - e\mathbf{A})^2 + m^2c^4}} - e\boldsymbol{\nabla}\Phi. \tag{9.247}$$

Or

$$\frac{d\mathbf{P}}{dt} = e\left(\frac{u_k}{c}\boldsymbol{\nabla} A_k - \boldsymbol{\nabla}\Phi\right). \tag{9.248}$$

This looks nothing like the Lorentz force law. Knowing that $\mathbf{P} - e\mathbf{A}/c$ is of significance (because we know where we started which is kind of a cheat), we can subtract derivatives of this from both sides, and use the convective derivative operator $d/dt = \partial/\partial t + \mathbf{u}\cdot\boldsymbol{\nabla}$ (ie. chain rule) yielding

$$\frac{d}{dt}(\mathbf{P} - e\mathbf{A}/c) = e\left(-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \frac{1}{c}(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{A} + \frac{u_k}{c}\boldsymbol{\nabla} A_k - \boldsymbol{\nabla}\Phi\right). \tag{9.249}$$

The first and last terms sum to the electric field, and we seen evaluating the Euler-Lagrange equations that the remainder is $u_k\boldsymbol{\nabla} A_k - (\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{A} = \mathbf{u}\times(\boldsymbol{\nabla}\times\mathbf{A})$. We have therefore gotten close to the familiar Lorentz force law, and have

$$\frac{d}{dt}(\mathbf{P} - e\mathbf{A}/c) = e\left(\mathbf{E} + \frac{\mathbf{u}}{c}\times\mathbf{B}\right). \tag{9.250}$$

The only untidy detail left is that $\mathbf{P} - e\mathbf{A}/c$ does not look much like $\gamma m\mathbf{u}$, what we recognize as the relativistically corrected momentum. We ought to have that implied somewhere and eq. (9.246) looks like the place. That turns out to be the case, and some rearrangement gives us this directly

$$\mathbf{P} - \frac{e}{c}\mathbf{A} = \frac{m\mathbf{u}}{\sqrt{1 - \mathbf{u}^2/c^2}}. \tag{9.251}$$

This completes the exercise, and we have now obtained the momentum part of the Lorentz force law. This is still unsatisfactory from a relativistic context since we do not have momentum and energy on equal footing. We likely need to utilize a covariant Lagrangian and Hamiltonian formulation to fix up that deficiency.

9.6 SOLUTIONS.

10 ROUTHIAN PROCEDURE.

10.1 MOTIVATION.

Attempting study of [4] section 7-2 on Routh's procedure has been giving me some trouble. It was not "sinking in", indicating a fundamental misunderstanding, or at least a requirement to work some examples. Here I pick a system, the spherical pendulum, which has the required ignorable coordinate, to illustrate the ideas for myself with something less abstract.

We see that a first attempt to work such an example leads to the wrong result and the reasons for this are explored.

10.2 SPHERICAL PENDULUM EXAMPLE.

The Lagrangian for the pendulum is

$$L = \frac{1}{2}mr^2\left(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta\right) - mgr(1 + \cos\theta), \tag{10.1}$$

and our conjugate momenta are therefore

$$\begin{aligned} p_\theta &= \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \\ p_\phi &= \frac{\partial L}{\partial \dot{\phi}} = mr^2 \sin^2\theta\dot{\phi}. \end{aligned} \tag{10.2}$$

That is enough to now formulate the Hamiltonian $H = \dot{\theta}p_\theta + \dot{\phi}p_\phi - L$, which is

$$H = H(\theta, p_\theta, p_\phi) = \frac{1}{2mr^2}(p_\theta)^2 + \frac{1}{2mr^2\sin^2\theta}(p_\phi)^2 + mgr(1 + \cos\theta). \tag{10.3}$$

We have got the ignorable coordinate ϕ here, since the Hamiltonian has no explicit dependence on it. In the Hamiltonian formalism the constant

of motion associated with this comes as a consequence of evaluating the Hamiltonian equations. For this system, those are

$$\begin{aligned}
\frac{\partial H}{\partial \theta} &= -\dot{p}_\theta \\
\frac{\partial H}{\partial \phi} &= -\dot{p}_\phi \\
\frac{\partial H}{\partial p_\theta} &= \dot{\theta} \\
\frac{\partial H}{\partial p_\phi} &= \dot{\phi},
\end{aligned} \tag{10.4}$$

These partials are

$$\begin{aligned}
-\dot{p}_\theta &= -mgr\sin\theta - \frac{\cos\theta}{2mr^2\sin^3\theta}(p_\phi)^2 \\
-\dot{p}_\phi &= 0 \\
\dot{\theta} &= \frac{1}{mr^2}p_\theta \\
\dot{\phi} &= \frac{1}{mr^2\sin^2\theta}p_\phi.
\end{aligned} \tag{10.5}$$

The second of these provides the integration constant, allowing us to write, $p_\phi = \alpha$. Once this is done, our Hamiltonian example is reduced to one complete set of conjugate coordinates,

$$H(\theta, p_\theta, \alpha) = \frac{1}{2mr^2}(p_\theta)^2 + \frac{1}{2mr^2\sin^2\theta}\alpha^2 + mgr(1+\cos\theta). \tag{10.6}$$

Goldstein notes that the behavior of the cyclic coordinate follows by integrating

$$\dot{q}_n = \frac{\partial H}{\partial \alpha}. \tag{10.7}$$

In this example $\alpha = p_\theta$, so this is really just one of our Hamiltonian equations

$$\dot{\phi} = \frac{\partial H}{\partial p_\phi}. \tag{10.8}$$

Okay, good. First part of the mission is accomplished. The setup for Routh's procedure no longer has anything mysterious to it. Now, Goldstein defines the Routhian as

$$R = p_i\dot{q}_i - L, \tag{10.9}$$

where the index i is summed only over the cyclic (ignorable) coordinates. For this spherical pendulum example, this is $q_i = \phi$, and $p_i = mr^2 \sin^2\theta\dot{\phi}$, for

$$R = \frac{1}{2}mr^2\left(-\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) + mgr(1+\cos\theta). \tag{10.10}$$

Now, we should also have for the non-cyclic coordinates, just like the Euler-Lagrange equations

$$\frac{\partial R}{\partial\theta} = \frac{d}{dt}\frac{\partial R}{\partial\dot{\theta}}. \tag{10.11}$$

Evaluating this we have

$$mr^2\sin\theta\cos\theta\dot{\phi}^2 - mgr\sin\theta = \frac{d}{dt}\left(-mr^2\dot{\theta}\right). \tag{10.12}$$

It would be reasonable now to compare this the θ Euler-Lagrange equations, but evaluating those we get

$$mr^2\sin\theta\cos\theta\dot{\phi}^2 + mgr\sin\theta = \frac{d}{dt}\left(mr^2\dot{\theta}\right). \tag{10.13}$$

Bugger. We have got a sign difference on the $\dot{\phi}^2$ term.

10.3 SIMPLER PLANAR EXAMPLE.

Having found an inconsistency with Routhian formalism and the concrete example of the spherical pendulum which has a cyclic coordinate as desired, let us step back slightly, and try a simpler example, artificially constructed

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x). \tag{10.14}$$

Our Hamiltonian and Routhian functions are

$$\begin{aligned} H &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + V(x) \\ R &= \frac{1}{2}m(-\dot{x}^2 + \dot{y}^2) + V(x). \end{aligned} \tag{10.15}$$

For the non-cyclic coordinate we should have

$$\frac{\partial R}{\partial x} = \frac{d}{dt}\frac{\partial R}{\partial\dot{x}}, \tag{10.16}$$

which is

$$V'(x) = \frac{d}{dt}\left(-m\dot{x}\right). \tag{10.17}$$

Okay, good, that is what is expected, and exactly what we get from the Euler-Lagrange equations. This looks good, so where did things go wrong in the spherical pendulum evaluation.

10.4 POLAR FORM EXAMPLE.

The troubles appear to come from when there is a velocity coupling in the Kinetic energy term. Let us try one more example with a simpler velocity coupling, using polar form coordinates in the plane, and a radial potential. Our Lagrangian, and conjugate momenta, and Hamiltonian, respectively are

$$\begin{aligned} L &= \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - V(r) \\ p_r &= m\dot{r} \\ p_\theta &= mr^2\dot{\theta} \\ H &= \frac{1}{2m}\left((p_r)^2 + \frac{1}{r^2}(p_\theta)^2\right) + V(r). \end{aligned} \tag{10.18}$$

Evaluation of the Euler-Lagrange equations gives us the equations of motion

$$\begin{aligned} \frac{d}{dt}\left(m\dot{r}\right) &= mr\dot{\theta}^2 - V'(r) \\ \frac{d}{dt}\left(mr^2\dot{\theta}\right) &= 0. \end{aligned} \tag{10.19}$$

Evaluation of the Hamiltonian equations $\partial_p H = \dot{q}$, $\partial_q H = -\dot{p}$ should give the same results. First for r this gives

$$\begin{aligned} \frac{1}{m}p_r &= \dot{r} \\ -\frac{1}{mr^3}(p_\theta)^2 + V'(r) &= -\dot{p}_r. \end{aligned} \tag{10.20}$$

The first just defines the canonical momentum (in this case the linear momentum for the radial aspect of the motion), and the second after some rearrangement is

$$mr(\dot{\theta})^2 - V'(r) = \frac{d}{dt}\left(m\dot{r}\right), \tag{10.21}$$

which is consistent with the Lagrangian approach. For the θ evaluation of the Hamiltonian equations we get

$$\begin{aligned}\frac{p_\theta}{mr^2} &= \dot{\theta} \\ 0 &= -\dot{p}_\theta.\end{aligned} \tag{10.22}$$

The first again, is implicitly, the definition of our canonical momentum (angular momentum in this case), while the second is the conservation condition on the angular momentum that we expect associated with this ignorable coordinate. So far so good. Everything is as it should be, and there is nothing new here. Just Lagrangian and Hamiltonian mechanics as usual. But we have two independently calculated results that are the same and the Routhian procedure should generate the same results.

Now, on to the Routhian. There we have a Hamiltonian like sum of $p\dot{q}$ terms over all cyclic coordinates, minus the Lagrangian. Here the θ coordinate is observed to be that cyclic coordinate, so this is

$$\begin{aligned}R &= p_\theta\dot{\theta} - L \\ &= mr^2\dot{\theta}^2 - \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + V(r) \\ &= \frac{1}{2}mr^2\dot{\theta}^2 - \frac{1}{2}m\dot{r}^2 + V(r).\end{aligned} \tag{10.23}$$

Now, this Routhian can be written in a few different ways. In particular for the $\dot{\theta}$ dependent term of the kinetic energy we can write

$$\frac{1}{2}mr^2\dot{\theta}^2 = \frac{1}{2mr^2}(p_\theta)^2 = \frac{1}{2}\dot{\theta}p_\theta. \tag{10.24}$$

Looking at the troubles obtaining the correct equations of motion from the Routhian, it appears likely that this freedom is where things go wrong. In the Cartesian coordinate description, where there was no coupling between the coordinates in the kinetic energy we had no such freedom. Looking back to Goldstein, I see that he writes the Routhian in terms of a set of explicit variables

$$R = R(q_1, \cdots q_n, p_1, \cdots p_s, \dot{q}_{s+1}, \cdots \dot{q}_n, t) = \sum_{i=1}^{s} \dot{q}_i p_i - L. \tag{10.25}$$

where $q_1, \cdots q_s$ were the cyclic coordinates. Additionally, taking the differential he writes

$$\begin{aligned} dR &= \sum_{i=1}^{s} \dot{q}_i dp_i - \sum_{i=1}^{s} \frac{\partial L}{\partial q_i} dq_i - \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt \\ &= \frac{\partial R}{\partial p_i} dp_i + \frac{\partial R}{\partial q_i} dq_i + \frac{\partial R}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial R}{\partial t} dt, \end{aligned} \tag{10.26}$$

with sums implied in the second total differential. It was term by term equivalence of these that led to the Routhian equivalent of the Euler-Lagrange equations for the non-cyclic coordinates, from which we should recover the desired equations of motion. Notable here is that we have no $\dot{q}_i$ for any of the cyclic coordinates q_i.

For this planar radial Lagrangian, it appears that we must write the Routhian, specifically as $R = R(r, \theta, p_\theta, \dot{r})$, so that we have no explicit dependence on the radial conjugate momentum. That is

$$R = \frac{1}{2mr^2}(p_\theta)^2 - \frac{1}{2}m\dot{r}^2 + V(r). \tag{10.27}$$

As a consequence of eq. (10.26) we should recover the equations of motion by evaluating $\delta R/\delta r = 0$, and doing so for eq. (10.27) we have

$$\frac{\delta R}{\delta r} = V'(r) - \frac{1}{mr^3}(p_\theta)^2 - \frac{d}{dt}\left(-m\dot{r}\right) = 0. \tag{10.28}$$

Good. This agrees with our result from the Lagrangian and Hamiltonian formalisms. On the other hand, if we evaluate this variational derivative for

$$R = \frac{1}{2}mr^2\dot{\theta}^2 - \frac{1}{2}m\dot{r}^2 + V(r), \tag{10.29}$$

something that is formally identical, but written in terms of the "wrong" variables, we get a result that is in fact wrong

$$\frac{\delta R}{\delta r} = mr\dot{\theta}^2 + V'(r) - \frac{d}{dt}\left(-m\dot{r}\right) = 0. \tag{10.30}$$

Here the term that comes from the $\dot{\theta}$ dependent part of the Kinetic energy has an incorrect sign. This was precisely the problem observed in the initial attempt to work the spherical pendulum equations of motion starting from the Routhian.

What variables to use to express the equations is a rather subtle difference, but if we do not get that exactly right the results are garbage. Next

step here is go back and revisit the spherical polar pendulum and verify that being more careful with the variables used to express R allows the correct answer to be obtained. That exercise is probably for a different day, and probably a paper only job.

Now, I note that Goldstein includes no problems for this Routhian formalism now that I look, and having worked an example successfully and seeing how we can go wrong, it is not quite clear what his point including this was. Perhaps that will become clearer later. I had guess that some of the value of this formalism could be once one attempts numerical solutions and finds the cyclic coordinates as a result of a linear approximation of the system equations around the neighborhood of some phase space point.

11 RIGID BODY MOTION.

11.1 RIGID BODY MOTION.

11.1.1 *Setup.*

We will consider either rigid bodies as in the connected by sticks fig. 11.1 or a body consisting of a continuous mass as in fig. 11.2 In the first figure

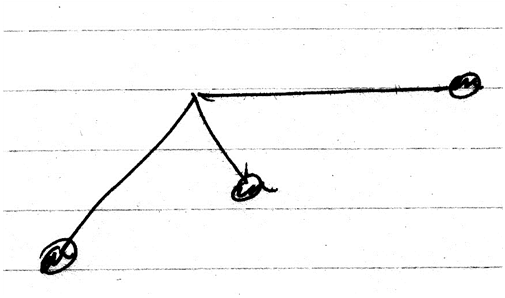


Figure 11.1: Rigid body of point masses.

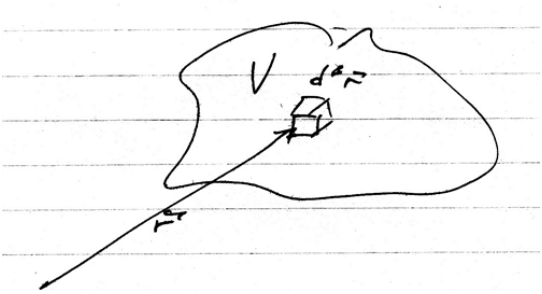


Figure 11.2: Rigid solid body of continuous mass.

our mass is made of discrete particles

$$M = \sum m_i. \tag{11.1}$$

whereas in the second figure with mass density $\rho(\mathbf{r})$ and a volume element $d^3\mathbf{r}$, our total mass is

$$M = \int_V \rho(\mathbf{r}) d^3\mathbf{r}. \tag{11.2}$$

11.1.2 *Degrees of freedom.*

How many numbers do we need to describe fixed body motion. Consider fig. 11.3

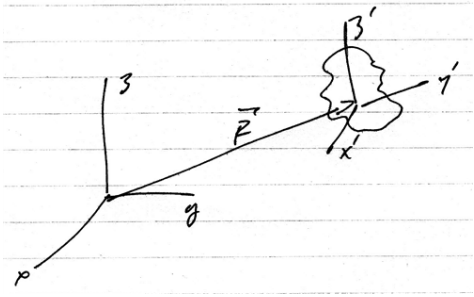


Figure 11.3: Body local coordinate system with vector to a fixed point in the body.

We will need to use six different numbers to describe the motion of a rigid body. We need three for the position of the body $\mathbf{R}_{CM}$ as a whole. We also need three degrees of freedom (in general) for the motion of the body at that point in space (how our local coordinate system at the body move at that point), describing the change of the orientation of the body as a function of time.

Note that the angle ϕ has not been included in any of the pictures because it is too messy with all the rest. Picture something like fig. 11.4

Let us express the position of the body in terms of that body's center of mass

$$\mathbf{R}_{CM} = \frac{\sum_i m_i \mathbf{r}_i}{\sum_j m_j}, \tag{11.3}$$

or for continuous masses

$$\mathbf{R}_{CM} = \frac{\int_V d^3\mathbf{r}' \mathbf{r}' \rho(\mathbf{r}')}{\int d^3\mathbf{r}'' \rho(\mathbf{r}'')}. \tag{11.4}$$

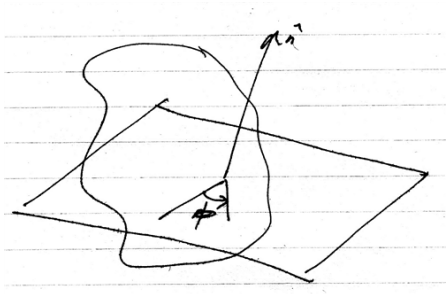


Figure 11.4: Rotation angle and normal in the body.

We consider the motion of point **P**, an arbitrary point in the body as in fig. 11.5, whos motion consists of

1. displacement of the CM $\mathbf{R}_{CM}$
2. rotation of $\mathbf{r}$ around some axis $\hat{\mathbf{n}}$ going through CM on some angle ϕ. (here $\hat{\mathbf{n}}$ is a unit vector).

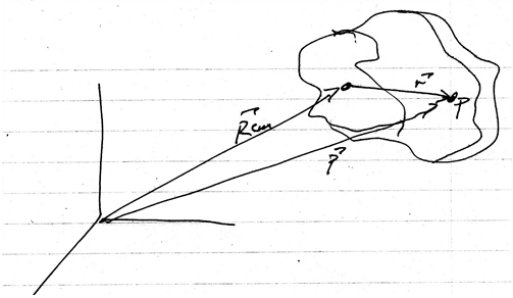


Figure 11.5: A point in the body relative to the center of mass.

From the picture we have

$$\boldsymbol{\rho} = \mathbf{R}_{CM} + \mathbf{r} \tag{11.5a}$$

$$d\boldsymbol{\rho} = d\mathbf{R}_{CM} + d\boldsymbol{\phi} \times \mathbf{r}, \tag{11.5b}$$

where

$$d\boldsymbol{\phi} = \hat{\mathbf{n}} d\phi. \tag{11.6}$$

Dividing by dt we have

$$\frac{d\boldsymbol{\rho}}{dt} = \frac{d\mathbf{R}_{CM}}{dt} + \frac{d\boldsymbol{\phi}}{dt} \times \mathbf{r}. \tag{11.7}$$

The total velocity of this point in the body is then

$$\mathbf{v} = \mathbf{V}_{CM} = \boldsymbol{\Omega}_{CM} \times \mathbf{r}. \tag{11.8}$$

where

$$\boldsymbol{\Omega}_{CM} = \frac{d\boldsymbol{\phi}}{dt} = \frac{d(\hat{\mathbf{n}}\phi)}{dt} = \text{angular velocity of the body.} \tag{11.9}$$

This circular motion is illustrated in fig. 11.6 Note that $\mathbf{v}$ is the velocity of

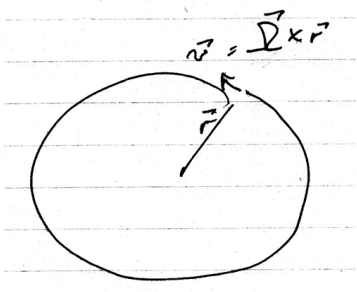


Figure 11.6: circular motion.

the particle with respect to the unprimed system.

We will spend a lot of time figuring out how to express $\boldsymbol{\Omega}_{CM}$.

Now let us consider a second point as in fig. 11.7

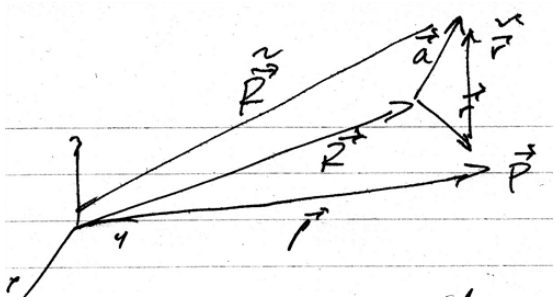


Figure 11.7: Two points in a rigid body.

$$\begin{aligned} \boldsymbol{\rho} &= \mathbf{R} + \mathbf{r} \\ \boldsymbol{\rho} &= \tilde{\mathbf{R}} + \tilde{\mathbf{r}} \\ \tilde{\mathbf{r}} &= \mathbf{r} + \mathbf{a}. \end{aligned} \tag{11.10}$$

we have

$$\begin{aligned}
\mathbf{v}_p &= \frac{d\boldsymbol{\rho}}{dt} \\
&= \frac{d\mathbf{r}}{dt} + \frac{d\boldsymbol{\phi}}{dt} \times \mathbf{r} \\
&= \frac{d\mathbf{r}}{dt} + \frac{d\boldsymbol{\phi}}{dt} \times (\tilde{\mathbf{r}} - \mathbf{a}) \\
&= \frac{d\mathbf{r}}{dt} - \frac{d\boldsymbol{\phi}}{dt} \times \mathbf{a} + \frac{d\boldsymbol{\phi}}{dt} \times \tilde{\mathbf{r}}.
\end{aligned} \tag{11.11}$$

$$\boxed{\mathbf{v}_p = \mathbf{V}_{CM} - \boldsymbol{\Omega}_{CM} \times \mathbf{a} + \boldsymbol{\Omega}_{CM} \times \tilde{\mathbf{r}}.} \tag{11.12}$$

Have another way that we can use to express the position of the point

$$\frac{d\boldsymbol{\rho}}{dt} = \frac{d\tilde{\mathbf{R}}}{dt} + \frac{d\tilde{\boldsymbol{\rho}}}{dt} \times \tilde{\mathbf{r}}_p, \tag{11.13}$$

or

$$\mathbf{v}_p = \mathbf{V}_A + \boldsymbol{\Omega}_A \times \tilde{\mathbf{r}}_p. \tag{11.14}$$

Equating with above, and noting that this holds for all $\tilde{\mathbf{r}}_p$, and noting that if $\tilde{\mathbf{r}}_p = 0$

$$\mathbf{V}_A = \mathbf{V}_{CM} - \boldsymbol{\Omega}_{CM} \times \mathbf{a}, \tag{11.15}$$

hence

$$\boldsymbol{\Omega}_{CM} \times \tilde{\mathbf{r}}_p = \boldsymbol{\Omega}_A \times \tilde{\mathbf{r}}_p. \tag{11.16}$$

or

$$\boldsymbol{\Omega}_{CM} = \boldsymbol{\Omega}_A. \tag{11.17}$$

The moral of the story is that the angular velocity $\boldsymbol{\Omega}$ is a characteristic of the system. It does not matter if it is calculated with respect to the center of mass or not.

See some examples in the notes.

11.2 KINETIC ENERGY.

For all P in the body we have

$$\mathbf{v}_p = \mathbf{V}_A + \boldsymbol{\Omega} \times \mathbf{r}_p. \tag{11.18}$$

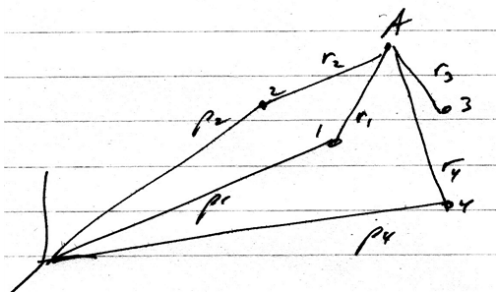


Figure 11.8: Kinetic energy setup relative to point A in the body.

here $\mathbf{V}_A$ is an arbitrary fixed point in the body as in fig. 11.8 The kinetic energy is

$$\begin{aligned} T &= \sum_a \frac{1}{2} m_a \boldsymbol{\rho}_a \\ &= \sum_a \frac{1}{2} \mathbf{v}_a^2 \\ &= \sum_a \frac{1}{2} \left(\mathbf{V}_A + \boldsymbol{\Omega} \times \mathbf{r}_a\right)^2 \\ &= \sum_a \frac{1}{2} \left(\mathbf{V}_A^2 + 2\mathbf{V}_A \cdot (\boldsymbol{\Omega} \times \mathbf{r}_a) + (\boldsymbol{\Omega} \times \mathbf{r}_a)^2\right). \end{aligned} \tag{11.19}$$

We see that if we take A to be the center of mass then our cross term

$$\begin{aligned} \sum_a m_a \mathbf{V}_A \cdot (\boldsymbol{\Omega} \times \mathbf{r}_a) &= \mathbf{V}_A \cdot \left(\boldsymbol{\Omega} \times \sum_a m_a \mathbf{r}_a\right) \\ &= \mathbf{V}_A \cdot (\boldsymbol{\Omega} \times \mathbf{R}_{CM}), \end{aligned} \tag{11.20}$$

which vanishes. With

$$\mu = \sum_a m_a, \tag{11.21}$$

we have

$$T = \frac{1}{2}\mu \mathbf{V}_{CM}^2 + \frac{1}{2} \sum_a (\boldsymbol{\Omega} \times \mathbf{r}_a) \cdot (\boldsymbol{\Omega} \times \mathbf{r}_a). \tag{11.22}$$

With

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}), \tag{11.23}$$

or

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{A}^2\mathbf{B}^2 - (\mathbf{A}\cdot\mathbf{B})^2. \tag{11.24}$$

Forgetting about the μ dependent term for now we have

$$T = \frac{1}{2}\sum_a m_a\left(\mathbf{\Omega}^2\mathbf{r}_a^2 - (\mathbf{\Omega}\cdot\mathbf{r}_a)^2\right). \tag{11.25}$$

Expanding this out with

$$\mathbf{r}_a = (r_{a_1} r_{a_2}, r_{a_3}) = \{r_{a_i}\}, \tag{11.26}$$

and

$$\mathbf{\Omega} = (\Omega_1\Omega_2, \Omega_3) = \{\Omega_i\}, \tag{11.27}$$

we have

$$T = \frac{1}{2}\sum_a m_a\left(\Omega_k\Omega_k r_{a_j} r_{a_j} - (\Omega_k r_{a_k})^2\right). \tag{11.28}$$

12 EULER ANGLES.

12.1 PICTORIALLY.

We want to look at some of the trig behind expressing general rotations. We can perform a general rotation by a sequence of successive rotations. One such sequence is a rotation around the z, x, z axes in sequence. Application of a rotation of angle ϕ takes us from our original fig. 12.1 frame to that of fig. 12.2. A second rotation around the (new) x axis by angle θ takes us to fig. 12.3, and finally a rotation of ψ around the (new) z axis, takes us to fig. 12.4.

A composite image of all of these rotations taken together can be found in fig. 12.5.

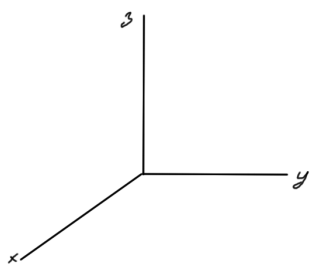


Figure 12.1: Initial frame.

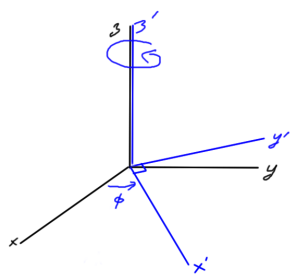


Figure 12.2: Rotation by ϕ around z axis.

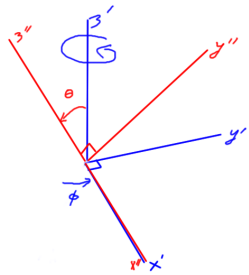


Figure 12.3: Rotation of θ around (new) x axis.

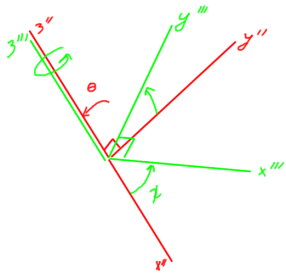


Figure 12.4: Rotation of ψ around (new) z axis.

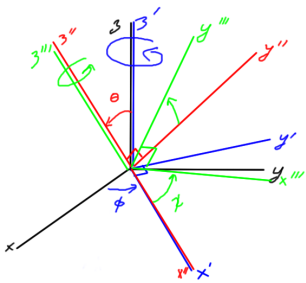


Figure 12.5: All three rotations superimposed.

12.2 RELATING THE TWO PAIRS OF COORDINATE SYSTEMS.

Let us look at this algebraically instead, using fig. 12.6 as a guide. Step 1.

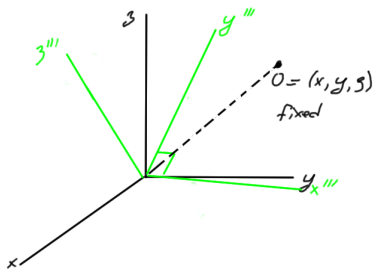


Figure 12.6: A point in two coordinate systems.

Rotation of ϕ around z

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{12.1}$$

Step 2. Rotation around x'.

$$\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}. \tag{12.2}$$

Step 3. Rotation around z''.

$$\begin{bmatrix} x''' \\ y''' \\ z''' \end{bmatrix} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix}. \tag{12.3}$$

So, our full rotation is the composition of the rotation matrices

$$\begin{bmatrix} x''' \\ y''' \\ z''' \end{bmatrix} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{12.4}$$

Let us introduce some notation and write this as

$$B_z(\alpha) = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12.5a}$$

$$B_x(\theta) = \begin{bmatrix} 1 & 0 & 0 & \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \end{bmatrix}, \tag{12.5b}$$

so that we have the mapping

$$\mathbf{r} \to B_z(\psi)B_x(\theta)B_z(\phi)\mathbf{r}. \tag{12.6}$$

Now let us write

$$\mathbf{r} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{12.7}$$

We will call

$$A(\psi, \theta, \phi) = B_z(\psi)B_x(\theta)B_z(\phi), \tag{12.8}$$

so that

$$x_i' = \sum_{j=1}^{3} A_{ij}x_j. \tag{12.9}$$

We will drop the explicit summation sign, so that the summation over repeated indices are implied

$$x_i' = A_{ij}x_j. \tag{12.10}$$

This matrix $A(\psi, \theta, \phi)$ is in fact a general parameterization of the 3×3 special orthogonal matrices. The set of three angles θ, ϕ, ψ parameterizes all rotations in 3dd space. Transformations that preserve $\mathbf{a} \cdot \mathbf{b}$ and have unit determinant. In symbols we must have

$$A^{\mathrm{T}}A = 1. \tag{12.11}$$

$$\det A = +1. \tag{12.12}$$

Having solved this auxiliary problem, we now want to compute the angular velocity.

We want to know how to express the coordinates of a point that is fixed in the body. i.e. We are fixing x'_i and now looking for x_i.

The coordinates of a point that has x', y' and z' in a body-fixed frame, in the fixed frame are x, y, z. That is given by just inverting the matrix

$$\begin{aligned}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= A^{-1}(\psi, \theta, \phi) \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} \\
&= B_z^{-1}(\phi) B_x^{-1}(\theta) B_z^{-1}(\psi) \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} \\
&= B_z^{\mathrm{T}}(\phi) B_x^{\mathrm{T}}(\theta) B_z^{\mathrm{T}}(\psi) \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix}.
\end{aligned} \tag{12.13}$$

Here we have used the fact that B_x and B_z are orthogonal, so that their inverses are just their transposes.

We have finally

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = B_z(-\phi) B_x(-\theta) B_z(-\psi) \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix}. \tag{12.14}$$

If we assume that ψ, θ and ϕ are functions of time, and compute $d\mathbf{r}/dt$. Starting with

$$x_i = [A^{-1}(\phi, \theta, \psi)]_{ij} x'_j, \tag{12.15}$$

$$\Delta x_i = \Big([A^{-1}(\phi + \Delta\phi, \theta + \Delta\theta, \psi + \Delta\psi)]_{ij} - [A^{-1}(\phi, \theta, \psi)]_{ij}\Big) x'_j. \tag{12.16}$$

For small changes, we can Taylor expand and retain only the first order terms. Doing that and dividing by Δt we have

$$\frac{dx_i}{dt} = \left(\frac{\partial}{\partial\psi}A_{ij}^{-1}\dot{\psi} + \frac{\partial}{\partial\theta}A_{ij}^{-1}\dot{\theta} + \frac{\partial}{\partial\phi}A_{ij}^{-1}\dot{\phi}\right)x'_j. \tag{12.17}$$

Now, we use

$$x'_j = A_{jl}x_l, \tag{12.18}$$

so that we have

$$\frac{dx_i}{dt} = \left(\left(\frac{\partial}{\partial\psi}A_{ij}^{-1}\right)A_{jl}\dot{\psi} + \left(\frac{\partial}{\partial\theta}A_{ij}^{-1}\right)A_{jl}\dot{\theta} + \left(\frac{\partial}{\partial\phi}A_{ij}^{-1}\right)A_{jl}\dot{\phi}\right)x_l. \tag{12.19}$$

We are looking for a relation of the form

$$\frac{d\mathbf{r}}{dt} = \mathbf{\Omega}\times\mathbf{r}. \tag{12.20}$$

We can write this as

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \left(\dot{\theta}\frac{\partial A^{-1}}{\partial\theta}A + \dot{\phi}\frac{\partial A^{-1}}{\partial\phi}A + \dot{\psi}\frac{\partial A^{-1}}{\partial\psi}A\right)\begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{12.21}$$

Actually doing this calculation is asked of us in HW6. The final answer is

$$\frac{dx_i}{dt} = \left(\dot{\phi}\epsilon_{ijk}j_n^{\phi} + \dot{\theta}\epsilon_{ijk}j_n^{\theta} + \dot{\psi}\epsilon_{ijk}j_n^{\psi}\right)x_j. \tag{12.22}$$

Here ϵ_{ijk} is the usual fully antisymmetric tensor with properties

$$\epsilon_{ijk} = \begin{cases} 0 & \text{when any of the indices are equal.} \\ 1 & \text{for any of } ijk = 123, 231, 312 \text{ (cyclic permutations of 123.} \\ -1 & \text{for any of } ijk = 213, 132, 321. \end{cases} \tag{12.23}$$

13 PARALLEL AXIS THEOREM.

We can express the kinetic energy as

$$T = \frac{1}{2}\sum_{i,j=1}^{3} \Omega_i I_{ij} \Omega_j, \tag{13.1}$$

where

$$I_{ij} = \sum_a m_a \left(\delta_{ij}\mathbf{r}_a^2 - r_{a_i} r_{a_j}\right). \tag{13.2}$$

Here a is a sum over all particles in the body.

If the body is continuous and $\rho(\mathbf{r})$ is the mass density then the mass inside is

$$m = \int d^3\mathbf{r}\rho(\mathbf{r}), \tag{13.3}$$

where we integrate over a volume element as in fig. 13.1.

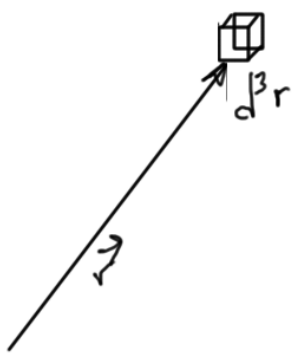


Figure 13.1: Volume element for continuous mass distribution.

For this continuous case we have

$$I_{ij} = \int_V d^3 r\rho(\mathbf{r})\left(\delta_{ij}\mathbf{r}^2 - r_i r_j\right). \tag{13.4}$$

Another property of I_{ij} is the parallel axis theorem (or as it is known in Europe and perhaps elsewhere, as the "Steiner theorem").

Let's consider a change of origin as in fig. 13.2. We write

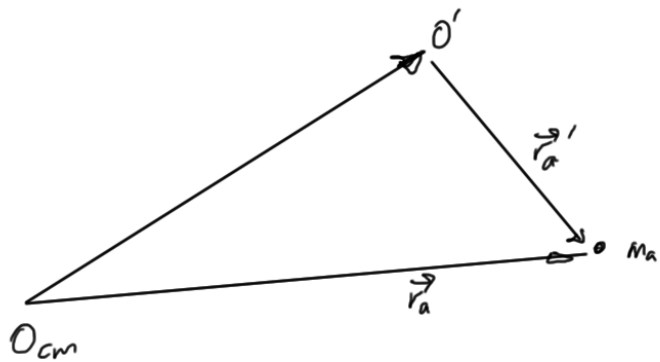


Figure 13.2: Shift of origin.

$$\mathbf{r}_a = \mathbf{r}'_a + \mathbf{b}, \tag{13.5}$$

and I'_{ij} for the inertia tensor with respect to O'. Write

$$\mathbf{r}_{a_i} = \mathbf{r}'_{a_i} + \mathbf{b}_i, \tag{13.6}$$

or

$$\mathbf{r}'_{a_i} = \mathbf{r}_{a_i} - \mathbf{b}_i, \tag{13.7}$$

so that

$$\begin{aligned} I'_{ij} &= \sum_a m_a \left(\delta_{ij}\mathbf{r}_a^2 - r'_{a_i} r'_{b_j}\right) \\ &= \sum_a m_a \left(\delta_{ij}(\mathbf{r}_a - \mathbf{b})^2 - (r_{a_i} - b_i)(r_{b_j} - b_j)\right) \\ &= \sum_a m_a \left(\delta_{ij}(r_{a_k} r_{a_k} - \mathbf{b}^2 - 2r_a b) - r_{a_i} r_{b_j} - b_i b_j + r_{a_i} b_j + r_{b_j} b_i\right), \end{aligned} \tag{13.8}$$

but, by definition of center of mass, we have

$$\sum_a m_a r'_{a_i} = 0, \tag{13.9}$$

so

$$
\begin{aligned}
I'_{ij} &= \sum_a m_a \left(\delta_{ij}\mathbf{r}_a^2 - r_{a_i} r_{a_j} - \cdots\right) \\
&= I_{ij} - 2\cancel{\left(\sum_a m_a \mathbf{r}_a \cdot \mathbf{b}\delta_{ij}\right)} + \mu\left(\delta_{ij}\mathbf{b}^2 - b_i b_j\right).
\end{aligned}
\tag{13.10}
$$

This is

$$
I'_{ij} = I_{ij}^{\mathrm{CM}} + \mu\left(\delta_{ij}\mathbf{b}^2 - b_i b_j\right). \tag{13.11}
$$

Some examples. Infinite cylinder rolling on a plane, with no slipping and no dissipation (heat?) as in fig. 13.3. Take the mass as uniform and

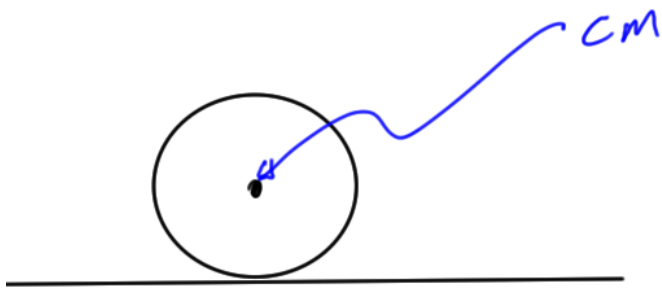


Figure 13.3: Infinite rolling cylinder on plane.

set up coordinates as in fig. 13.4.

No slip means on revolution, the center of mass moves $2\pi R$. We have one degree of freedom: ϕ.

$$
|\Omega| = \dot{\phi} = \frac{d\phi}{dt}. \tag{13.12}
$$

This is the angular velocity.

$$
\frac{\Delta\phi}{\Delta x} = \frac{2\pi}{2\pi R}. \tag{13.13}
$$

so

$$
\Delta x = R\Delta\phi. \tag{13.14}
$$

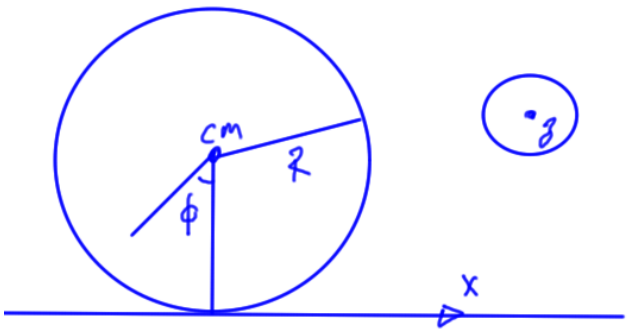


Figure 13.4: Coordinates for infinite cylinder.

The kinetic energy is

$$\begin{aligned} T &= \frac{1}{2}\mu V_{\text{CM}}^2 + \frac{1}{2}\Omega_3^2 I_{33} \\ &= \frac{1}{2}\mu V_{\text{CM}}^2 + \frac{1}{2}\Omega^2 I. \end{aligned} \tag{13.15}$$

$$V_{\text{CM}} = \frac{\Delta x}{\Delta t} = R\frac{\Delta \phi}{\Delta t} = R\Omega = R\dot{\phi}, \tag{13.16}$$

so

$$T = \frac{1}{2}\mu R^2 \dot{\phi}^2 + \frac{1}{2}\dot{\phi}^2 I. \tag{13.17}$$

(can calculate I : See notes or derive).

Now suppose the CM is displaced as in fig. 13.5. Perhaps a hollow tube with a blob attached as in fig. 13.6, where the torque is now due to gravity. This can have more interesting motion. Example: Oscillation. This is a typical test question, where calculation of the frequency of oscillation is requested. Such a question would probably be posed with the geometry of fig. 13.7. Recall for a general body as in fig. 13.8. Write

$$\mathbf{r} = \mathbf{a} + \mathbf{r}', \tag{13.18}$$

and

$$\mathbf{v} = \mathbf{V}_{\text{CM}} + \mathbf{\Omega} \times \mathbf{r}, \tag{13.19}$$

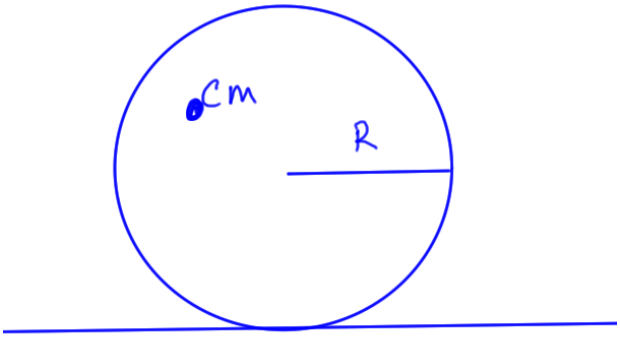


Figure 13.5: Displaced CM for infinite cylinder.

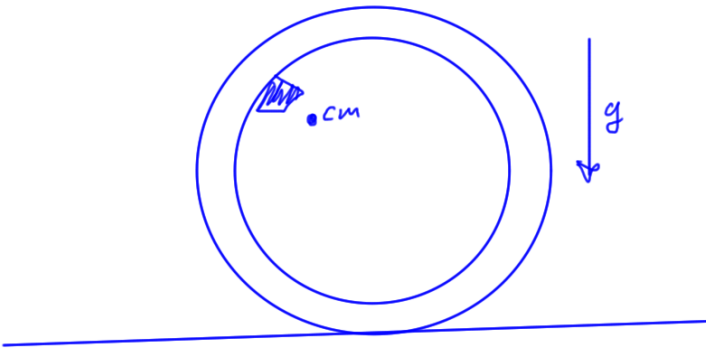


Figure 13.6: Hollow tube with blob.

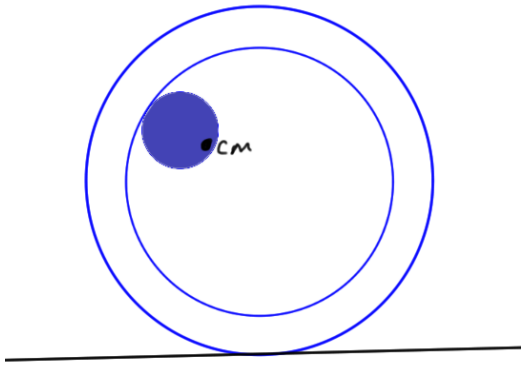


Figure 13.7: Hollow tube with cylindrical blob.

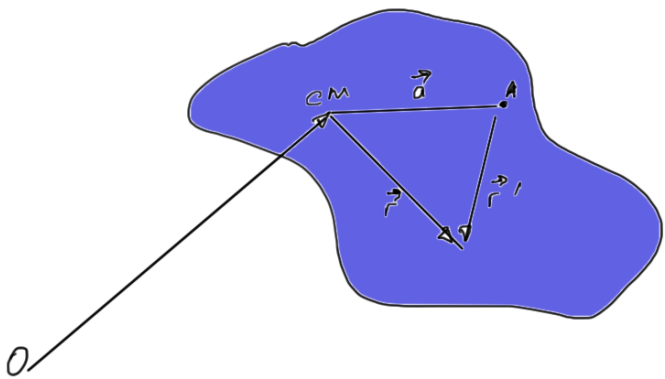


Figure 13.8: general body coordinates.

or

$$\mathbf{v} = \mathbf{V}_{\text{CM}} + \mathbf{\Omega} \times \mathbf{a} + \mathbf{\Omega} \times \mathbf{r}'. \tag{13.20}$$

Here $\mathbf{V}_{\text{CM}}$ is the velocity of the origin A.

If $\mathbf{V}_{\text{CM}}$ and $\mathbf{\Omega}$ are perpendicular always there always exists $\mathbf{a}$ such that A is at rest.

Another example is a cone on plane or rod as in fig. 13.9. (this is another

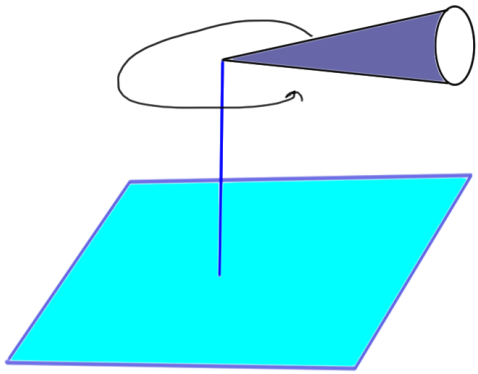


Figure 13.9: Cone on rod.

typical test question).

For cylinder that point is the contact between plane and cylinder. This is called the momentary axis of rotation: fig. 13.10. Using this is a very useful trick.

Aside. more interesting is the cone viewed from above as in fig. 13.11. Coordinates for this problem as in fig. 13.12. Using eq. (13.20) we have

$$V_{\text{CM}} = \mathbf{\Omega} \times \mathbf{b} = \dot{\phi}\hat{\mathbf{z}} \times \mathbf{b}, \tag{13.21}$$

where this followed from

$$\mathbf{v} = \mathbf{\Omega} \times \mathbf{r}'. \tag{13.22}$$

Here $\mathbf{r}'$ is the vector from axes of momentary rotation to point.

Our kinetic energy is

$$T = \frac{1}{2}\mu V_{\text{CM}}^2 + \frac{I}{2}\dot{\phi}^2, \tag{13.23}$$

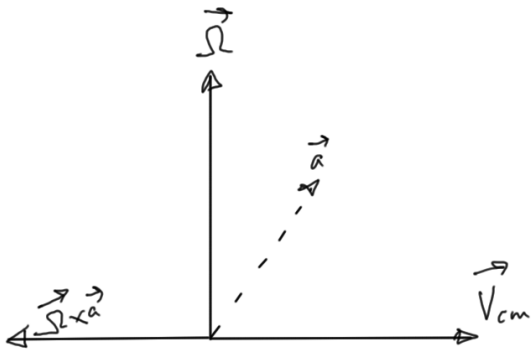


Figure 13.10: Momentary axes of rotation.

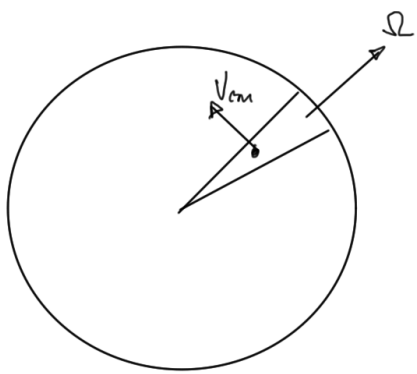


Figure 13.11: Cone from above.

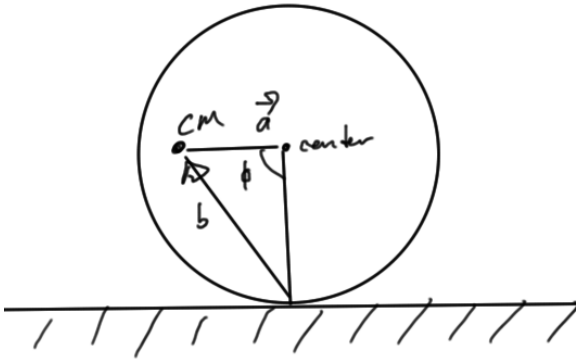


Figure 13.12: Momentary axes of rotation for cone on stick.

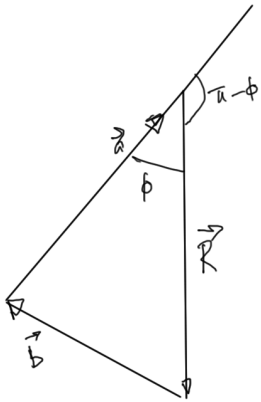


Figure 13.13: Coordinates.

and our coordinates are fig. 13.13.

$$\mathbf{V}_{\mathrm{CM}} = \mathbf{\Omega} \times \mathbf{b}. \tag{13.24}$$

$$\begin{aligned} |\mathbf{V}_{\mathrm{CM}}| &= \left|\dot{\phi}\right||\mathbf{b}| \\ &= \dot{\phi}|\mathbf{b}| \times \text{moving unit vector in x y plane} \\ &= \dot{\phi}\sqrt{\mathbf{a}^2 + \mathbf{R}^2 + 2\mathbf{a}\cdot\mathbf{R}} \\ &= \dot{\phi}\sqrt{a^2 + R^2 + 2aR\cos(\pi - \phi)}. \end{aligned} \tag{13.25}$$

For

$$\begin{aligned} T &= \frac{\mu}{2}\dot{\phi}^2\left(a^2 + R^2 + 2aR\cos(\pi - \phi)\right) + \frac{I}{2}\dot{\phi}^2 \\ &= \frac{1}{2}\dot{\phi}^2\left(\mu\left(a^2 + R^2 + 2aR\cos(\pi - \phi)\right) + I\right), \end{aligned} \tag{13.26}$$

and

$$L = T - \mu g \overbrace{\boxed{(R - a\cos\phi)}}^{\text{Height of CM above plane}}. \tag{13.27}$$

This gravity portion accounts for the torque producing interesting effects.

14 PHASE SPACE AND TRAJECTORIES.

14.1 PHASE SPACE AND PHASE TRAJECTORIES.

The phase space and phase trajectories are the space of p's and q's of a mechanical system (always even dimensional, with as many p's as q's for N particles in 3d: 6N dimensional space).

The state of a mechanical system ≡ the point in phase space. Time evolution ≡ a curve in phase space.

Example: 1 dim system, say a harmonic oscillator.

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \tag{14.1}$$

Our phase space can be illustrated as an ellipse as in fig. 14.1 where the

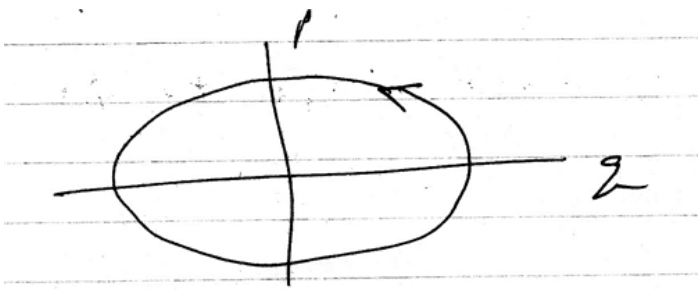


Figure 14.1: Harmonic oscillator phase space trajectory.

phase space trajectories of the SHO. The equation describing the ellipse is

$$E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2, \tag{14.2}$$

which we can put into standard elliptical form as

$$1 = \left(\frac{p}{\sqrt{2mE}}\right)^2 + \left(\sqrt{\frac{m}{2E}}\omega\right)q^2. \tag{14.3}$$

14.1.1 *Applications of H.*

- Classical stat mech.
- transition into QM via Poisson brackets.
- mathematical theorems about phase space "flow".
- perturbation theory.

14.1.2 *Poisson brackets.*

Poisson brackets arises very naturally if one asks about the time evolution of a function $f(p, q, t)$ on phase space.

$$\begin{aligned} \frac{d}{dt} f(p_i, q_i, t) &= \sum_i \frac{\partial f}{\partial p_i}\frac{\partial p_i}{\partial t} + \frac{\partial f}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial f}{\partial t} \\ &= \sum_i -\frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} + \frac{\partial f}{\partial t}. \end{aligned} \tag{14.4}$$

Define the commutator of H and f as

$$[H, f] = \sum_i \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial f}{\partial p_i}. \tag{14.5}$$

This is the Poisson bracket of $H(p, q, t)$ with $f(p, q, t)$, defined for arbitrary functions on phase space.

Note that other conventions for sign exist (apparently in Landau and Lifshitz uses the opposite).

So we have

$$\frac{d}{dt} f(p_i, q_i, t) = [H, f] + \frac{\partial f}{\partial t}. \tag{14.6}$$

Corollaries:

If f has no explicit time dependence $\partial f/\partial t = 0$ and if $[H, f] = 0$, then f is an integral of motion.

In QM conserved quantities are the ones that commute with the Hamiltonian operator.

To see the analogy better, recall def of Poisson bracket

$$[f, g] = \sum_i \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}. \tag{14.7}$$

Properties of Poisson bracket

- antisymmetric

$$[f, g] = -[g, f]. \tag{14.8}$$

- linear

$$\begin{aligned}[af + bh, g] &= a[f, g] + b[h, g] \\ [g, af + bh] &= a[g, f] + b[g, h].\end{aligned} \tag{14.9}$$

14.1.2.1 *Example. Compute p, q. commutators*

$$\begin{aligned}[p_i, p_j] &= \sum_k \frac{\partial p_i}{\partial p_k}\cancel{\frac{\partial p_j}{\partial q_k}} - \cancel{\frac{\partial p_i}{\partial q_k}}\frac{\partial p_j}{\partial p_k} \\ &= 0.\end{aligned} \tag{14.10}$$

So

$$[p_i, p_j] = 0. \tag{14.11}$$

Similarly $[q_i, q_j] = 0$.

How about

$$\begin{aligned}[q_i, p_j] &= \sum_k \cancel{\frac{\partial q_i}{\partial p_k}}\cancel{\frac{\partial p_j}{\partial q_k}} - \frac{\partial q_i}{\partial q_k}\frac{\partial p_j}{\partial p_k} \\ &= -\sum_k \delta_{ik}\delta_{jk} \\ &= -\delta_{ij}.\end{aligned} \tag{14.12}$$

So

$$[q_i, p_j] = -\delta_{ij}. \tag{14.13}$$

This provides a systematic (axiomatic) way to "quantize" a classical mechanics system, where we make replacements

$$\begin{aligned}q_i &\to \hat{q}_i \\ p_i &\to \hat{p}_i,\end{aligned} \tag{14.14}$$

and

$$\begin{aligned}[q_i, p_j] = -\delta_{ij} &\to [q_i, p_j] = i\hbar\delta_{ij} \\ H(p, q, t) &\to \hat{H}(\hat{p}, \hat{q}, t).\end{aligned} \tag{14.15}$$

So

$$\frac{[\hat{q}_i, \hat{p}_j]}{-i\hbar} = -\delta_{ij}. \tag{14.16}$$

Our quantization of time evolution is therefore

$$\begin{aligned} \frac{d}{dt}\hat{q}_i &= \frac{1}{-i\hbar}\left[\hat{H}, \hat{q}_i\right] \\ \frac{d}{dt}\hat{p}_i &= \frac{1}{-i\hbar}\left[\hat{H}, \hat{p}_i\right]. \end{aligned} \tag{14.17}$$

These are the Heisenberg equations of motion in QM.

14.1.2.2 *Conserved quantities.*

For conserved quantities f, functions of p's q's, we have

$$[f, H] = 0. \tag{14.18}$$

Considering the components M_i, where

$$\mathbf{M} = \mathbf{r} \times \mathbf{p}, \tag{14.19}$$

We can show eq. (14.25) that our Poisson brackets obey

$$\begin{aligned} [M_x, M_y] &= -M_z \\ [M_y, M_z] &= -M_x \\ [M_z, M_x] &= -M_y. \end{aligned} \tag{14.20}$$

(Prof Poppitz was not sure if he had the sign of this right for the sign convention he happened to be using for Poisson brackets in this lecture, but it appears he had it right).

These are the analogue of the momentum commutator relationships from QM right here in classical mechanics.

Considering the symmetries that lead to this conservation relationship, it is actually possible to show that rotations in 4D space lead to these symmetries and the conservation of the Runge-Lenz vector.

14.2 ADIABATIC CHANGES IN PHASE SPACE AND CONSERVED QUANTITIES.

In fig. 14.2 where we have

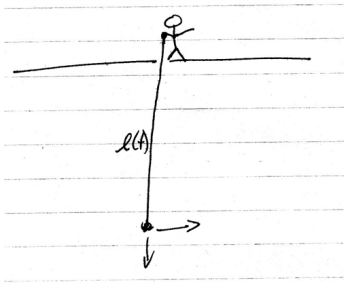


Figure 14.2: Variable length pendulum.

$$T = \frac{2\pi}{\omega(t)} = \sqrt{\frac{l(t)}{g}}. \tag{14.21}$$

Imagine that we change the length $l(t)$ very slowly so that

$$T\frac{1}{l}\frac{dl}{dt} \ll 1. \tag{14.22}$$

where T is the period of oscillation. This is what is called an adiabatic change, where the change of ω is small over a period. It turns out that if this rate of change is slow, then there is actually an invariant, and

$$\frac{E}{\omega}, \tag{14.23}$$

is the so-called "adiabatic invariant". There is an important application to this (and some relations to QM). Imagine that we have a particle bounded by two walls, where the walls are moved very slowly as in fig. 14.3

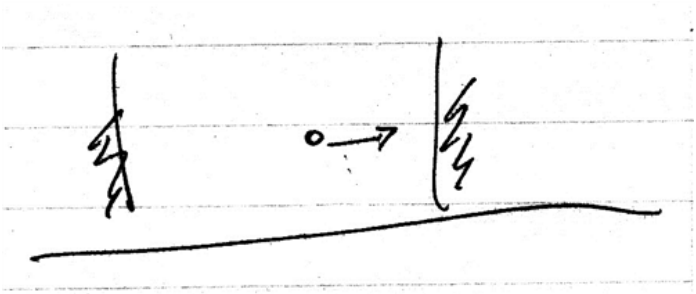


Figure 14.3: Particle constrained by slowly moving walls.

This can be used to derive the adiabatic equation for an ideal gas (also using the equipartition theorem).

14.3 APPENDIX I. POISSON BRACKETS OF ANGULAR MOMENTUM.

Let us verify the angular momentum relations of eq. (14.20) above (summation over k implied):

$$
\begin{aligned}
[M_i, M_j] &= \frac{\partial M_i}{\partial p_k}\frac{\partial M_j}{\partial x_k} - \frac{\partial M_i}{\partial x_k}\frac{\partial M_j}{\partial p_k} \\
&= \epsilon_{abi}\epsilon_{rsj}\frac{\partial x_a p_b}{\partial p_k}\frac{\partial x_r p_s}{\partial x_k} - \epsilon_{abi}\epsilon_{rsj}\frac{\partial x_a p_b}{\partial x_k}\frac{\partial x_r p_s}{\partial p_k} \\
&= \epsilon_{abi}\epsilon_{rsj}x_a\frac{\partial p_b}{\partial p_k}p_s\frac{\partial x_r}{\partial x_k} - \epsilon_{abi}\epsilon_{rsj}p_b\frac{\partial x_a}{\partial x_k}x_r\frac{\partial p_s}{\partial p_k} \\
&= \epsilon_{abi}\epsilon_{rsj}x_a\delta_{kb}p_s\delta_{kr} - \epsilon_{abi}\epsilon_{rsj}p_b\delta_{ka}x_r\delta_{sk} \\
&= \epsilon_{abi}\epsilon_{rsj}x_a p_s\delta_{br} - \epsilon_{abi}\epsilon_{rsj}p_b x_r\delta_{as} \\
&= \epsilon_{ari}\epsilon_{rsj}x_a p_s - \epsilon_{sbi}\epsilon_{rsj}p_b x_r \\
&= -\delta_{ai}^{[sj]}x_a p_s - \delta_{bi}^{[jr]}p_b x_r \\
&= -\left(\delta_{as}\delta_{ij} - \delta_{aj}\delta_{is}\right)x_a p_s - \left(\delta_{bj}\delta_{ir} - \delta_{br}\delta_{ij}\right)p_b x_r \\
&= -\delta_{as}\delta_{ij}x_a p_s + \delta_{aj}\delta_{is}x_a p_s - \delta_{bj}\delta_{ir}p_b x_r + \delta_{br}\delta_{ij}p_b x_r \\
&= -\cancel{x_s p_s\delta_{ij}} + x_j p_i - p_j x_i + \cancel{p_b x_b\delta_{ij}}.
\end{aligned}
\tag{14.24}
$$

So, as claimed, if $i \neq j \neq k$ we have

$$
[M_i, M_j] = -M_k. \tag{14.25}
$$

14.4 APPENDIX II. EOM FOR THE VARIABLE LENGTH PENDULUM.

Since we have referred to a variable length pendulum above, let us recall what form the EOM for this system take. With cylindrical coordinates as in fig. 14.4, and a spring constant $\omega_0^2 = k/m$ our Lagrangian is

$$
L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{1}{2}m\omega_0^2 r^2 - mgr(1 - \cos\theta). \tag{14.26}
$$

The EOM follows immediately

$$
\begin{aligned}
P_\theta &= \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \\
P_r &= \frac{\partial L}{\partial \dot{r}} = m\dot{r} \\
\frac{dP_\theta}{dt} &= \frac{\partial L}{\partial \theta} = -mgr\sin\theta \\
\frac{dP_r}{dt} &= \frac{\partial L}{\partial r} = mr\dot{\theta}^2 - m\omega_0^2 r - mg(1 - \cos\theta).
\end{aligned}
\tag{14.27}
$$

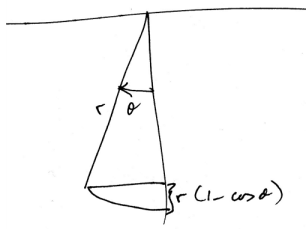


Figure 14.4: phaseSpaceAndTrajectoriesFig4.

Or

$$\begin{aligned} \frac{d}{dt}\left(r^2\dot{\theta}\right) &= -gr\sin\theta \\ \frac{d}{dt}\left(\dot{r}\right) &= r\left(\dot{\theta}^2 - \omega_0^2\right) - g(1-\cos\theta). \end{aligned} \tag{14.28}$$

Even in the small angle limit this is not a terribly friendly looking system

$$\begin{aligned} r\ddot{\theta} + 2\dot{\theta}\dot{r} + g\theta &= 0 \\ \ddot{r} - r\dot{\theta}^2 + r\omega_0^2 &= 0. \end{aligned} \tag{14.29}$$

However, in the first equation of this system

$$\ddot{\theta} + 2\dot{\theta}\frac{\dot{r}}{r} + \frac{1}{r}g\theta = 0, \tag{14.30}$$

we do see the $\dot{r}/r$ dependence mentioned in class, and see how this being small will still result in something that approximately has the form of a SHO.

15 CONSERVED QUANTITIES.

15.1 RUNGE-LENZ VECTOR CONSERVATION.

15.1.1 *Motivation.*

Notes from Prof. Poppitz's phy354 classical mechanics lecture on the Runge-Lenz vector, a less well known conserved quantity for the 3D $1/r$ potentials that can be used to solve the Kepler problem.

15.1.2 *Motivation: The Kepler problem.*

We can plug away at the Lagrangian in cylindrical coordinates and find eventually

$$\int_{\phi_0}^{\phi} d\phi = \int_{r_0}^{r} \frac{M}{mr^2} \frac{dr}{\sqrt{\frac{2}{M}\left(E - U + \frac{M^2}{2mr^2}\right)}}, \tag{15.1}$$

but this can be messy to solve, where we get elliptic integrals or worse, depending on the potential. For the special case of the 3D problem where the potential has a $1/r$ form, this is what Prof. Poppitz called "super-integrable". With $2N - 1 = 5$ conserved quantities to be found, we have got one more. Here the form of that last conserved quantity is given, called the Runge-Lenz vector, and we verify that it is conserved.

15.1.3 *Runge-Lenz vector.*

Given a potential

$$U = -\frac{\alpha}{r}, \tag{15.2}$$

and a Lagrangian

$$\begin{aligned} L &= \frac{m\dot{r}^2}{2} + \frac{1}{2}\frac{M_z^2}{mr^2} - U \\ M_z &= mr^2\dot{\phi}^2, \end{aligned} \tag{15.3}$$

and writing the angular momentum as

$$\mathbf{M} = m\mathbf{r} \times \mathbf{v}, \tag{15.4}$$

the Runge-Lenz vector

$$\mathbf{A} = \mathbf{v} \times \mathbf{M} - \alpha \hat{\mathbf{r}}, \tag{15.5}$$

is a conserved quantity.

15.1.3.1 *Verify the conservation assumption.*

Let us show that the conservation assumption is correct

$$\frac{d}{dt}(\mathbf{v} \times \mathbf{M}) = \frac{d\mathbf{v}}{dt} \times \mathbf{M} + \mathbf{v} \times \cancel{\frac{d\mathbf{M}}{dt}}. \tag{15.6}$$

Here, we note that angular momentum conservation is really $d\mathbf{M}/dt = 0$, so we are left with only the acceleration term, which we can rewrite in terms of the Euler-Lagrange equation

$$\begin{aligned}
\frac{d}{dt}(\mathbf{v} \times \mathbf{M}) &= -\frac{1}{m}\boldsymbol{\nabla} U \times M \\
&= -\frac{1}{m}\frac{\partial U}{\partial r}\hat{\mathbf{r}} \times M \\
&= -\frac{1}{m}\frac{\partial U}{\partial r}\hat{\mathbf{r}} \times (m\mathbf{r} \times \mathbf{v}) \\
&= -\frac{\partial U}{\partial r}\hat{\mathbf{r}} \times (\mathbf{r} \times \mathbf{v}).
\end{aligned} \tag{15.7}$$

We can compute the double cross product

$$\begin{aligned}
(\mathbf{a} \times (\mathbf{b} \times \mathbf{c}))_i &= a_m b_r c_s \epsilon_{rst} \epsilon_{mti} \\
&= a_m b_r c_s \delta_{im}^{[rs]} \\
&= a_m b_i c_m - a_m b_m c_i.
\end{aligned} \tag{15.8}$$

For

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}. \tag{15.9}$$

Plugging this we have

$$\begin{aligned}
\frac{d}{dt}(\mathbf{v} \times \mathbf{M}) &= \frac{\partial U}{\partial r}\left((\hat{\mathbf{r}} \cdot \mathbf{r})\mathbf{v} - (\hat{\mathbf{r}} \cdot \mathbf{v})\mathbf{r}\right) \\
&= \left(\frac{\alpha}{r^2}\right)\left(r\mathbf{v} - \frac{1}{r}(\mathbf{r} \cdot \mathbf{v})\mathbf{r}\right) \\
&= \alpha\left(\frac{\mathbf{v}}{r} - \frac{(\mathbf{r} \cdot \mathbf{v})\mathbf{r}}{r^3}\right).
\end{aligned} \tag{15.10}$$

Now let us look at the other term. We will need the derivative of $\hat{\mathbf{r}}$

$$\begin{aligned}
\frac{d\hat{\mathbf{r}}}{dt} &= \frac{d}{dt}\frac{\mathbf{r}}{r} \\
&= \frac{\mathbf{v}}{r} + \mathbf{r}\frac{d\frac{1}{r}}{dt} \\
&= \frac{\mathbf{v}}{r} - \frac{\mathbf{r}}{r^2}\frac{dr}{dt} \\
&= \frac{\mathbf{v}}{r} - \frac{\mathbf{r}}{r^2}\frac{d\sqrt{\mathbf{r}\cdot\mathbf{r}}}{dt} \\
&= \frac{\mathbf{v}}{r} - \frac{\mathbf{r}}{r^2}\frac{\mathbf{v}\cdot\mathbf{r}}{\sqrt{\mathbf{r}^2}} \\
&= \frac{\mathbf{v}}{r} - \frac{\mathbf{r}}{r^3}\mathbf{v}\cdot\mathbf{r}.
\end{aligned} \tag{15.11}$$

Putting all the bits together we have now verified the conservation statement

$$\frac{d}{dt}\left(\mathbf{v}\times\mathbf{M} - \alpha\hat{\mathbf{r}}\right) = \alpha\left(\frac{\mathbf{v}}{r} - \frac{(\mathbf{r}\cdot\mathbf{v})\mathbf{r}}{r^3}\right) - \alpha\left(\frac{\mathbf{v}}{r} - \frac{\mathbf{r}}{r^3}\mathbf{v}\cdot\mathbf{r}\right) = 0. \tag{15.12}$$

With

$$\frac{d}{dt}\left(\mathbf{v}\times\mathbf{M} - \alpha\hat{\mathbf{r}}\right) = 0, \tag{15.13}$$

our vector must be some constant vector. Let us write this

$$\mathbf{v}\times\mathbf{M} - \alpha\hat{\mathbf{r}} = \alpha\mathbf{e}, \tag{15.14}$$

so that

$$\boxed{\mathbf{v}\times\mathbf{M} = \alpha\left(\mathbf{e}+\hat{\mathbf{r}}\right).} \tag{15.15}$$

Dotting eq. (15.15) with $\mathbf{M}$ we find

$$\begin{aligned}
\alpha\mathbf{M}\cdot\left(\mathbf{e}+\hat{\mathbf{r}}\right) &= \mathbf{M}\cdot(\mathbf{v}\times\mathbf{M}) \\
&= 0.
\end{aligned} \tag{15.16}$$

With $\hat{\mathbf{r}}$ lying in the plane of the trajectory (perpendicular to $\mathbf{M}$), we must also have $\mathbf{e}$ lying in the plane of the trajectory. Now we can dot eq. (15.15) with $\mathbf{r}$ to find

$$\begin{aligned}
\mathbf{r}\cdot(\mathbf{v}\times\mathbf{M}) &= \alpha\mathbf{r}\cdot\left(\mathbf{e}+\hat{\mathbf{r}}\right) \\
&= \alpha\left(re\cos(\phi-\phi_0)+r\right) \\
\mathbf{M}\cdot(\mathbf{r}\times\mathbf{v}) &= \\
\mathbf{M}\cdot\frac{\mathbf{M}}{m} &= \\
\frac{\mathbf{M}^2}{m} &=
\end{aligned} \tag{15.17}$$

This is

$$\frac{\mathbf{M}^2}{m} = \alpha r\left(1+e\cos(\phi-\phi_0)\right). \tag{15.18}$$

This is a kind of curious implicit relationship, since ϕ is also a function of r. Recall that the kinetic portion of our Lagrangian was

$$\frac{1}{2}m\left(\dot{r}^2+r^2\dot{\phi}^2\right), \tag{15.19}$$

so that our angular momentum was

$$M_\phi = \frac{\partial}{\partial\dot{\phi}}\left(\frac{1}{2}mr^2\dot{\phi}^2\right) = mr^2\dot{\phi}, \tag{15.20}$$

with no ϕ dependence in the Lagrangian we have

$$\frac{d}{dt}\left(mr^2\dot{\phi}\right) = 0, \tag{15.21}$$

or

$$\mathbf{M} = mr^2\dot{\phi}\hat{\mathbf{z}} = \text{constant}. \tag{15.22}$$

Our dynamics are now fully specified, even if this not completely explicit

$$\boxed{\begin{aligned}
r &= \frac{M^2}{m\alpha}\frac{1}{1+e\cos(\phi-\phi_0)} \\
\frac{d\phi}{dt} &= \frac{M}{mr^2}.
\end{aligned}} \tag{15.23}$$

What we can do is rearrange and separate variables

$$\frac{1}{r^2} = \frac{m^2\alpha^2}{M^4}(1+e\cos(\phi-\phi_0))^2 = \frac{m}{M}\frac{d\phi}{dt}, \tag{15.24}$$

to find

$$t - t_0 = \frac{M^3}{m\alpha^3}\int_{\phi_0}^{\phi} d\phi \frac{1}{(1+e\cos(\phi-\phi_0))^2} = \frac{M^3}{m\alpha^3}\int_0^{\phi-\phi_0} du \frac{1}{(1+e\cos u)^2}. \tag{15.25}$$

Now, at least $\phi = \phi(t)$ is specified implicitly.

We can also use the first of these to determine the magnitude of the radial velocity

$$\begin{aligned}
\frac{dr}{dt} &= -\frac{M^2}{m\alpha}\frac{1}{(1+e\cos(\phi-\phi_0))^2}(-e\sin(\phi-\phi_0))\frac{d\phi}{dt} \\
&= \frac{eM^2}{m\alpha}\frac{1}{(1+e\cos(\phi-\phi_0))^2}\sin(\phi-\phi_0)\frac{M}{mr^2} \\
&= \frac{eM^3}{m^2\alpha r^2}\frac{1}{(1+e\cos(\phi-\phi_0))^2}\sin(\phi-\phi_0) \\
&= \frac{eM^3}{m^2\alpha r^2}\left(\frac{mr\alpha}{M^2}\right)^2\sin(\phi-\phi_0) \\
&= \frac{e}{M}\sin(\phi-\phi_0),
\end{aligned} \tag{15.26}$$

with this, we can also find the energy

$$\begin{aligned}
E &= \dot{r}(m\dot{r}) + \dot{\phi}\left(mr^2\dot{\phi}\right) - \left(\frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 - U\right) \\
&= \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 + U \\
&= \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 - \frac{\alpha}{r} \\
&= \frac{1}{2}m\frac{e^2}{M^2}\sin^2(\phi-\phi_0) + \frac{1}{2mr^2}M^2 - \frac{\alpha}{r}.
\end{aligned} \tag{15.27}$$

Or

$$E = \frac{m}{2M^2}(\mathbf{e}\times\hat{\mathbf{r}})^2 + \frac{1}{2mr^2}M^2 - \frac{\alpha}{r}. \tag{15.28}$$

Is this what was used in class to state the relation

$$e = \sqrt{1 + \frac{2EM^2}{m\alpha^2}}. \tag{15.29}$$

It is not obvious exactly how that is obtained, but we can go back to eq. (15.23) to eliminate the $e^2\sin^2\Delta\phi$ term

$$E = \frac{1}{2}m\frac{1}{M^2}\left(e^2 - \left(\frac{M^2}{rm\alpha} - 1\right)^2\right) + \frac{1}{2mr^2}M^2 - \frac{\alpha}{r}. \tag{15.30}$$

Presumably this simplifies to the desired result (or there is other errors made in that prevent that).

15.2 SOLUTIONS.

16 FIELD LAGRANGIANS.

This chapter will cover

- Derivation of the relativistic form of the Euler-Lagrange field equations from the covariant form of the action,
- Derivation of Maxwell's equation (in it's Space Time Algebra (STA) form) from the Maxwell Lagrangian,
- Relationship of the STA Maxwell Lagrangian to the tensor equivalent,
- Relationship of the STA form of Maxwell's equation to it's tensor equivalents,
- Relationship of the STA Maxwell's equation to it's conventional Gibbs form.
- Show that we may use a multivector valued Lagrangian with all of F^2, not just the scalar part.

It is assumed that the reader is thoroughly familiar with the STA formalism, and if that is not the case, there is no better reference than [2].

16.1 FIELD ACTION.

Theorem 16.1: Relativistic Euler-Lagrange field equations.

Let $\phi \to \phi + \delta\phi$ be any variation of the field, such that the variation $\delta\phi = 0$ vanishes at the boundaries of the action integral

$$S = \int d^4x \mathcal{L}(\phi, \partial_\nu \phi).$$

The extreme value of the action is found when the Euler-Lagrange equations

$$0 = \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)},$$

are satisfied. For a Lagrangian with multiple field variables, there will be one such equation for each field.

Proof. To ease the visual burden, designate the variation of the field by $\delta\phi = \epsilon$, and perform a first order expansion of the varied Lagrangian

$$\begin{aligned} \mathcal{L} &\to \mathcal{L}(\phi + \epsilon, \partial_\nu(\phi + \epsilon)) \\ &= \mathcal{L}(\phi, \partial_\nu \phi) + \frac{\partial \mathcal{L}}{\partial \phi}\epsilon + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}\partial_\nu \epsilon. \end{aligned} \tag{16.1}$$

The variation of the Lagrangian is

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \phi}\epsilon + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}\partial_\nu \epsilon \\ &= \frac{\partial \mathcal{L}}{\partial \phi}\epsilon + \partial_\nu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}\epsilon\right) - \epsilon\partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}, \end{aligned} \tag{16.2}$$

which we may plug into the action integral to find

$$\delta S = \int d^4x \epsilon \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}\right) + \int d^4x \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}\epsilon\right). \tag{16.3}$$

The last integral can be evaluated along the dx^ν direction, leaving

$$\int d^3x \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}\epsilon\Big|_{\Delta x^\nu}, \tag{16.4}$$

where $d^3x = dx^\alpha dx^\beta dx^\gamma$ is the product of differentials that does not include dx^ν. By construction, ϵ vanishes on the boundary of the action integral so eq. (16.4) is zero. The action takes its extreme value when

$$\begin{aligned} 0 &= \delta S \\ &= \int d^4x \epsilon \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}\right). \end{aligned} \tag{16.5}$$

The proof is complete after noting that this must hold for all variations of the field ϵ, which means that we must have

$$0 = \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}. \tag{16.6}$$

□

16.2 MAXWELL'S EQUATION.

Armed with the Euler-Lagrange equations, we can apply them to the Maxwell's equation Lagrangian, which we will claim has the following form.

Theorem 16.2: Maxwell's equation Lagrangian.

Application of the Euler-Lagrange equations to the Lagrangian

$$\mathcal{L} = -\frac{\epsilon_0 c}{2} F \cdot F + J \cdot A,$$

where $F = \nabla \wedge A$, yields the vector portion of Maxwell's equation

$$\nabla \cdot F = \frac{1}{\epsilon_0 c} J,$$

which implies

$$\nabla F = \frac{1}{\epsilon_0 c} J.$$

This is Maxwell's equation.

Proof. We wish to apply all of the Euler-Lagrange equations simultaneously (i.e. once for each of the four A_μ components of the potential), and cast it into four-vector form

$$0 = \gamma_\nu \left(\frac{\partial}{\partial A_\nu} - \partial_\mu \frac{\partial}{\partial(\partial_\mu A_\nu)} \right) \mathcal{L}. \tag{16.7}$$

Since our Lagrangian splits nicely into kinetic and interaction terms, this gives us

$$0 = \gamma_\nu \left(\frac{\partial(A \cdot J)}{\partial A_\nu} + \frac{\epsilon_0 c}{2} \partial_\mu \frac{\partial(F \cdot F)}{\partial(\partial_\mu A_\nu)} \right). \tag{16.8}$$

The interaction term above is just

$$\gamma_\nu \frac{\partial(A \cdot J)}{\partial A_\nu} = \gamma_\nu \frac{\partial(A_\mu J^\mu)}{\partial A_\nu} = \gamma_\nu J^\nu = J, \tag{16.9}$$

but the kinetic term takes a bit more work. Let's start with evaluating

$$\begin{aligned}
\frac{\partial(F\cdot F)}{\partial(\partial_\mu A_\nu)} &= \frac{\partial F}{\partial(\partial_\mu A_\nu)}\cdot F + F\cdot\frac{\partial F}{\partial(\partial_\mu A_\nu)} \\
&= 2\frac{\partial F}{\partial(\partial_\mu A_\nu)}\cdot F \\
&= 2\frac{\partial(\partial_\alpha A_\beta)}{\partial(\partial_\mu A_\nu)}\left(\gamma^\alpha\wedge\gamma^\beta\right)\cdot F \\
&= 2\left(\gamma^\mu\wedge\gamma^\nu\right)\cdot F.
\end{aligned} \tag{16.10}$$

We hit this with the μ-partial and expand as a scalar selection to find

$$\begin{aligned}
\partial_\mu\frac{\partial(F\cdot F)}{\partial(\partial_\mu A_\nu)} &= 2\left(\partial_\mu\gamma^\mu\wedge\gamma^\nu\right)\cdot F \\
&= -2(\gamma^\nu\wedge\nabla)\cdot F \\
&= -2\langle(\gamma^\nu\wedge\nabla)F\rangle \\
&= -2\langle\gamma^\nu\nabla F - \cancel{\gamma^\nu\cdot\nabla F}\rangle \\
&= -2\gamma^\nu\cdot\left(\nabla\cdot F\right).
\end{aligned} \tag{16.11}$$

Putting all the pieces together yields

$$\begin{aligned}
0 &= J - \epsilon_0 c\gamma_\nu\left(\gamma^\nu\cdot\left(\nabla\cdot F\right)\right) \\
&= J - \epsilon_0 c\left(\nabla\cdot F\right),
\end{aligned} \tag{16.12}$$

but

$$\begin{aligned}
\nabla\cdot F &= \nabla F - \nabla\wedge F \\
&= \nabla F - \nabla\wedge(\nabla\wedge A) \\
&= \nabla F,
\end{aligned} \tag{16.13}$$

so the multivector field equations for this Lagrangian are

$$\nabla F = \frac{1}{\epsilon_0 c}J, \tag{16.14}$$

as claimed. □

Exercise 16.1 Tensor formalism.

Cast the Lagrangian of theorem 16.2 into the conventional tensor form

$$\mathcal{L} = \frac{\epsilon_0 c}{4}F_{\mu\nu}F^{\mu\nu} + A^\mu J_\mu. \tag{16.15}$$

Also show that the four-vector component of Maxwell's equation $\nabla \cdot F = J/(\epsilon_0 c)$ is equivalent to the conventional tensor form of the Gauss-Ampere law

$$\partial_\mu F^{\mu\nu} = \frac{1}{\epsilon_0 c} J^\nu, \tag{16.16}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ as usual. Also show that the trivector component of Maxwell's equation $\nabla \wedge F = 0$ is equivalent to the tensor form of the Gauss-Faraday law

$$\partial_\alpha \left(\epsilon^{\alpha\beta\mu\nu} F_{\mu\nu} \right) = 0. \tag{16.17}$$

Exercise 16.2 Tensor and Gibbs forms of Maxwell's equations.

Using the tensor identifications of eq. (7.43) and

$$J^\mu = \left(c\rho, \mathbf{J} \right), \tag{16.26}$$

the reader should satisfy themselves that the traditional Gibbs form of Maxwell's equations can be recovered from eq. (16.16).

Exercise 16.3 Grad and curl form of Maxwell's equations.

With $J = c\rho\gamma_0 + J^k \gamma_k$ and $F = \mathbf{E} + Ic\mathbf{B}$ show that Maxwell's equation, as stated in theorem 16.2 expand to the conventional div and curl expressions for Maxwell's equations.

Exercise 16.4 Alternative multivector Lagrangian.

Show that a scalar+pseudoscalar Lagrangian of the following form

$$\mathcal{L} = -\frac{\epsilon_0 c}{2} F^2 + J \cdot A,$$

which omits the scalar selection of the Lagrangian in theorem 16.2, also represents Maxwell's equation. Discuss the scalar and pseudoscalar components of F^2, and show why the pseudoscalar inclusion is irrelevant.

16.3 SOLUTIONS.

Answer for Exercise 16.1

To show the Lagrangian correspondence we must expand $F \cdot F$ in coordinates

$$
\begin{aligned}
F \cdot F &= (\nabla \wedge A) \cdot (\nabla \wedge A) \\
&= \left((\gamma^\mu \partial_\mu) \wedge (\gamma^\nu A_\nu)\right) \cdot \left((\gamma^\alpha \partial_\alpha) \wedge (\gamma^\beta A_\beta)\right) \\
&= \left(\gamma^\mu \wedge \gamma^\nu\right) \cdot \left(\gamma_\alpha \wedge \gamma_\beta\right) (\partial_\mu A_\nu)(\partial^\alpha A^\beta) \\
&= \left(\delta^\mu{}_\beta \delta^\nu{}_\alpha - \delta^\mu{}_\alpha \delta^\nu{}_\beta\right) (\partial_\mu A_\nu)(\partial^\alpha A^\beta) \\
&= -\partial_\mu A_\nu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) \\
&= -\partial_\mu A_\nu F^{\mu\nu} \\
&= -\frac{1}{2} \left(\partial_\mu A_\nu F^{\mu\nu} + \partial_\nu A_\mu F^{\nu\mu}\right) \\
&= -\frac{1}{2} \left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) F^{\mu\nu} \\
&= -\frac{1}{2} F_{\mu\nu} F^{\mu\nu}.
\end{aligned}
\tag{16.18}
$$

With a substitution of this and $A \cdot J = A_\mu J^\mu$ back into the Lagrangian, we recover the tensor form of the Lagrangian.

To recover the tensor form of Maxwell's equation, we first split it into vector and trivector parts

$$
\nabla \cdot F + \nabla \wedge F = \frac{1}{\epsilon_0 c} J. \tag{16.19}
$$

Now the vector component may be expanded in coordinates by dotting both sides with γ^ν to find

$$
\frac{1}{\epsilon_0 c} \gamma^\nu \cdot J = J^\nu, \tag{16.20}
$$

and

$$
\begin{aligned}
\gamma^\nu \cdot (\nabla \cdot F) &= \partial_\mu \gamma^\nu \cdot \left(\gamma^\mu \cdot \left(\gamma_\alpha \wedge \gamma_\beta\right) \partial^\alpha A^\beta\right) \\
&= \left(\delta^\mu{}_\alpha \delta^\nu{}_\beta - \delta^\nu{}_\alpha \delta^\mu{}_\beta\right) \partial_\mu \partial^\alpha A^\beta \\
&= \partial_\mu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) \\
&= \partial_\mu F^{\mu\nu}.
\end{aligned}
\tag{16.21}
$$

Equating eq. (16.20) and eq. (16.21) finishes the first part of the job. For the trivector component, we have

$$\begin{aligned} 0 &= \nabla \wedge F \\ &= (\gamma^\mu \partial_\mu) \wedge \left(\gamma^\alpha \wedge \gamma^\beta\right) \partial_\alpha A_\beta \\ &= \frac{1}{2}(\gamma^\mu \partial_\mu) \wedge \left(\gamma^\alpha \wedge \gamma^\beta\right) F_{\alpha\beta}. \end{aligned} \tag{16.22}$$

Wedging with γ^τ and then multiplying by $-2I$ we find

$$0 = -\left(\gamma^\mu \wedge \gamma^\alpha \wedge \gamma^\beta \wedge \gamma^\tau\right) I \partial_\mu F_{\alpha\beta}, \tag{16.23}$$

but

$$\gamma^\mu \wedge \gamma^\alpha \wedge \gamma^\beta \wedge \gamma^\tau = -I\epsilon^{\mu\alpha\beta\tau}, \tag{16.24}$$

which leaves us with

$$\epsilon^{\mu\alpha\beta\tau} \partial_\mu F_{\alpha\beta} = 0, \tag{16.25}$$

as expected.

Answer for Exercise 16.2

The reader is referred to Exercise 3.4 "Electrodynamics, variational principle." from [14].

Answer for Exercise 16.3

To obtain Maxwell's equations in their traditional vector forms, we pre-multiply both sides with γ_0

$$\gamma_0 \nabla F = \frac{1}{\epsilon_0 c} \gamma_0 J, \tag{16.27}$$

and then select each grade separately. First observe that the RHS above has scalar and bivector components, as

$$\gamma_0 J = c\rho + J^k \gamma_0 \gamma_k. \tag{16.28}$$

In terms of the spatial bivector basis $\mathbf{e}_k = \gamma_k \gamma_0$, the RHS of eq. (16.27) is

$$\gamma_0 \frac{J}{\epsilon_0 c} = \frac{\rho}{\epsilon_0} - \mu_0 c \mathbf{J}. \tag{16.29}$$

For the LHS, first note that

$$\begin{aligned}\gamma_0\nabla &= \gamma_0\left(\gamma_0\partial^0 + \gamma_k\partial^k\right)\\ &= \partial_0 - \gamma_0\gamma_k\partial_k\\ &= \frac{1}{c}\frac{\partial}{\partial t} + \boldsymbol{\nabla}.\end{aligned} \tag{16.30}$$

We can express all the the LHS of eq. (16.27) in the bivector spatial basis, so that Maxwell's equation in multivector form is

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \boldsymbol{\nabla}\right)(\mathbf{E} + Ic\mathbf{B}) = \frac{\rho}{\epsilon_0} - \mu_0 c\mathbf{J}. \tag{16.31}$$

Selecting the scalar, vector, bivector, and trivector grades of both sides (in the spatial basis) gives the following set of respective equations

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \frac{\rho}{\epsilon_0} \tag{16.32a}$$

$$\frac{1}{c}\partial_t\mathbf{E} + Ic\boldsymbol{\nabla}\wedge\mathbf{B} = -\mu_0 c\mathbf{J} \tag{16.32b}$$

$$\boldsymbol{\nabla}\wedge\mathbf{E} + I\partial_t\mathbf{B} = 0 \tag{16.32c}$$

$$Ic\boldsymbol{\nabla}\cdot B = 0, \tag{16.32d}$$

which we can rewrite after some duality transformations (and noting that $\mu_0\epsilon_0 c^2 = 1$), we have

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \frac{\rho}{\epsilon_0} \tag{16.33a}$$

$$\boldsymbol{\nabla}\times\mathbf{B} - \mu_0\epsilon_0\frac{\partial\mathbf{E}}{\partial t} = \mu_0\mathbf{J} \tag{16.33b}$$

$$\boldsymbol{\nabla}\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = 0 \tag{16.33c}$$

$$\boldsymbol{\nabla}\cdot B = 0, \tag{16.33d}$$

which are Maxwell's equations in their traditional form.

Answer for Exercise 16.4

The quantity $F^2 = F \cdot F + F \wedge F$ has both scalar and pseudoscalar [1] components, which can be seen if we expand it in terms of the electric and magnetic fields

$$\begin{aligned} F^2 &= (\mathbf{E} + Ic\mathbf{B})^2 \\ &= \mathbf{E}^2 - c^2\mathbf{B}^2 + Ic\left(\mathbf{E}\mathbf{B} + \mathbf{B}\mathbf{E}\right) \\ &= \mathbf{E}^2 - c^2\mathbf{B}^2 + 2Ic\mathbf{E}\cdot\mathbf{B}. \end{aligned} \tag{16.34}$$

Both the scalar and pseudoscalar parts of F^2 are Lorentz invariant, a requirement of our Lagrangian, but most Maxwell equation Lagrangians only include the scalar $\mathbf{E}^2 - c^2\mathbf{B}^2$ component of the field square. If we allow the Lagrangian to be multivector valued, and evaluate the Euler-Lagrange equations, we quickly find the same results

$$\begin{aligned} 0 &= \gamma_\nu \left(\frac{\partial}{\partial A_\nu} - \partial_\mu \frac{\partial}{\partial(\partial_\mu A_\nu)} \right) \mathcal{L} \\ &= \gamma_\nu \left(J^\nu + \frac{\epsilon_0 c}{2} \partial_\mu \left((\gamma^\mu \wedge \gamma^\nu) F + F(\gamma^\mu \wedge \gamma^\nu) \right) \right). \end{aligned} \tag{16.35}$$

Here some steps are skipped, building on our previous scalar Euler-Lagrange evaluation experience. We have a symmetric product of two bivectors, which we can express as a 0,4 grade selection, since

$$\langle XF \rangle_{0,4} = \frac{1}{2}\left(XF + FX\right), \tag{16.36}$$

for any two bivectors X, F. This leaves

$$\begin{aligned} 0 &= J + \epsilon_0 c \gamma_\nu \langle (\nabla \wedge \gamma^\nu) F \rangle_{0,4} \\ &= J + \epsilon_0 c \gamma_\nu \langle -\gamma^\nu \nabla F + \cancel{(\gamma^\nu \cdot \nabla) F} \rangle_{0,4} \\ &= J - \epsilon_0 c \gamma_\nu \left(\gamma^\nu \cdot (\nabla \cdot F) + \gamma^\nu \wedge \nabla \wedge F \right). \end{aligned} \tag{16.37}$$

However, since $\nabla \wedge F = \nabla \wedge \nabla \wedge A = 0$, we see that there is no contribution from the $F \wedge F$ pseudoscalar component of the Lagrangian, and we are left with

$$\begin{aligned} 0 &= J - \epsilon_0 c (\nabla \cdot F) \\ &= J - \epsilon_0 c \nabla F, \end{aligned} \tag{16.38}$$

which is Maxwell's equation, as before.

1 Unlike vectors, a bivector wedge in 4D with itself need not be zero (example: $\gamma_0\gamma_1 + \gamma_2\gamma_3$ wedged with itself).

17

REVIEW: FIELD LAGRANGIANS.

17.0.0.1 *Schrödinger's equation*

Problem 11.3 in [4] is to take the Lagrangian

$$\begin{aligned}\mathcal{L} &= \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi^* + V\psi\psi^* + \frac{\hbar}{2i}\left(\psi^*\partial_t\psi - \psi\partial_t\psi^*\right)\\ &= \frac{\hbar^2}{2m}\partial_k\psi\partial_k\psi^* + V\psi\psi^* + \frac{\hbar}{2i}\left(\psi^*\partial_t\psi - \psi\partial_t\psi^*\right).\end{aligned} \tag{17.1}$$

treating ψ, and ψ^* as separate fields and show that Schrödinger's equation and its conjugate follows. (note: I have added a 1/2 fact in the commutator term that was not in the Goldstein problem. Believe that to have been a typo in the original (first edition)). We have

$$\frac{\partial\mathcal{L}}{\partial\psi^*} = V\psi + \frac{\hbar}{2i}\partial_t\psi. \tag{17.2}$$

and canonical momenta

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial(\partial_m\psi^*)} &= \frac{\hbar^2}{2m}\partial_m\psi\\ \frac{\partial\mathcal{L}}{\partial(\partial_t\psi^*)} &= -\frac{\hbar}{2i}\psi.\end{aligned} \tag{17.3}$$

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\psi^*} &= \sum_m \partial_m\frac{\partial\mathcal{L}}{\partial(\partial_m\psi^*)} + \partial_t\frac{\partial\mathcal{L}}{\partial(\partial_t\psi^*)}\\ V\psi + \frac{\hbar}{2i}\partial_t\psi &= \frac{\hbar^2}{2m}\sum_m\partial_{mm}\psi - \frac{\hbar}{2i}\frac{\partial\psi}{\partial t}.\end{aligned} \tag{17.4}$$

which is the desired result

$$-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi + V\psi = \hbar i\frac{\partial\psi}{\partial t}. \tag{17.5}$$

The conjugate result

$$-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi^* + V\psi^* = -\hbar i\frac{\partial\psi^*}{\partial t}. \tag{17.6}$$

follows by inspection since all terms except the time partial are symmetric in ψ and ψ^*. The time partial has a negation in sign from the commutator of the Lagrangian.

FIXME: Goldstein also wanted the Hamiltonian, but I do not know what that is yet. Got to go read the earlier parts of the book!

17.0.0.2 *Relativistic Schrödinger's equation*

The wiki article on Noether's theorem lists the relativistic quantum Lagrangian in the form

$$\mathcal{L} = -\eta^{\mu\nu}\partial_\mu\psi\partial_\nu\psi^* + \frac{m^2c^2}{\hbar^2}\psi\psi^*. \tag{17.7}$$

That article uses $\hbar = c = 1$, and appears to use a $- + ++$ metric, both of which are adjusted for here. Calculating the derivatives

$$\frac{\partial\mathcal{L}}{\partial\psi^*} = \frac{m^2c^2}{\hbar^2}\psi. \tag{17.8}$$

$$\begin{aligned}\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi^*)} &= -\partial_\mu\left(\eta^{\alpha\beta}\partial_\alpha\psi\frac{\partial}{\partial(\partial_\mu\psi^*)}\partial_\beta\psi^*\right)\\ &= -\partial_\mu\left(\eta^{\alpha\mu}\partial_\alpha\psi\right)\\ &= -\partial_\mu\partial^\mu\psi.\end{aligned} \tag{17.9}$$

So we have

$$\partial_\mu\partial^\mu\psi = \frac{-m^2c^2}{\hbar^2}\psi. \tag{17.10}$$

With the metric dependency made explicit this is

$$\left(\boldsymbol{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{(\partial t)^2}\right)\psi = \frac{m^2c^2}{\hbar^2}\psi. \tag{17.11}$$

Much different looking than the classical time dependent Schrödinger's equation in eq. (17.6). [23] has a nice discussion about this equation and its relation to the non-relativistic Schrödinger's equation.

Exercise 17.1 One dimensional wave equation.

The Lagrangian for a one dimensional wave is derived in [4] using a limiting argument applied to an infinite sequence of connected masses on springs. The result is

$$\mathcal{L} = \frac{1}{2}\left(\mu\left(\frac{\partial\eta}{\partial t}\right)^2 - Y\left(\frac{\partial\eta}{\partial x}\right)^2\right). \tag{17.12}$$

Here η was the displacement from the equilibrium position, μ is the mass line density and Y is Young's modulus.

Using this Lagrangian, find the equations of the field, showing that it has the expected form.

Exercise 17.2 Wave equation in higher dimensions.

For a string or film or other wavy material with more degrees of freedom than a string with back and forth motion, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}\left(\mu\left(\frac{\partial\eta}{\partial t}\right)^2 - Y\sum_i\left(\frac{\partial\eta}{\partial x^i}\right)^2\right). \tag{17.15}$$

Evaluate the equations for the field.

Exercise 17.3 Non-relativistic QM Lagrangian.

The non-relativistic Lagrangian given by [4] (pr. 11.3) is

$$\mathcal{L} = \frac{\hbar^2}{2m}(\boldsymbol{\nabla}\psi)\cdot(\boldsymbol{\nabla}\psi^*) + V\psi\psi^* + i\hbar\left(\psi\partial_t\psi^* - \psi^*\partial_t\psi\right). \tag{17.18}$$

Show that

$$\left(\frac{-\hbar^2}{2m}\boldsymbol{\nabla}^2 + V\right)\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{17.19}$$

Exercise 17.4 Klein-Gordon Lagrangian.

The Klein-Gordon Lagrangian is

$$\mathcal{L} = -(\nabla\psi)\cdot(\nabla\psi^*) + \frac{m^2c^2}{\hbar^2}\psi\psi^*. \tag{17.20}$$

Evaluate the Euler-Lagrange equations to show that this describes the Klein-Gordon scalar wave equation

$$\left(\frac{\hbar^2}{2m}\nabla^2 + \frac{1}{2}mc^2\right)\psi = 0. \tag{17.21}$$

Exercise 17.5 Dirac equation.

The Lagrangian for the Dirac equation is

$$\mathcal{L} = mc\bar{\psi}\psi - \frac{1}{2}i\hbar(\bar{\psi}\gamma^\mu(\partial_\mu\psi) - (\partial_\mu\bar{\psi})\gamma^\mu\psi). \tag{17.22}$$

Where $\bar{\psi} = \gamma_0\tilde{\psi}$, and $\tilde{\psi}$ is the reversed field spinor. Show that evaluating the Euler-Lagrange equations yield

$$i\hbar\nabla\psi = \pm mc\psi. \tag{17.23}$$

17.1 SOLUTIONS.

Answer for Exercise 17.1

Taking derivatives confirms that this is the correct form. The Euler-Lagrange equations for this equation are:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\eta} &= \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\frac{\partial\eta}{\partial t}} + \frac{\partial}{\partial x}\frac{\partial\mathcal{L}}{\partial\frac{\partial\eta}{\partial x}} \\ 0 &= \frac{\partial}{\partial t}\mu\frac{\partial\eta}{\partial t} - \frac{\partial}{\partial x}Y\frac{\partial\eta}{\partial x}.\end{aligned} \tag{17.13}$$

Which has the expected form

$$\mu\frac{\partial^2\eta}{(\partial t)^2} - Y\frac{\partial^2\eta}{(\partial x)^2} = 0. \tag{17.14}$$

Answer for Exercise 17.2

Calculating the Euler-Lagrange equations gives

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\eta} &= \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\frac{\partial\eta}{\partial t}} + \sum_i\frac{\partial}{\partial x^i}\frac{\partial\mathcal{L}}{\partial\frac{\partial\eta}{\partial x^i}} \\ 0 &= \frac{\partial}{\partial t}\mu\frac{\partial\eta}{\partial t} - \sum_i\frac{\partial}{\partial x^i}Y\frac{\partial\eta}{\partial x^i}.\end{aligned} \tag{17.16}$$

This has the expected form

$$\mu\frac{\partial^2\eta}{(\partial t)^2} - Y\sum_i\frac{\partial^2\eta}{(\partial x^i)^2} = 0. \tag{17.17}$$

18 NOETHER'S THEOREM FOR FIELDS.

18.1 NOETHER'S THEOREM.

18.1.1 *Derivation.*

It was seen in 8 that Noether's law for a line integral action was shown to essentially be an application of the chain rule, coupled with an application of the Euler-Lagrange equations.

For a field Lagrangian a similar conservation statement can be made, where it takes the form of a divergence relationship instead of derivative with respect to the integration parameter associated with the line integral.

The following derivation follows [2], but is dumbed down to the scalar field variable case, and additional details are added.

The Lagrangian to be considered is

$$\mathcal{L} = \mathcal{L}(\psi, \partial_\mu \psi), \tag{18.1}$$

and the single field case is sufficient to see how this works. Consider the following transformation:

$$\begin{aligned} \psi &\to f(\psi, \alpha) = \psi' \\ \mathcal{L}' &= \mathcal{L}(f, \partial_\mu f). \end{aligned} \tag{18.2}$$

Taking derivatives of the transformed Lagrangian with respect to the free transformation variable α, we have

$$\frac{d\mathcal{L}'}{d\alpha} = \frac{\partial \mathcal{L}}{\partial f}\frac{\partial f}{\partial \alpha} + \sum_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}\frac{\partial(\partial_\mu f)}{\partial \alpha}. \tag{18.3}$$

The Euler-Lagrange field equations for the transformed Lagrangian are

$$\frac{\partial \mathcal{L}}{\partial f} = \sum_\mu \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}. \tag{18.4}$$

For some for background discussion, examples, and derivation of the field form of Noether's equation see **??**. Now substitute back into eq. (18.3) for

$$\begin{aligned}\frac{d\mathcal{L}'}{d\alpha} &= \sum_\mu \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}\right)\frac{\partial f}{\partial \alpha} + \sum_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}\frac{\partial(\partial_\mu f)}{\partial \alpha} \\ &= \sum_\mu \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}\right)\frac{\partial f}{\partial \alpha} + \sum_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}\partial_\mu \frac{\partial f}{\partial \alpha}.\end{aligned} \tag{18.5}$$

Using the product rule we have

$$\begin{aligned}\frac{d\mathcal{L}'}{d\alpha} &= \sum_\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}\frac{\partial f}{\partial \alpha}\right) \\ &= \sum_\mu \gamma^\mu \partial_\mu \cdot \left(\gamma_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f)}\frac{\partial f}{\partial \alpha}\right) \\ &= \nabla \cdot \left(\gamma_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi')}\frac{\partial \psi'}{\partial \alpha}\right).\end{aligned} \tag{18.6}$$

Here the field does not have to be a relativistic field which could be implied by the use of the standard symbols for relativistic four vector basis $\{\gamma_\mu\}$ of STA. This is really a statement that one can form a gradient in the field variable configuration space using any appropriate reciprocal basis pair.

Noether's law for a field Lagrangian is a statement that if the transformed Lagrangian is unchanged (invariant) by some type of parametrized field variable transformation, then with $J' = J'^\mu \gamma_\mu$ one has

$$\frac{d\mathcal{L}'}{d\alpha} = \nabla \cdot J' = 0. \tag{18.7a}$$

$$J'^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi')}\frac{\partial \psi'}{\partial \alpha}. \tag{18.7b}$$

FIXME: GAFP evaluates things at $\alpha = 0$ where that is the identity case. I think this is what allows them to drop the primes later. Must think this through.

18.1.2 *Examples.*

18.1.2.1 *Klein-Gordan Lagrangian invariance under phase change.*

The Klein-Gordan Lagrangian, a relativistic relative of the Schrödinger equation is

$$\mathcal{L} = \eta^{\mu\nu}\partial_\mu\psi\partial_\nu\psi^* - m^2\psi\psi^*. \tag{18.8}$$

FIXME: fixed sign above. Adjust the remainder below. This provides a simple example application of the field form of Noether's equation, for a transformation that involves a phase change

$$\begin{aligned}\psi &\to \psi' = e^{i\theta}\psi \\ \psi^* &\to \psi^{*\prime} = e^{-i\theta}\psi^*.\end{aligned} \tag{18.9}$$

This transformation leaves the Lagrangian unchanged, so there is an associated conserved quantity.

$$\begin{aligned}\frac{\partial\psi'}{\partial\theta} &= i\psi' \\ \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi')} &= \eta^{\mu\nu}\partial_\nu\psi'^* = \partial^\mu\psi'^*.\end{aligned} \tag{18.10}$$

Summing all the field partials, treating ψ, and ψ^* as separate field variables the divergence conservation statement is

$$\partial_\mu \overbrace{\boxed{\left(\partial^\mu\psi'^* i\psi' - \partial^\mu\psi' i\psi'^*\right)}}^{J'^\mu} = 0. \tag{18.11}$$

Dropping primes and writing $J = \gamma_\mu J^\mu$, this is

$$\begin{aligned}J &= i(\psi\nabla\psi^* - \psi^*\nabla\psi) \\ \nabla\cdot J &= 0.\end{aligned} \tag{18.12}$$

Apparently with charge added this quantity actually represents electric current density. It will be interesting to learn some quantum mechanics and see how this works.

18.1.2.2 *Lorentz boost and rotation invariance of Maxwell Lagrangian.*

$$\begin{aligned}\mathcal{L} &= -\left\langle (\nabla \wedge A)^2 \right\rangle + \kappa A \cdot J \\ &= \partial_\mu A_\nu (\partial^\mu A^\nu - \partial^\nu A^\mu) + \kappa A_\sigma J^\sigma .\end{aligned} \tag{18.13a}$$

$$\kappa = \frac{2}{\epsilon_0 c}. \tag{18.13b}$$

The rotation and boost invariance of the Maxwell Lagrangian was demonstrated in D. Following E write the Lorentz boost or rotation in exponential form.

$$L(x) = \exp(-\alpha i/2) x \exp(\alpha i/2), \quad \Lambda = \exp(-\alpha i/2). \tag{18.14}$$

where i is a unit spatial bivector for a rotation of $-\alpha$ radians, and a boost with rapidity α when i is a spacetime unit bivector.

Introducing the transformation

$$A \to A' = \Lambda A \Lambda^\dagger. \tag{18.15}$$

The change in A' with respect to α is

$$\frac{\partial A'}{\partial \alpha} = -iA' + A'i = 2A' \cdot i = 2A'_\sigma \gamma^\sigma \cdot i. \tag{18.16}$$

Next we want to compute

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial(\partial_\mu A'_\nu)} &= \frac{\partial}{\partial(\partial_\mu A'_\nu)} \left(\partial_\alpha A'_\beta (\partial^\alpha A'^\beta - \partial^\beta A'^\alpha) + \kappa A'_\sigma J^\sigma \right) \\ &= \left(\frac{\partial}{\partial(\partial_\mu A'_\nu)} \partial_\alpha A'_\beta \right) \left(\partial^\alpha A'^\beta - \partial^\beta A'^\alpha \right) \\ &+ \partial^\alpha A'^\beta \frac{\partial}{\partial(\partial_\mu A'_\nu)} \left(\partial_\alpha A'_\beta - \partial_\beta A'_\alpha \right)) \\ &= \left(\frac{\partial}{\partial(\partial_\mu A'_\nu)} \partial_\mu A'_\nu \right) (\partial^\mu A'^\nu - \partial^\nu A'^\mu) \\ &+ \partial^\mu A'^\nu \frac{\partial}{\partial(\partial_\mu A'_\nu)} \partial_\mu A'_\nu \\ &- \partial^\nu A'^\mu \frac{\partial}{\partial(\partial_\mu A'_\nu)} \partial_\mu A'_\nu \\ &= 2 \left(\partial^\mu A'^\nu - \partial^\nu A'^\mu \right) \\ &= 2F^{\mu\nu}.\end{aligned} \tag{18.17}$$

Employing the vector field form of Noether's equation as in eq. (18.37) the conserved current C components are

$$\begin{aligned} C^\mu &= 2(\gamma_\nu F^{\mu\nu}) \cdot (2A \cdot i) \\ &\propto (\gamma_\nu F^{\mu\nu}) \cdot (A \cdot i) \\ &\propto (\gamma^\mu \cdot F) \cdot (A \cdot i). \end{aligned} \tag{18.18}$$

Or

$$C = \gamma_\mu((\gamma^\mu \cdot F) \cdot (A \cdot i)). \tag{18.19}$$

Here C was used instead of J for the conserved current vector since J is already taken for the current charge density itself.

18.1.2.3 *Questions.*

FIXME: What is this quantity? It has the look of angular momentum, or torque, or an inertial tensor. Does it have a physical significance? Can the i be factored out of the expression, leaving a conserved quantity that is some linear function only of F, and A (this was possible in the Lorentz force Lagrangian for the same invariance considerations).

18.1.2.4 *Expansion for x-axis boost.*

As an example to get a feel for eq. (18.19), lets expand this for a specific spacetime boost plane. Using the x-axis that is $i = \gamma_1 \wedge \gamma_0$

First expanding the potential projection one has

$$\begin{aligned} A \cdot i &= (A_\mu \gamma^\mu) \cdot (\gamma_1 \wedge \gamma_0) \\ &= A_1 \gamma_0 - A_0 \gamma_1. \end{aligned} \tag{18.20}$$

Next the μ component of the field is

$$\begin{aligned} \gamma^\mu \cdot F &= \frac{1}{2} F^{\alpha\beta} \gamma^\mu \cdot (\gamma_\alpha \wedge \gamma_\beta) \\ &= \frac{1}{2} F^{\mu\beta} \gamma_\beta - \frac{1}{2} F^{\alpha\mu} \gamma_\alpha \\ &= F^{\mu\alpha} \gamma_\alpha. \end{aligned} \tag{18.21}$$

So the μ component of the conserved vector is

$$\begin{aligned} C^\mu &= (\gamma^\mu \cdot F) \cdot (A \cdot i) \\ &= (F^{\mu\alpha} \gamma_\alpha) \cdot (A_1 \gamma_0 - A_0 \gamma_1) \\ &= (F^{\mu\alpha} \gamma_\alpha) \cdot (A^0 \gamma^1 - A^1 \gamma^0) \end{aligned} \tag{18.22}$$

.

Therefore the conservation statement is

$$\begin{aligned} C^\mu &= F^{\mu 1} A^0 - F^{\mu 0} A^1 \\ \partial_\mu C^\mu &= 0. \end{aligned} \tag{18.23}$$

Let us write out the components of eq. (18.23) explicitly, to perhaps get a better feel for them.

$$\begin{aligned} C^0 &= F^{01} A^0 = -E_x \phi \\ C^1 &= -F^{10} A^1 = -E_x A_x \\ C^2 &= F^{21} A^0 - F^{20} A^1 = B_z \phi - E_y A_x \\ C^3 &= F^{31} A^0 - F^{30} A^1 = -B_y \phi - E_z A_x. \end{aligned} \tag{18.24}$$

Well, that is not particularly enlightening looking after all.

18.1.2.5 *Expansion for rotation or boost.*

Suppose that one takes $i = \gamma^\mu \wedge \gamma^\nu$, so that we have a symmetry for a boost if one of μ or ν is zero, and rotational symmetry otherwise.

This gives

$$\begin{aligned} A \cdot i &= (A^\alpha \gamma_\alpha) \cdot (\gamma^\mu \wedge \gamma^\nu) \\ &= A^\mu \gamma^\nu - A^\nu \gamma^\mu, \end{aligned} \tag{18.25a}$$

$$\begin{aligned} C^\alpha &= (\gamma^\alpha \cdot F) \cdot (A \cdot i) \\ &= (F^{\alpha\beta} \gamma_\beta) \cdot (A^\mu \gamma^\nu - A^\nu \gamma^\mu), \end{aligned} \tag{18.25b}$$

$$C^\alpha = F^{\alpha\nu} A^\mu - F^{\alpha\mu} A^\nu. \tag{18.25c}$$

For a rotation in the a, b, plane with $\mu = a$, and $\nu = b$ (say), lets write out the C^α components explicitly in terms of $\mathbf{E}$ and $\mathbf{B}$ components, also writing $0 < d$, $a \neq d \neq b$. That is

$$\begin{aligned} C^0 &= F^{0b} A^a - F^{0a} A^b = E^a A^b - E^b A^a \\ C^1 &= F^{1b} A^a - F^{1a} A^b \\ C^2 &= F^{2b} A^a - F^{2a} A^b \\ C^3 &= F^{3b} A^a - F^{3a} A^b. \end{aligned} \tag{18.26}$$

Only the first term of this reduces nicely. Suppose we additionally write $a = 1$, $b = 2$ to make things more concrete. Then we have

$$\begin{aligned}
C^0 &= F^{02}A^1 - F^{01}A^2 = E_x A_y - E_y A_x = (\mathbf{E}\times\mathbf{A})_z \\
C^1 &= F^{12}A^1 - F^{11}A^2 = -B_z A_x \\
C^2 &= F^{22}A^1 - F^{21}A^2 = B_z A_x \\
C^3 &= F^{32}A^1 - F^{31}A^2 = B_x A_x + B_y A_y.
\end{aligned} \tag{18.27}$$

The time-like component of whatever this vector is the z component of a cross product (spatial component of the $\mathbf{E}\times\mathbf{A}$ product in the direction of the normal to the rotational plane), but what is the rest?

18.1.2.6 *Conservation statement.*

Returning to eq. (18.25c), the conservation statement can be calculated as

$$\begin{aligned}
0 &= \partial_\alpha C^\alpha \\
&= \partial_\alpha F^{\alpha\nu}A^\mu - \partial_\alpha F^{\alpha\mu}A^\nu + F^{\alpha\nu}\partial_\alpha A^\mu - F^{\alpha\mu}\partial_\alpha A^\nu.
\end{aligned} \tag{18.28}$$

But the grade one terms of the Maxwell equation in tensor form is

$$\partial_\mu F^{\mu\alpha} = J^\alpha/\epsilon_0 c. \tag{18.29}$$

So we have

$$\begin{aligned}
0 &= \frac{1}{\epsilon_0 c}\left(J^\nu A^\mu - J^\mu A^\nu\right) + F_\alpha{}^\nu \partial^\alpha A^\mu - F_\alpha{}^\mu \partial^\alpha A^\nu \\
&= \frac{1}{\epsilon_0 c}\left(J^\nu A^\mu - J^\mu A^\nu\right) + F_\alpha{}^\nu F^{\alpha\mu} - F_\alpha{}^\mu F^{\alpha\nu}.
\end{aligned} \tag{18.30}$$

This first part is some sort of current-potential torque like beastie. That second part, the squared field term is what? I do not see an obvious way to reduce it to something more structured.

18.1.3 *Multivariable derivation.*

For completion sake, cut and pasted with most discussion omitted, the multiple field variable case follows in the same fashion as the single field variable Lagrangian.

$$\mathcal{L} = \mathcal{L}(\psi_\sigma, \partial_\mu \psi_\sigma). \tag{18.31}$$

The transformation is now:

$$\begin{aligned}\psi_\sigma &\to f_\sigma(\psi_\sigma, \alpha) = \psi'_\sigma \\ \mathcal{L}' &= \mathcal{L}(f_\sigma, \partial_\mu f_\sigma).\end{aligned} \tag{18.32}$$

Taking derivatives:

$$\frac{d\mathcal{L}'}{d\alpha} = \sum_\sigma \frac{\partial \mathcal{L}}{\partial f_\sigma}\frac{\partial f_\sigma}{\partial \alpha} + \sum_{\mu,\sigma} \frac{\partial \mathcal{L}}{\partial(\partial_\mu f_\sigma)}\frac{\partial(\partial_\mu f_\sigma)}{\partial \alpha}. \tag{18.33}$$

Again, making the Euler-Lagrange substitution of eq. (18.4) (with $f \to f_\sigma$) back into eq. (18.33) gives

$$\begin{aligned}\frac{d\mathcal{L}'}{d\alpha} &= \sum_\sigma \left(\sum_\mu \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f_\sigma)}\right)\frac{\partial f_\sigma}{\partial \alpha} + \sum_{\mu,\sigma} \frac{\partial \mathcal{L}}{\partial(\partial_\mu f_\sigma)}\frac{\partial(\partial_\mu f_\sigma)}{\partial \alpha} \\ &= \sum_{\mu,\sigma}\left(\left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu f_\sigma)}\right)\frac{\partial f_\sigma}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu f_\sigma)}\partial_\mu \frac{\partial f_\sigma}{\partial \alpha}\right) \\ &= \sum_{\mu,\sigma} \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu f_\sigma)}\frac{\partial f_\sigma}{\partial \alpha}\right) \\ &= \sum_\mu \gamma^\mu \partial_\mu \cdot \left(\sum_{\sigma,\nu} \gamma_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu f_\sigma)}\frac{\partial f_\sigma}{\partial \alpha}\right) \\ &= \nabla \cdot \left(\sum_{\sigma,\nu} \gamma_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \psi'_\sigma)}\frac{\partial \psi'_\sigma}{\partial \alpha}\right).\end{aligned} \tag{18.34}$$

Or

$$\frac{d\mathcal{L}'}{d\alpha} = \nabla \cdot J' = 0, \tag{18.35a}$$

$$J' = J'^\mu \gamma_\mu, \tag{18.35b}$$

$$J'^\mu = \sum_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi'_\sigma)}\frac{\partial \psi'_\sigma}{\partial \alpha}. \tag{18.35c}$$

A notational convenience for vector valued fields, in particular as we have in the electrodynamic Lagrangian for the vector potential, the chain rule summation in eq. (18.35) above can be replaced with a dot product.

$$J'^\mu = \gamma_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi'_\sigma)} \cdot \frac{\partial \gamma^\sigma \psi'_\sigma}{\partial \alpha}. \tag{18.36}$$

Dropping primes for convenience, and writing $\Psi = \gamma^\sigma \psi_\sigma$ for the vector field variable, the field form of Noether's law takes the form

$$J = \gamma_\mu \left(\gamma_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_\sigma)} \cdot \frac{\partial \Psi}{\partial \alpha} \right), \tag{18.37a}$$

$$\nabla \cdot J = 0. \tag{18.37b}$$

That is, a current vector with respect to this configuration space divergence is conserved when the Lagrangian field transformation is invariant.

18.2 TRANSLATION AND ROTATION NOETHER FIELD CURRENTS.

18.2.1 *Motivation.*

The article [21] details the calculation for a conserved current associated with an incremental Poincare transformation. Instead of starting with the canonical energy momentum tensor (arising from spacetime translation) which is not symmetric but can be symmetrized with other arguments, the paper of interest obtains the symmetric energy momentum tensor for Maxwell's equations directly.

I believe that I am slowly accumulating the tools required to understand this paper. One such tool is likely the exponential rotational generator examined in [9], utilizing the angular momentum operator.

Here I review the Noether conservation relations and the associated Noether currents for a single parameter alteration of the Lagrangian, incremental spacetime translation of the Lagrangian, and incremental Lorentz transform of the Lagrangian.

By reviewing these I hope that understanding the referenced article will be easier, or I independently understand (in my own way) how to apply similar techniques to the incremental Poincare transformed Lagrangian.

18.2.2 *Field Euler-Lagrange equations.*

The extremization of the action integral

$$S = \int \mathcal{L} d^4 x, \tag{18.38}$$

can be dealt with (following Feynman) as a first order Taylor expansion and integration by parts exercise. A single field variable example serves to illustrate. A first order Lagrangian of a single field variable has the form

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi). \tag{18.39}$$

Let us vary the field $\phi \to \phi + \bar{\phi}$ around the stationary field $\bar{\phi}$, inducing a corresponding variation in the action

$$\begin{aligned} S + \delta S &= \int \mathcal{L}(\phi + \bar{\phi}, \partial_\mu(phi + \bar{\phi}) d^4 x \\ &= \int d^4 x \left(\mathcal{L}(\bar{\phi}, \partial_\mu \bar{\phi}) + \bar{\phi} \frac{\partial \mathcal{L}}{\partial \phi} + \partial_\mu \bar{\phi} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} + \cdots \right). \end{aligned} \tag{18.40}$$

Neglecting any second or higher order terms the change in the action from the assumed solution is

$$\delta S = \int d^4 x \left(\bar{\phi} \frac{\partial \mathcal{L}}{\partial \phi} + \partial_\mu \bar{\phi} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right). \tag{18.41}$$

This is now integrable by parts yielding

$$\delta S = \int d^3 x \left(\bar{\phi} \partial_\mu \mathcal{L} \big|_{\partial x^\mu} \right) + \int d^4 x \bar{\phi} \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right). \tag{18.42}$$

Here $d^3 x$ is taken to mean that part of the integration not including dx_μ. The field $\bar{\phi}$ is always required to vanish on the boundary as in the dynamic Lagrangian arguments, so the first integral is zero. If the remainder is zero for all fields $\bar{\phi}$, then the inner term must be zero, and we the field Euler-Lagrange equations as a result

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0. \tag{18.43}$$

When we have multiple field variables, say A_ν, the chain rule expansion leading to eq. (18.41) will have to be modified to sum over all the field variables, and we end up instead with

$$\delta S = \int d^4 x \sum_\nu \overline{A_\nu} \left(\frac{\partial \mathcal{L}}{\partial A_\nu} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right). \tag{18.44}$$

So for $\delta S = 0$ for all $\bar{A}_\nu$ we have a set of equations, one for each ν

$$\frac{\partial \mathcal{L}}{\partial A_\nu} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = 0. \tag{18.45}$$

18.2.3 *Field Noether currents.*

The single parameter Noether conservation equation again is mainly application of the chain rule. Illustrating with the one field variable case, with an altered field variable $\phi \to \phi'(\theta)$, and

$$\mathcal{L}' = \mathcal{L}(\phi', \partial_\mu \phi'). \tag{18.46}$$

Examining the change of $\mathcal{L}'$ with θ we have

$$\frac{d\mathcal{L}'}{d\theta} = \frac{\partial \mathcal{L}}{\partial \phi'}\frac{\partial \phi'}{\partial \theta} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi')}\frac{\partial(\partial_\mu \phi')}{\partial \theta}. \tag{18.47}$$

For the last term we can switch up the order of differentiation

$$\begin{aligned}\frac{\partial(\partial_\mu \phi')}{\partial \theta} &= \frac{\partial}{\partial \theta}\frac{\partial \phi'}{\partial x^\mu} \\ &= \frac{\partial}{\partial x^\mu}\frac{\partial \phi'}{\partial \theta}.\end{aligned} \tag{18.48}$$

Additionally, with substitution of the Euler-Lagrange equations in the first term we have

$$\frac{d\mathcal{L}'}{d\theta} = \left(\frac{\partial}{\partial x^\mu}\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi')}\right)\frac{\partial \phi'}{\partial \theta} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi')}\frac{\partial}{\partial x^\mu}\frac{\partial \phi'}{\partial \theta}. \tag{18.49}$$

But this can be directly anti-differentiated yielding the Noether conservation equation

$$\frac{d\mathcal{L}'}{d\theta} = \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi')}\frac{\partial \phi'}{\partial \theta}\right). \tag{18.50}$$

With multiple field variables we will have a term in the chain rule expansion for each field variable. The end result is pretty much the same, but we have to sum over all the fields

$$\frac{d\mathcal{L}'}{d\theta} = \sum_\nu \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A'_\nu)}\frac{\partial A'_\nu}{\partial \theta}\right). \tag{18.51}$$

Unlike the field Euler-Lagrange equations we have just one here, not one for each field variable. In this multivariable case, expression in vector form can eliminate the sum over field variables. With $A' = A'_\nu \gamma^\nu$, we have

$$\frac{d\mathcal{L}'}{d\theta} = \frac{\partial}{\partial x^\mu}\left(\gamma_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A'_\nu)} \cdot \frac{\partial A'}{\partial \theta}\right). \tag{18.52}$$

With an evaluation at $\theta = 0$, we have finally

$$\left.\frac{d\mathcal{L}'}{d\theta}\right|_{\theta=0} = \frac{\partial}{\partial x^\mu}\left(\gamma_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \cdot \left.\frac{\partial A'}{\partial \theta}\right|_{\theta=0}\right). \tag{18.53}$$

When the Lagrangian alteration is independent of θ (i.e. is invariant), it is said that there is a symmetry. By eq. (18.53) we have a conserved quantity associated with this symmetry, some quantity, say J that has a zero divergence. That is

$$\begin{aligned} J^\mu &= \gamma_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \cdot \left.\frac{\partial A'}{\partial \theta}\right|_{\theta=0} \\ 0 &= \partial_\mu J^\mu. \end{aligned} \tag{18.54}$$

18.2.4 *Spacetime translation symmetries and Noether currents.*

Considering the effect of spacetime translation on the Lagrangian we examine the application of the first order linear Taylor series expansion shifting the vector parameters by an increment a. The Lagrangian alteration is

$$\mathcal{L} \to e^{a\cdot\nabla}\mathcal{L} \approx \mathcal{L} + a \cdot \nabla \mathcal{L}. \tag{18.55}$$

Similar to the addition of derivatives to the Lagrangians of dynamics, we can add in some types of total derivatives $\partial_\mu F^\mu$ to the Lagrangian without changing the resulting field equations (i.e. there is an associated "symmetry" for this Lagrangian alteration). The directional derivative $a \cdot \nabla \mathcal{L} = a^\mu \partial_\mu \mathcal{L}$ appears to be an example of a total derivative alteration that leaves the Lagrangian unchanged.

18.2.4.1 *On the symmetry.*

The fact that this translation necessarily results in the same field equations is not necessarily obvious. Using one of the simplest field Lagrangians, that of the Coulomb electrostatic law, we can illustrate that this is true in at least one case, and also see what is required in the general case

$$\mathcal{L} = \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 - \frac{1}{\epsilon_0}\rho\phi = \frac{1}{2}\sum_m (\partial_m \phi)^2 - \frac{1}{\epsilon_0}\rho\phi. \tag{18.56}$$

With partials written $\partial_m f = f_m$ we summarize the field Euler-Lagrange equations using the variational derivative

$$\frac{\delta}{\delta\phi} = \frac{\partial}{\partial\phi} - \sum_m \partial_m \frac{\partial}{\partial\phi_m}. \tag{18.57}$$

Where the extremum condition $\delta\mathcal{L}/\delta\phi = 0$ produces the field equations.

For the Coulomb Lagrangian without (spatial) translation, we have

$$\frac{\delta\mathcal{L}}{\delta\phi} = -\frac{1}{\epsilon_0}\rho - \partial_{mm}\phi. \tag{18.58}$$

So the extremum condition $\delta\mathcal{L}/\delta\phi = 0$ gives

$$\boldsymbol{\nabla}^2\phi = -\frac{1}{\epsilon_0}\rho. \tag{18.59}$$

Equivalently, and probably more familiar, we write $\mathbf{E} = -\boldsymbol{\nabla}\phi$, and get the differential form of Coulomb's law in terms of the electric field

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \frac{1}{\epsilon_0}\rho. \tag{18.60}$$

To consider the translation case we have to first evaluate the first order translation produced by the directional derivative. This is

$$\begin{aligned}\mathbf{a}\cdot\boldsymbol{\nabla}\mathcal{L} &= \sum_m a_m\partial_m\mathcal{L} \\ &= -\frac{\mathbf{a}}{\epsilon_0}\cdot(\rho\boldsymbol{\nabla}\phi + \phi\boldsymbol{\nabla}\rho).\end{aligned} \tag{18.61}$$

For the translation to be a symmetry the evaluation of the variational derivative must be zero. In this case we have

$$\begin{aligned}\frac{\delta}{\delta\phi}\mathbf{a}\cdot\boldsymbol{\nabla}\mathcal{L} &= -\frac{\mathbf{a}}{\epsilon_0}\cdot\frac{\delta}{\delta\phi}(\rho\boldsymbol{\nabla}\phi + \phi\boldsymbol{\nabla}\rho) \\ &= -\sum_m \frac{a_m}{\epsilon_0}\frac{\delta}{\delta\phi}(\rho\partial_m\phi + \phi\partial_m\rho) \\ &= -\sum_m \frac{a_m}{\epsilon_0}\left(\frac{\partial}{\partial\phi} - \sum_k \partial_k\frac{\partial}{\partial\phi_k}\right)(\rho\phi_m + \phi\rho_m).\end{aligned} \tag{18.62}$$

We see that the ϕ partials select only ρ derivatives whereas the ϕ_k partials select only the ρ term. All told we have zero

$$\begin{aligned}\left(\frac{\partial}{\partial\phi} - \sum_k \partial_k\frac{\partial}{\partial\phi_k}\right)(\rho\phi_m + \phi\rho_m) &= \rho_m - \sum_k \partial_k\rho\delta_{km} \\ &= \rho_m - \partial_m\rho \\ &= 0.\end{aligned} \tag{18.63}$$

This example illustrates that we have a symmetry provided we can "commute" the variational derivative with the gradient

$$\frac{\delta}{\delta\phi}\mathbf{a}\cdot\boldsymbol{\nabla}\mathcal{L} = \mathbf{a}\cdot\boldsymbol{\nabla}\frac{\delta\mathcal{L}}{\delta\phi}. \tag{18.64}$$

Since $\delta\mathcal{L}/\delta\phi = 0$ by construction, the resulting field equations are unaltered by such a modification.

Are there conditions where this commutation is not possible? Some additional exploration on symmetries associated with addition of derivatives to field Lagrangians was made previously in F. After all was said and done, the conclusion motivated by this simple example was also reached. Namely, we require the commutation condition eq. (18.64) between the variational derivative and the gradient of the Lagrangian.

18.2.4.2 *Existence of a symmetry for translational variation.*

Considering an example Lagrangian we found that there was a symmetry provided we could commute the variational derivative with the gradient, as in eq. (18.64) What this really means is not clear in general and a better answer to the existence question for incremental translation can be had by considering the transformation of the action directly around the stationary fields.

Without really any loss of generality we can consider an action with a four dimensional spacetime volume element, and apply the incremental translation operator to this

$$\begin{aligned}\int d^4xa\cdot\nabla\mathcal{L}(A^\beta+\overline{A}^\beta,\partial_\alpha A^\beta+\partial_\alpha\overline{A}^\beta)\\ = \int d^4xa\cdot\nabla\mathcal{L}(\overline{A}^\beta,\partial_\alpha\overline{A}^\beta)+\int d^4xa\cdot\nabla\left(\frac{\partial\mathcal{L}}{\partial A^\beta}\overline{A^\beta}+\frac{\partial\mathcal{L}}{\partial(\partial_\alpha A^\beta)}\partial_\alpha\overline{A^\beta}\right)+\cdots\end{aligned} \tag{18.65}$$

For the first term we have $a \cdot \nabla \int d^4x \mathcal{L}(\overline{A}^\beta, \partial_\alpha \overline{A}^\beta)$, but this integral is our stationary action. The remainder, to first order in the field variables, can then be expanded and integrated by parts

$$\begin{aligned}
&\int d^4x a^\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial A^\beta} \overline{A^\beta} + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\alpha \overline{A^\beta} \right) \\
&\quad = \int d^4x a^\mu \left(\left(\partial_\mu \frac{\partial \mathcal{L}}{\partial A^\beta} \right) \overline{A^\beta} + \frac{\partial \mathcal{L}}{\partial A^\beta} \left(\partial_\mu \overline{A^\beta} \right) \right) \\
&\quad + \int d^4x a^\mu \left(\left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \right) \partial_\alpha \overline{A^\beta} + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \left(\partial_\mu \partial_\alpha \overline{A^\beta} \right) \right) \\
&\quad = \int d^4x \left(\left(a^\mu \partial_\mu \frac{\partial \mathcal{L}}{\partial A^\beta} \right) \overline{A^\beta} - \left(\partial_\mu a^\mu \frac{\partial \mathcal{L}}{\partial A^\beta} \right) \overline{A^\beta} \right) \\
&\qquad + \int d^4x \left(\left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \right) \partial_\alpha \overline{A^\beta} - \left(\partial_\mu a^\mu \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \right) \partial_\alpha \overline{A^\beta} \right).
\end{aligned} \tag{18.66}$$

Since a^μ are constants, this is zero, so there can be no contribution to the field equations by the addition of the translation increment to the Lagrangian.

18.2.4.3 *Noether current derivation.*

With the assumption that the Lagrangian translation induces a symmetry, we can proceed with the calculation of the Noether current. This procedure for deriving the Noether current for an incremental spacetime translation follows along similar lines as the scalar alteration considered previously.

We start with the calculation of the first order alteration, expanding the derivatives. Let us work with a multiple field Lagrangian $\mathcal{L} = \mathcal{L}(A^\beta, \partial_\alpha A^\beta)$ right from the start

$$\begin{aligned}
a \cdot \nabla \mathcal{L} &= a^\mu \partial_\mu \mathcal{L} \\
&= a^\mu \left(\frac{\partial \mathcal{L}}{\partial A^\sigma} \frac{\partial A^\sigma}{\partial x^\mu} + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \frac{\partial(\partial_\alpha A^\beta)}{\partial x^\mu} \right).
\end{aligned} \tag{18.67}$$

Using the Euler-Lagrange field equations in the first term, and switching integration order in the second this can be written as a single derivative

$$\begin{aligned}
a \cdot \nabla \mathcal{L} &= a^\mu \left(\partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \frac{\partial A^\beta}{\partial x^\mu} + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\alpha \frac{\partial A^\beta}{\partial x^\mu} \right) \\
&= a^\mu \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \frac{\partial A^\beta}{\partial x^\mu} \right)
\end{aligned} \tag{18.68}$$

In the scalar Noether current we were able to form an similar expression, but one that was a first order derivative that could be set to zero, to fix the conservation relationship. Here there is no such freedom, but we can sneakily subtract $a \cdot \nabla \mathcal{L}$ from itself to calculate such a zero

$$0 = \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} a^\mu \frac{\partial A^\beta}{\partial x^\mu} - a^\alpha \mathcal{L} \right). \tag{18.69}$$

Since this must hold for any vector a, we have the freedom to choose the simplest such vector, a unit vector $a = \gamma_\nu$, for which $a^\mu = \delta^\mu{}_\nu$. Our current and its zero divergence relationship then becomes

$$\begin{aligned} T^\alpha{}_\nu &= \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\nu A^\beta - \delta^\alpha{}_\nu \mathcal{L} \\ 0 &= \partial_\alpha T^\alpha{}_\nu. \end{aligned} \tag{18.70}$$

This is not the symmetric energy momentum tensor that we want in the electrodynamics context although it can be obtained from it by adding just the right zero.

18.2.4.4 *Canonical energy momentum tensor and Lagrangian gradient.*

In [2] many tensor quantities are not written in index form, but instead using a vector notation. In particular, the symmetric energy momentum tensor is expressed as

$$T(a) = -\frac{\epsilon_0}{2} F a F. \tag{18.71}$$

where the usual tensor form following by taking dot products with γ^μ and substituting $a = \gamma^\nu$. The conservation equation for the canonical energy momentum tensor of eq. (18.70) can be put into a similar vector form

$$\begin{aligned} T(a) &= \gamma_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} (a \cdot \nabla) A^\beta - a \mathcal{L} \\ 0 &= \nabla \cdot T(a). \end{aligned} \tag{18.72}$$

The adjoint $\overline{T}$ of the tensor can be calculated from the definition

$$\nabla \cdot T(a) = a \cdot \overline{T}(\nabla). \tag{18.73}$$

Somewhat unintuitively, this is a function of the gradient. Playing around with factoring out the displacement vector a from eq. (18.72) that the energy momentum adjoint essentially provides an expansion of the gradient of the Lagrangian. To prepare, let us introduce some helper notation

$$\Pi_\beta \equiv \gamma_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)}. \tag{18.74}$$

With this our Noether current equation becomes

$$\begin{aligned}\nabla \cdot T(a) &= \langle \nabla T(a) \rangle \\ &= \left\langle \nabla(\Pi_\beta (a \cdot \nabla) A^\beta - a \nabla \mathcal{L}) \right\rangle \\ &= \left\langle \nabla \left(\frac{1}{2} \Pi_\beta (a(\nabla A^\beta) + (\nabla A^\beta) a) - a\mathcal{L} \right) \right\rangle.\end{aligned} \tag{18.75}$$

Cyclic permutation of the vector products $\langle abc \rangle = \langle cab \rangle$ can be used in the scalar selection. This is a little more tractable with some helper notation for the A^β gradients, say $v^\beta = \nabla A^\beta$. Because of the operator nature of the gradient once the vector order is permuted we have to allow for the gradient to act left or right or both, so arrows are used to disambiguate this where appropriate.

$$\begin{aligned}\nabla \cdot T(a) &= \left\langle \nabla \left(\frac{1}{2} \Pi_\beta a v^\beta + \Pi_\beta v^\beta a \right) - \nabla \mathcal{L} a \right\rangle \\ &= \left\langle \left(\frac{1}{2} v^\beta \overset{\leftrightarrow}{\nabla} \Pi_\beta \frac{1}{2} \nabla (\Pi_\beta v^\beta) - \nabla \mathcal{L} \right) a \right\rangle \\ &= a \cdot \left(\frac{1}{2} \left\langle v^\beta \overset{\leftrightarrow}{\nabla} \Pi_\beta + \nabla (\Pi_\beta v^\beta) \right\rangle_1 - \nabla \mathcal{L} \right).\end{aligned} \tag{18.76}$$

This dotted with quantity is the adjoint of the canonical energy momentum tensor

$$\bar{T}(\nabla) = \frac{1}{2} \left\langle v^\beta \overset{\leftrightarrow}{\nabla} \Pi_\beta + \nabla (\Pi_\beta v^\beta) \right\rangle_1 - \nabla \mathcal{L}. \tag{18.77}$$

This can however, be expanded further. First tackling the bidirectional gradient vector term we can utilize the property that the reverse of a vector leaves the vector unchanged. This gives us

$$\begin{aligned}\left\langle v^\beta \overset{\leftrightarrow}{\nabla} \Pi_\beta \right\rangle_1 &= \left\langle v^\beta (\overrightarrow{\nabla} \Pi_\beta) \right\rangle_1 + \left\langle (v^\beta \overleftarrow{\nabla}) \Pi_\beta \right\rangle_1 \\ &= \left\langle v^\beta (\overrightarrow{\nabla} \Pi_\beta) \right\rangle_1 + \left\langle \Pi_\beta (\overrightarrow{\nabla} v^\beta) \right\rangle_1.\end{aligned} \tag{18.78}$$

In the remaining term, using the Hestenes overdot notation clarify the scope of the operator, we have

$$\bar{T}(\nabla) = \frac{1}{2} \left(\left\langle v^\beta (\nabla \Pi_\beta) \right\rangle_1 + \left\langle \Pi_\beta (\nabla v^\beta) \right\rangle_1 + \left\langle (\nabla \Pi_\beta) v^\beta \right\rangle_1 + \left\langle \nabla' \Pi_\beta v^{\beta'} \right\rangle_1 \right) - \nabla \mathcal{L}. \tag{18.79}$$

The grouping of the first and third terms above simplifies nicely

$$\begin{aligned}
&\frac{1}{2}\left\langle v^\beta(\nabla\Pi_\beta)\right\rangle_1 + \frac{1}{2}\left\langle(\nabla\Pi_\beta)v^\beta\right\rangle_1 \\
&\quad = v^\beta(\nabla\cdot\Pi_\beta) + \frac{1}{2}\left\langle v^\beta(\nabla\wedge\Pi_\beta)\right\rangle_1 + \left\langle(\nabla\wedge\Pi_\beta)v^\beta\right\rangle_1.
\end{aligned} \tag{18.80}$$

Since $a(b\wedge c) + (b\wedge c)a = 2a\wedge b\wedge c$, which is purely a trivector, the vector grade selection above is zero. This leaves the adjoint reduced to

$$\overline{T}(\nabla) = v^\beta(\nabla\cdot\Pi_\beta) + \frac{1}{2}\left(\left\langle\Pi_\beta(\nabla v^\beta)\right\rangle_1 + \left\langle\nabla'\Pi_\beta v^{\beta'}\right\rangle_1\right) - \nabla\mathcal{L}. \tag{18.81}$$

For the remainder vector grade selection operators we have something that is of the following form

$$\frac{1}{2}\langle abc + bac\rangle_1 = (a\cdot b)c. \tag{18.82}$$

And we are finally able to put the adjoint into a form that has no remaining grade selection operators

$$\begin{aligned}
\overline{T}(\nabla) &= (\nabla A^\beta)(\nabla\cdot\Pi_\beta) + (\Pi_\beta\cdot\nabla)(\nabla A^\beta) - \nabla\mathcal{L} \\
&= (\nabla A^\beta)(\overrightarrow{\nabla}\cdot\Pi_\beta) + (\nabla A^\beta)(\overleftarrow{\nabla}\cdot\Pi_\beta) - \nabla\mathcal{L} \\
&= (\nabla A^\beta)(\overleftrightarrow{\nabla}\cdot\Pi_\beta) - \nabla\mathcal{L}.
\end{aligned} \tag{18.83}$$

Recapping, we have for the tensor and its adjoint

$$\begin{aligned}
0 &= \nabla\cdot T(a) = a\cdot\overline{T}(\nabla) \\
\Pi_\beta &\equiv \gamma_\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\alpha A^\beta)} \\
T(a) &= \Pi_\beta(a\cdot\nabla)A^\beta - a\nabla\mathcal{L} \\
\overline{T}(\nabla) &= (\nabla A^\beta)(\overleftrightarrow{\nabla}\cdot\Pi_\beta) - \nabla\mathcal{L}.
\end{aligned} \tag{18.84}$$

For the adjoint, since $a\cdot\overline{T}(\nabla) = 0$ for all a, we must also have $\overline{T}(\nabla) = 0$, which means the adjoint of the canonical energy momentum tensor really provides not much more than a recipe for computing the Lagrangian gradient

$$\nabla\mathcal{L} = (\nabla A^\beta)(\overleftrightarrow{\nabla}\cdot\Pi_\beta). \tag{18.85}$$

Having seen the adjoint notation, it was natural to see what this was for a multiple scalar field variable Lagrangian, even if it is not intrinsically

useful. Observe that the identity eq. (18.85), obtained so laboriously, is not more than syntactic sugar for the chain rule expansion of the Lagrangian partials (plus application of the Euler-Lagrange field equations). We could obtain this directly if desired much more easily than by factoring out a from $\nabla \cdot T(a) = 0$.

$$\begin{aligned}
\partial_\mu \mathcal{L} &= \frac{\partial \mathcal{L}}{\partial A^\beta} \partial_\mu A^\beta + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\mu \partial_\alpha A^\beta \\
&= \left(\partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \right) \partial_\mu A^\beta + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\alpha \partial_\mu A^\beta \\
&= \partial_\alpha \left(\left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \right) \partial_\mu A^\beta \right).
\end{aligned} \tag{18.86}$$

Summing over μ for the gradient, this reproduces eq. (18.85), with much less work

$$\begin{aligned}
\nabla \mathcal{L} &= \gamma^\mu \partial_\mu \mathcal{L} \\
&= \partial_\alpha \left(\left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \right) (\nabla A^\beta) \right) \\
&= (\Pi_\beta \cdot \overleftrightarrow{\nabla})(\nabla A^\beta).
\end{aligned} \tag{18.87}$$

Observe that the Euler-Lagrange field equations are implied in this relationship, so perhaps it has some utility. Also note that while it is simpler to directly compute this, without having started with the canonical energy momentum tensor, we would not know how the two of these were related.

18.2.5 *Noether current, infinitesimal Lorentz transformation.*

Let us assume that we can use the exponential generator of rotations

$$e^{(i \cdot x) \cdot \nabla} = 1 + (i \cdot x) \cdot \nabla + \cdots \tag{18.88}$$

to alter a Lagrangian density. In particular, that we can use the first order approximation of this Taylor series, applying the incremental rotation operator $(i \cdot x) \cdot \nabla = i \cdot (x \wedge \nabla)$ to transform the Lagrangian.

$$\mathcal{L} \to \mathcal{L} + (i \cdot x) \cdot \nabla \mathcal{L}. \tag{18.89}$$

Suppose that we parametrize the rotation bivector i using two perpendicular unit vectors u, and v. Here perpendicular is in the sense $uv = -vu$ so

that $i = u \wedge v = uv$. For the bivector expressed this way our incremental rotation operator takes the form

$$\begin{aligned}(i \cdot x) \cdot \nabla &= ((u \wedge v) \cdot x) \cdot \nabla \\ &= (u(v \cdot x) - v(u \cdot x)) \cdot \nabla \\ &= (v \cdot x)u \cdot \nabla - (u \cdot x))v \cdot \nabla.\end{aligned} \tag{18.90}$$

The operator is reduced to a pair of torque-like scaled directional derivatives, and we have already examined the Noether currents for the translations induced by the directional derivatives. It is not unreasonable to take exactly the same approach to consider rotation symmetries as we did for translation. We found for incremental translations

$$a \cdot \nabla \mathcal{L} = \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} (a \cdot \nabla) A^\beta \right). \tag{18.91}$$

So for incremental rotations the change to the Lagrangian is

$$(i \cdot x) \cdot \nabla \mathcal{L} = (v \cdot x)\partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} (u \cdot \nabla) A^\beta \right) - (u \cdot x)\partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} (v \cdot \nabla) A^\beta \right). \tag{18.92}$$

Since the choice to make u and v both unit vectors and perpendicular has been made, there is really no loss in generality to align these with a pair of the basis vectors, say $u = \gamma_\mu$ and $v = \gamma_\nu$. The incremental rotation operator is reduced to

$$\begin{aligned}(i \cdot x) \cdot \nabla &= (\gamma_\nu \cdot x)\gamma_\mu \cdot \nabla - (\gamma_\mu \cdot x))\gamma_\nu \cdot \nabla \\ &= x_\nu \partial_\mu - x_\mu \partial_\nu.\end{aligned} \tag{18.93}$$

Similarly the change to the Lagrangian is

$$(i \cdot x) \cdot \nabla \mathcal{L} = x_\nu \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\mu A^\beta \right) - x_\mu \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\nu A^\beta \right). \tag{18.94}$$

Subtracting the two, essentially forming $(i \cdot x) \cdot \nabla \mathcal{L} - (i \cdot x) \cdot \nabla \mathcal{L} = 0$, we have

$$0 = x_\nu \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\mu A^\beta - \delta^\alpha{}_\mu \mathcal{L} \right) - x_\mu \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\nu A^\beta - \delta^\alpha{}_\nu \mathcal{L} \right). \tag{18.95}$$

We previously wrote

$$T^\alpha{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\nu A^\beta - \delta^\alpha{}_\nu \mathcal{L}. \tag{18.96}$$

for the Noether current of spacetime translation, and with that our conservation equation becomes

$$0 = x_\nu \partial_\alpha T^\alpha{}_\mu - x_\mu \partial_\alpha T^\alpha{}_\nu. \tag{18.97}$$

As is, this does not really appear to say much, since we previously also found $\partial_\alpha T^\alpha{}_\nu = 0$. We appear to need a way to pull the x coordinates into the derivatives to come up with a more interesting statement. A test expansion of $\nabla \cdot (i \cdot x)\mathcal{L}$ to see what is left over compared to $(i \cdot x) \cdot \nabla \mathcal{L}$ shows that there is in fact no difference, and we actually have the identity

$$i \cdot (x \wedge \nabla)\mathcal{L} = (i \cdot x) \cdot \nabla \mathcal{L} = \nabla \cdot (i \cdot x)\mathcal{L}. \tag{18.98}$$

This suggests that we can pull the x coordinates into the derivatives of eq. (18.97) as in

$$0 = \partial_\alpha \left(T^\alpha{}_\mu x_\nu - T^\alpha{}_\nu x_\mu \right). \tag{18.99}$$

However, expanding this derivative shows that this is fact not the case. Instead we have

$$\begin{aligned} \partial_\alpha \left(T^\alpha{}_\mu x_\nu - T^\alpha{}_\nu x_\mu \right) &= T^\alpha{}_\mu \partial_\alpha x_\nu - T^\alpha{}_\nu \partial_\alpha x_\mu \\ &= T^\alpha{}_\mu \eta_{\alpha\nu} - T^\alpha{}_\nu \eta_{\alpha\mu} \\ &= T_{\nu\mu} - T_{\mu\nu}. \end{aligned} \tag{18.100}$$

So instead of a Noether current, following the procedure used to calculate the spacetime translation current, we have only a mediocre compromise

$$\begin{aligned} M^\alpha{}_{\mu\nu} &\equiv T^\alpha{}_\mu x_\nu - T^\alpha{}_\nu x_\mu \\ \partial_\alpha M^\alpha{}_{\mu\nu} &= T_{\nu\mu} - T_{\mu\nu}. \end{aligned} \tag{18.101}$$

Jackson [7] ends up with a similar index upper expression

$$M^{\alpha\beta\gamma} \equiv T^{\alpha\beta} x^\gamma - T^{\alpha\gamma} x^\beta. \tag{18.102}$$

and then uses a requirement for vanishing 4-divergence of this quantity

$$0 = \partial_\alpha M^{\alpha\beta\gamma}. \tag{18.103}$$

to symmetries this tensor by subtracting off all the antisymmetric portions. The differences compared to Jackson with upper verses lower indices are

minor for we can follow the same arguments and arrive at the same sort of $0 - 0 = 0$ result as we had in eq. (18.97)

$$0 = x^\nu \partial_\alpha T^{\alpha\mu} - x^\mu \partial_\alpha T^{\alpha\nu}. \tag{18.104}$$

The only difference is that our not-really-a-conservation equation becomes

$$\partial_\alpha M^{\alpha\mu\nu} = T^{\nu\mu} - T^{\mu\nu}. \tag{18.105}$$

18.2.5.1 *An example of the symmetry.*

While not a proof that application of the incremental rotation operator is a symmetry, an example at least provides some comfort that this is a reasonable thing to attempt. Again, let us consider the Coulomb Lagrangian

$$\mathcal{L} = \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 - \frac{1}{\epsilon_0}\rho\phi. \tag{18.106}$$

For this we have

$$\begin{aligned}\mathcal{L}' &= \mathcal{L} + (i \cdot \mathbf{x}) \cdot \boldsymbol{\nabla}\mathcal{L} \\ &= \mathcal{L} - (i \cdot \mathbf{x}) \cdot \frac{1}{\epsilon_0}\left(\rho\boldsymbol{\nabla}\phi + \phi\boldsymbol{\nabla}\rho\right).\end{aligned} \tag{18.107}$$

If the variational derivative of the incremental rotation contribution is zero, then we have a symmetry.

$$\begin{aligned}\frac{\delta}{\delta\phi}(i \cdot \mathbf{x}) \cdot \boldsymbol{\nabla}\mathcal{L} &= (i \cdot \mathbf{x}) \cdot \frac{1}{\epsilon_0}\boldsymbol{\nabla}\rho - \sum_m \partial_m \left((i \cdot \mathbf{x}) \cdot \frac{1}{\epsilon_0}\rho\mathbf{e}_m\right) \\ &= (i \cdot \mathbf{x}) \cdot \frac{1}{\epsilon_0}\boldsymbol{\nabla}\rho - \boldsymbol{\nabla} \cdot \left((i \cdot \mathbf{x})\frac{1}{\epsilon_0}\rho\right).\end{aligned} \tag{18.108}$$

As found in eq. (18.98), we have $(i \cdot \mathbf{x}) \cdot \boldsymbol{\nabla} = \boldsymbol{\nabla} \cdot (i \cdot \mathbf{x})$, so we have

$$\frac{\delta}{\delta\phi}(i \cdot \mathbf{x}) \cdot \boldsymbol{\nabla}\mathcal{L} = 0. \tag{18.109}$$

for this specific Lagrangian as expected.

Note that the test expansion I used to state eq. (18.98) was done using only the bivector $i = \gamma_\mu \wedge \gamma_\nu$. An expansion with $i = u^\alpha u^\beta \gamma_\alpha \wedge \gamma_\beta$

shows that this is also the case in shows that this is true more generally. Specifically, this expansion gives

$$\begin{aligned}\nabla \cdot (i \cdot x)\mathcal{L} &= (i \cdot x) \cdot \nabla \mathcal{L} + (\eta_{\alpha\beta} - \eta_{\beta\alpha})u^\alpha v^\beta \mathcal{L} \\ &= (i \cdot x) \cdot \nabla \mathcal{L}.\end{aligned} \tag{18.110}$$

(since the metric tensor is symmetric).

Loosely speaking, the geometric reason for this is that $\nabla \cdot f(x)$ takes its maximum (or minimum) when $f(x)$ is colinear with x and is zero when $f(x)$ is perpendicular to x. The vector $i \cdot x$ is a combined projection and 90 degree rotation in the plane of the bivector, and the divergence is left with no colinear components to operate on.

While this commutation of the $i \cdot \mathbf{x}$ with the divergence operator did not help with finding the Noether current, it does at least show that we have a symmetry. Demonstrating the invariance for the general Lagrangian (at least the single field variable case) likely follows the same procedure as in this specific example above.

18.2.5.2 *General existence of the rotational symmetry.*

The example above hints at a general method to demonstrate that the incremental Lorentz transform produces a symmetry. It will be sufficient to consider the variation around the stationary field variables for the change due to the action from the incremental rotation operator. That is

$$\delta S = \int d^4x (i \cdot x) \cdot \nabla \mathcal{L}(A^\beta + \bar{A}^\beta, \partial_\alpha A^\beta + \partial_\alpha \bar{A}^\beta). \tag{18.111}$$

Performing a first order Taylor expansion of the Lagrangian around the stationary field variables we have

$$\begin{aligned}\delta S &= \int d^4x (i \cdot x) \cdot \gamma^\mu \partial_\mu \mathcal{L}(A^\beta + \bar{A}^\beta, \partial_\alpha A^\beta + \partial_\alpha \bar{A}^\beta) \\ &= \int d^4x (i \cdot x) \cdot \gamma^\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial A^\beta} \bar{A}^\beta + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} (\partial_\alpha \bar{A}^\beta) \right) \\ &= \int d^4x (i \cdot x) \cdot \gamma^\mu \\ &\qquad \left(\left(\partial_\mu \frac{\partial \mathcal{L}}{\partial A^\beta} \right) \bar{A}^\beta + \frac{\partial \mathcal{L}}{\partial A^\beta} \partial_\mu \bar{A}^\beta \right) \\ &+ \int d^4x (i \cdot x) \cdot \gamma^\mu \\ &\qquad \left(\left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \right) (\partial_\alpha \bar{A}^\beta) + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha A^\beta)} \partial_\mu (\partial_\alpha \bar{A}^\beta) \right).\end{aligned} \tag{18.112}$$

Doing the integration by parts we have

$$
\begin{aligned}
\delta S &= \int d^4x \bar{A}^\beta \gamma^\mu \cdot \left((i \cdot x) \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial A^\beta} \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial A^\beta} (i \cdot x) \right) \right) \\
&+ \int d^4x (\partial_\alpha \bar{A}^\beta) \gamma^\mu \cdot \left((i \cdot x) \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\alpha A^\beta)} \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\alpha A^\beta)} (i \cdot x) \right) \right) \\
&= \int d^4x \bar{A}^\beta \left((i \cdot x) \cdot \nabla \frac{\partial \mathcal{L}}{\partial A^\beta} - \nabla \cdot (i \cdot x) \frac{\partial \mathcal{L}}{\partial A^\beta} \right) \\
&\quad + (\partial_\alpha \bar{A}^\beta) \left((i \cdot x) \cdot \nabla \frac{\partial \mathcal{L}}{\partial (\partial_\alpha A^\beta)} - \nabla \cdot (i \cdot x) \frac{\partial \mathcal{L}}{\partial (\partial_\alpha A^\beta)} \right).
\end{aligned}
\tag{18.113}
$$

Since $(i \cdot x) \cdot \nabla f = \nabla \cdot (i \cdot x) f$ for any f, there is no change to the resulting field equations due to this incremental rotation, so we have a symmetry for any Lagrangian that is first order in its derivatives.

18.3 SOLUTIONS.

A MATHEMATICA NOTEBOOKS.

These Mathematica notebooks, some just trivial ones used to generate figures, others more elaborate, and perhaps some even polished, can be found in

https://github.com/peeterjoot/mathematica/tree/master/classicalmechanics/.

The free Wolfram CDF player, is capable of read-only viewing these notebooks to some extent.

Files saved explicitly as CDF have interactive content that can be explored with the CDF player.

- Feb 19, 2012 infiniteCylinderPotential.cdf

  Attempt at evaluating the potential for an infinite cylinder.

- Feb 24, 2012 psIIp4InfPlanePotTakeIII.cdf

  Attempt at evaluating the potential for an infinite plane. Experimenting with using mathematica to produce decent documents, as well as trying a variation of the previous calculation where I used $R^2 \sim e$.

  The final output is not as nice as latex, but the save as latex option seems promising. New Mathematica tools used in this notebook include HoldForm, TraditionalForm, and ReleaseHold, which can be used to generate traditional form by default for scratch display generation.

  Note that cut-and-pasting URLS in comments as I've been doing get mangled and can't be followed. Switched the ones in this doc to Insert->Hyperlink instead.

- Feb 27, 2012 psIIp4InfCylPot.cdf

  Attempt evaluation of a cylindrical potential.

  New Mathematica methods used: HoldForm, Assuming, Assumptions.

- Jan 26, 2016 multisphericalPendulum.nb

  calculate the matrix products from the papers to verify (and as it turns out, correct).

- Nov 4, 2020 reciprocalMatrix.nb

  This demonstrates solving for a two vector reciprocal frame basis, using the STA metric.

B

SPHERICAL N-PENDULUM PROBLEM.

Abstract. The dynamics of chain like objects can be idealized as a multiple pendulum, treating the system as a set of point masses, joined by rigid massless connecting rods, and frictionless pivots. The double planar pendulum and single mass spherical pendulum problems are well treated in Lagrangian physics texts, but due to complexity a similar treatment of the spherical N-pendulum problem is not pervasive. We show that this problem can be tackled in a direct fashion, even in the general case with multiple masses and no planar constraints. A matrix factorization of the kinetic energy into allows an explicit and compact specification of the Lagrangian. Once that is obtained the equations of motion for this generalized pendulum system follow directly.

B.1 INTRODUCTION.

Derivation of the equations of motion for a planar motion constrained double pendulum system and a single spherical pendulum system are given as problems or examples in many texts covering Lagrangian mechanics. Setup of the Lagrangian, particularly an explicit specification of the system kinetic energy, is the difficult aspect of the multiple mass pendulum problem. Each mass in the system introduces additional interaction coupling terms, complicating the kinetic energy specification. In this paper, we use matrix algebra to determine explicitly the Lagrangian for the spherical N pendulum system, and to evaluate the Euler-Lagrange equations for the system.

It is well known that the general specification of the kinetic energy for a system of independent point masses takes the form of a symmetric quadratic form [4] [5]. However, actually calculating that energy explicitly for the general N-pendulum is likely thought too pedantic for even the most punishing instructor to inflict on students as a problem or example.

Given a 3×1 coordinate vector of velocity components for each mass relative to the position of the mass it is connected to, we can factor this as a $(3 \times 2)(2 \times 1)$ product of matrices where the 2×1 matrix is a vector of angular velocity components in the spherical polar representation. The remaining matrix factor contains all the trigonometric dependence. Such a grouping can be used to tidily separate the kinetic energy into an explicit quadratic form, sandwiching a symmetric matrix between two vectors of generalized velocity coordinates.

This paper is primarily a brute force and direct attack on the problem. It contains no new science, only a systematic treatment of a problem that is omitted from mechanics texts, yet conceptually simple enough to deserve treatment.

The end result of this paper is a complete and explicit specification of the Lagrangian and evaluation of the Euler-Lagrange equations for the chain-like N spherical pendulum system. While this end result is essentially nothing more than a non-linear set of coupled differential equations, it is believed that the approach used to obtain it has some elegance. Grouping all the rotational terms of the kinetic into a symmetric kernel appears to be a tidy way to tackle multiple discrete mass problems. At the very least, the calculation performed can show that a problem perhaps thought to be too messy for a homework exercise yields nicely to an organized and systematic attack.

B.2 DIVING RIGHT IN.

We make the simplifying assumptions of point masses, rigid massless connecting rods, and frictionless pivots.

B.2.1 *Single spherical pendulum.*

Using polar angle θ and azimuthal angle ϕ, and writing $S_\theta = \sin\theta$, $C_\phi = \cos\phi$ and so forth, we have for the coordinate vector on the unit sphere

$$\hat{\mathbf{r}} = \begin{bmatrix} C_\phi S_\theta \\ S_\phi S_\theta \\ C_\theta \end{bmatrix}. \tag{B.1}$$

The Lagrangian for the pendulum is then

$$L = \frac{1}{2} m l \dot{\hat{\mathbf{r}}}^T \dot{\hat{\mathbf{r}}} - m g l C_\theta. \tag{B.2}$$

This is somewhat unsatisfying since the unit vector derivatives have not been evaluated. Doing so we get

$$\dot{\hat{\mathbf{r}}} = \begin{bmatrix} C_\phi C_\theta \dot{\theta} - S_\phi S_\theta \dot{\phi} \\ S_\phi C_\theta \dot{\theta} + C_\phi S_\theta \dot{\phi} \\ -S_\theta \dot{\theta} \end{bmatrix}. \tag{B.3}$$

This however, is an ugly beastie to take the norm of as is. It is straightforward to show that this norm is just

$$\dot{\hat{\mathbf{r}}}^{\mathrm{T}} \dot{\hat{\mathbf{r}}} = \dot{\theta}^2 + S_\theta^2 \dot{\phi}^2, \tag{B.4}$$

however, the brute force multiplication that leads to this result is not easily generalized to the multiple pendulum problem. Instead of actually expanding this now, lets defer that until later and instead write for a coordinate vector of angular velocity components

$$\Omega = \begin{bmatrix} \dot{\theta} \dot{\phi} \end{bmatrix}. \tag{B.5}$$

Now the unit polar derivative eq. (B.3) can be factored as

$$\begin{aligned} \dot{\hat{\mathbf{r}}} &= A^{\mathrm{T}} \Omega \\ A &= \begin{bmatrix} C_\phi C_\theta & S_\phi C_\theta & -S_\theta \\ -S_\phi S_\theta & C_\phi S_\theta & 0 \end{bmatrix}. \end{aligned} \tag{B.6a}$$

Our Lagrangian now takes the explicit form

$$\begin{aligned} L &= \frac{1}{2} m l \Omega^{\mathrm{T}} A A^{\mathrm{T}} \Omega - m g l C_\theta \\ A A^{\mathrm{T}} &= \begin{bmatrix} 1 & 0 \\ 0 & S_\theta^2 \end{bmatrix}. \end{aligned} \tag{B.7a}$$

B.2.2 *Spherical double pendulum.*

Before generalizing to N links, consider the double pendulum. Let the position of each of the k-th mass (with $k = 1, 2$) be

$$\mathbf{u}_k = \mathbf{u}_{k-1} + l_k \hat{\mathbf{r}}_k = \sum_{j=1}^{k} l_k \hat{\mathbf{r}}_k. \tag{B.8}$$

The unit vectors from the origin to the first mass, or from the first mass to the second have derivatives

$$\dot{\hat{\mathbf{r}}}_k = A_k{}^{\mathrm{T}}\dot{\mathbf{\Theta}}_k, \tag{B.9}$$

where

$$\begin{aligned} A_k &= \begin{bmatrix} C_{\phi_k}C_{\theta_k} & S_{\phi_k}C_{\theta_k} & -S_{\theta_k} \\ -S_{\phi_k}S_{\theta_k} & C_{\phi_k}S_{\theta_k} & 0 \end{bmatrix} \\ \mathbf{\Theta}_k &= \begin{bmatrix} \theta_k \\ \phi_k \end{bmatrix}. \end{aligned} \tag{B.10}$$

Since

$$\frac{d\mathbf{u}_k}{dt} = \sum_{j=1}^{k} l_j A_j^{\mathrm{T}}\dot{\Theta}_j, \tag{B.11}$$

The squared velocity of each mass is

$$\left|\frac{d\mathbf{u}_k}{dt}\right|^2 = \sum_{r,s=1}^{k} l_r l_s \dot{\Theta}_r^{\mathrm{T}} A_r A_s^{\mathrm{T}} \dot{\Theta}_s. \tag{B.12}$$

To see the structure of this product, it is helpful to expand this sum completely, something that is feasible for this $N = 2$ case. First for $k = 1$ we have just

$$\left|\frac{d\mathbf{u}_1}{dt}\right|^2 = l_1^2 \dot{\Theta}_1^{\mathrm{T}} A_1 A_1^{\mathrm{T}} \dot{\Theta}_1, \tag{B.13}$$

and for $k = 2$ we have

$$\begin{aligned} \left|\frac{d\mathbf{u}_2}{dt}\right|^2 &= l_1^2 \dot{\Theta}_1^{\mathrm{T}} A_1 A_1^{\mathrm{T}} \dot{\Theta}_1 + l_2^2 \dot{\Theta}_2^{\mathrm{T}} A_2 A_2^{\mathrm{T}} \dot{\Theta}_2 + l_1 l_2 \dot{\Theta}_1^{\mathrm{T}} A_1 A_2^{\mathrm{T}} \dot{\Theta}_2 + l_2 l_1 \dot{\Theta}_2^{\mathrm{T}} A_2 A_1^{\mathrm{T}} \dot{\Theta}_1 \\ &= \left(l_1^2 \dot{\Theta}_1^{\mathrm{T}} A_1 A_1^{\mathrm{T}} + l_2 l_1 \dot{\Theta}_2^{\mathrm{T}} A_2 A_1^{\mathrm{T}}\right)\dot{\Theta}_1 + \left(l_2^2 \dot{\Theta}_2^{\mathrm{T}} A_2 A_2^{\mathrm{T}} + l_1 l_2 \dot{\Theta}_1^{\mathrm{T}} A_1 A_2^{\mathrm{T}}\right)\dot{\Theta}_2 \\ &= \begin{bmatrix} \dot{\Theta}_1^{\mathrm{T}} & \dot{\Theta}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} l_1^2 A_1 A_1^{\mathrm{T}} \\ l_2 l_1 A_2 A_1^{\mathrm{T}} \end{bmatrix} \dot{\Theta}_1 + \begin{bmatrix} \dot{\Theta}_1^{\mathrm{T}} & \dot{\Theta}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} l_1 l_2 A_1 A_2^{\mathrm{T}} \\ l_2^2 A_2 A_2^{\mathrm{T}} \end{bmatrix} \dot{\Theta}_2 \\ &= \begin{bmatrix} \dot{\Theta}_1^{\mathrm{T}} & \dot{\Theta}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} l_1^2 A_1 A_1^{\mathrm{T}} & l_1 l_2 A_1 A_2^{\mathrm{T}} \\ l_2 l_1 A_2 A_1^{\mathrm{T}} & l_2^2 A_2 A_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \dot{\Theta}_1 \\ \dot{\Theta}_2 \end{bmatrix}. \end{aligned} \tag{B.14}$$

Observe that these can be summarized by writing

$$
\begin{aligned}
B_1^{\mathrm{T}} &= \begin{bmatrix} l_1 A_1^{\mathrm{T}} & 0 \end{bmatrix} \\
B_2^{\mathrm{T}} &= \begin{bmatrix} l_1 A_1^{\mathrm{T}} & l_2 A_2^{\mathrm{T}} \end{bmatrix} \\
\Theta &= \begin{bmatrix} \dot{\Theta}_1 \\ \dot{\Theta}_2 \end{bmatrix} \\
\dot{\mathbf{u}}_k &= \dot{\Theta}^{\mathrm{T}} B_k B_k^{\mathrm{T}} \dot{\Theta}.
\end{aligned}
\tag{B.15}
$$

The kinetic energy for particle one is

$$
\begin{aligned}
K_1 &= \frac{1}{2} m_1 \dot{\Theta}^{\mathrm{T}} B_1 B_1^{\mathrm{T}} \dot{\Theta} \\
&= \dot{\Theta}^{\mathrm{T}} \begin{bmatrix} m_1 l_1^2 A_1 A_1^{\mathrm{T}} & 0 \\ 0 & 0 \end{bmatrix} \dot{\Theta},
\end{aligned}
\tag{B.16}
$$

and for the second particle

$$
K_2 = \frac{1}{2} m_2 \dot{\Theta}^{\mathrm{T}} B_2 B_2^{\mathrm{T}} \dot{\Theta} = \frac{1}{2} m_2 \dot{\Theta}^{\mathrm{T}} \begin{bmatrix} l_1^2 A_1 A_1^{\mathrm{T}} & l_1 l_2 A_1 A_2^{\mathrm{T}} \\ l_2 l_1 A_2 A_1^{\mathrm{T}} & l_2^2 A_2 A_2^{\mathrm{T}} \end{bmatrix} \dot{\Theta}. \tag{B.17}
$$

Summing these we have

$$
K = \frac{1}{2} \dot{\Theta}^{\mathrm{T}} \begin{bmatrix} (m_1 + m_2) l_1^2 A_1 A_1^{\mathrm{T}} & m_2 l_1 l_2 A_1 A_2^{\mathrm{T}} \\ m_2 l_2 l_1 A_2 A_1^{\mathrm{T}} & m_2 l_2^2 A_2 A_2^{\mathrm{T}} \end{bmatrix} \dot{\Theta}. \tag{B.18}
$$

For the mass sums let

$$
\mu_k \equiv \sum_{j=k}^{2} m_j, \tag{B.19}
$$

so

$$
K = \frac{1}{2} \dot{\Theta}^{\mathrm{T}} \begin{bmatrix} \mu_1 l_1^2 A_1 A_1^{\mathrm{T}} & \mu_2 l_1 l_2 A_1 A_2^{\mathrm{T}} \\ \mu_2 l_2 l_1 A_2 A_1^{\mathrm{T}} & \mu_2 l_2^2 A_2 A_2^{\mathrm{T}} \end{bmatrix} \dot{\Theta}. \tag{B.20}
$$

If the matrix of quadradic factors is designated Q, so that

$$
K = \frac{1}{2} \dot{\Theta}^{\mathrm{T}} Q \dot{\Theta}, \tag{B.21}
$$

then the (i,j) element of the matrix Q is given by

$$
Q_{ij} = \mu_{\max(i,j)} l_i l_j A_i A_j^{\mathrm{T}}. \tag{B.22}
$$

For the potential energy, things are simplest if that energy is measured from the $z = 0$ plane. The potential energy for mass 1 is

$$T_1 = m_1 g l_1 \cos\theta_1, \tag{B.23}$$

and the potential energy for mass 2 is

$$T_2 = m_2 g \left(l_1 \cos\theta_1 + l_2 \cos\theta_2 \right). \tag{B.24}$$

The total potential energy for the system is

$$\begin{aligned} T &= (m_1 + m_2) g l_1 \cos\theta_1 + m_2 g l_2 \cos\theta_2 \\ &= \sum_{k=1}^{2} \mu_k g l_k \cos\theta_k. \end{aligned} \tag{B.25}$$

B.2.3 *N spherical pendulum.*

Having written things out explicitly for the two particle case, the generalization to N particles is straightforward

$$\begin{aligned} \boldsymbol{\Theta}^{\mathrm{T}} &= \begin{bmatrix} \boldsymbol{\Theta}_1^{\mathrm{T}} & \boldsymbol{\Theta}_2^{\mathrm{T}} & \cdots & \boldsymbol{\Theta}_N^{\mathrm{T}} \end{bmatrix} \\ Q_{ij} &= \mu_{\max(i,j)} l_i l_j A_i A_j^{\mathrm{T}} \\ K &= \frac{1}{2} \dot{\Theta}^{\mathrm{T}} Q \dot{\Theta} \\ \mu_k &= \sum_{j=k}^{N} m_j \\ \Phi &= g \sum_{k=1}^{N} \mu_k l_k \cos\theta_k \\ L &= K - \Phi. \end{aligned} \tag{B.26}$$

After some expansion one can find that the block matrices $A_i A_j^{\mathrm{T}}$ contained in Q are

$$A_i A_j^{\mathrm{T}} = \begin{bmatrix} C_{\phi_j - \phi_i} C_{\theta_i} C_{\theta_j} + S_{\theta_i} S_{\theta_j} & -S_{\phi_j - \phi_i} C_{\theta_i} S_{\theta_j} \\ S_{\phi_j - \phi_i} C_{\theta_j} S_{\theta_i} & C_{\phi_j - \phi_i} S_{\theta_i} S_{\theta_j} \end{bmatrix}. \tag{B.27}$$

The diagonal blocks are particularly simple and have no ϕ dependence

$$A_i A_i^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta_i \end{bmatrix}. \tag{B.28}$$

B.3 EVALUATING THE EULER-LAGRANGE EQUATIONS.

It will be convenient to group the Euler-Lagrange equations into a column vector form, with a column vector of generalized coordinates and derivatives, and position and velocity gradients in the associated phase space

$$\mathbf{q} \equiv \left[q_r\right]_r \tag{B.29a}$$

$$\dot{\mathbf{q}} \equiv \left[\dot{q}_r\right]_r \tag{B.29b}$$

$$\nabla_{\mathbf{q}} L \equiv \left[\frac{\partial L}{\partial q_r}\right]_r \tag{B.29c}$$

$$\nabla_{\dot{\mathbf{q}}} L \equiv \left[\frac{\partial L}{\partial \dot{q}_r}\right]_r. \tag{B.29d}$$

In this form the Euler-Lagrange equations take the form of a single vector equation

$$\nabla_{\mathbf{q}} L = \frac{d}{dt} \nabla_{\dot{\mathbf{q}}} L. \tag{B.30}$$

We are now set to evaluate these generalized phase space gradients. For the acceleration terms our computation reduces nicely to a function of only Q

$$\begin{aligned} \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}_a} &= \frac{1}{2}\frac{d}{dt}\left(\frac{\partial \dot{\mathbf{\Theta}}}{\partial \dot{\theta}_a}^{\mathrm{T}} Q \dot{\mathbf{\Theta}} + \dot{\mathbf{\Theta}}^{\mathrm{T}} Q \frac{\partial \dot{\mathbf{\Theta}}}{\partial \dot{\theta}_a}\right) \\ &= \frac{d}{dt}\left(\left[\delta_{ac}\begin{bmatrix}1 & 0\end{bmatrix}\right]_c Q \dot{\mathbf{\Theta}}\right), \end{aligned} \tag{B.31}$$

and

$$\begin{aligned} \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}_a} &= \frac{1}{2}\frac{d}{dt}\left(\frac{\partial \dot{\mathbf{\Theta}}}{\partial \dot{\phi}_a}^{\mathrm{T}} Q \dot{\mathbf{\Theta}} + \dot{\mathbf{\Theta}}^{\mathrm{T}} Q \frac{\partial \dot{\mathbf{\Theta}}}{\partial \dot{\phi}_a}\right) \\ &= \frac{d}{dt}\left(\left[\delta_{ac}\begin{bmatrix}0 & 1\end{bmatrix}\right]_c Q \dot{\mathbf{\Theta}}\right). \end{aligned} \tag{B.32}$$

The last groupings above made use of $Q = Q^{\mathrm{T}}$, and in particular $(Q + Q^{\mathrm{T}})/2 = Q$. We can now form a column matrix putting all the angular velocity gradient in a tidy block matrix representation

$$\nabla_{\dot{\boldsymbol{\Theta}}} L = \left[\left[\begin{array}{c} \frac{\partial L}{\partial \dot{\theta}_r} \\ \frac{\partial L}{\partial \dot{\phi}_r} \end{array}\right]\right]_r = Q\dot{\boldsymbol{\Theta}}. \tag{B.33}$$

A small aside on Hamiltonian form. This velocity gradient is also the conjugate momentum of the Hamiltonian, so if we wish to express the Hamiltonian in terms of conjugate momenta, we require invertability of Q at the point in time that we evaluate things. Writing

$$P_{\boldsymbol{\Theta}} = \nabla_{\dot{\boldsymbol{\Theta}}} L, \tag{B.34}$$

and noting that $(Q^{-1})^{\mathrm{T}} = Q^{-1}$, we get for the kinetic energy portion of the Hamiltonian

$$K = \frac{1}{2} P_{\boldsymbol{\Theta}}{}^{\mathrm{T}} Q^{-1} P_{\boldsymbol{\Theta}}. \tag{B.35}$$

Now, the invertiblity of Q cannot be taken for granted. Even in the single particle case we do not have invertiblity. For the single particle case we have

$$Q = ml^2 \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}, \tag{B.36}$$

so at $\theta = \pm\pi/2$ this quadratic form is singular, and the planar angular momentum becomes a constant of motion. Returning to the evaluation of the Euler-Lagrange equations, the problem is now reduced to calculating the right hand side of the following system

$$\frac{d}{dt}\left(Q\dot{\boldsymbol{\Theta}}\right) = \left[\left[\begin{array}{c} \frac{\partial L}{\partial \theta_r} \\ \frac{\partial L}{\partial \phi_r} \end{array}\right]\right]_r. \tag{B.37}$$

With back substitution of eq. (B.27), and eq. (B.28) we have a complete and explicit matrix expansion of the left hand side. For the right hand side taking the θ_a and ϕ_a derivatives respectively we get

$$\frac{\partial L}{\partial \theta_a} = \frac{1}{2}\dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left[\mu_{\max(r,c)} l_r l_c \left(\frac{\partial A_r}{\partial \theta_a} A_c^{\mathrm{T}} + A_r \frac{\partial A_c}{\partial \theta_a}{}^{\mathrm{T}}\right)\right]_{rc} \dot{\boldsymbol{\Theta}} - g\mu_a l_a \sin\theta_a, \tag{B.38a}$$

$$\frac{\partial L}{\partial \phi_a} = \frac{1}{2}\dot{\mathbf{\Theta}}^{\mathrm{T}}\left[\mu_{\max(r,c)} l_r l_c \left(\frac{\partial A_r}{\partial \phi_a} A_c^{\mathrm{T}} + A_r \frac{\partial A_c}{\partial \phi_a}^{\mathrm{T}}\right)\right]_{rc} \dot{\mathbf{\Theta}}. \tag{B.38b}$$

So to proceed we must consider the $A_r A_c^{\mathrm{T}}$ partials. A bit of thought shows that the matrices of partials above are mostly zeros. Illustrating by example, consider $\partial Q/\partial \theta_2$, which in block matrix form is

$$\begin{bmatrix} 0 & \frac{1}{2}\mu_2 l_1 l_2 A_1 \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}} & 0 & \dots & 0 \\ \frac{1}{2}\mu_2 l_2 l_1 \frac{\partial A_2}{\partial \theta_2} A_1^{\mathrm{T}} & \frac{1}{2}\mu_2 l_2 l_2 \left(A_2 \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}} + \frac{\partial A_2}{\partial \theta_2} A_2^{\mathrm{T}}\right) & \frac{1}{2}\mu_3 l_2 l_3 \frac{\partial A_2}{\partial \theta_2} A_3^{\mathrm{T}} & \dots & \frac{1}{2}\mu_N l_2 l_N \frac{\partial A_2}{\partial \theta_2} A_N^{\mathrm{T}} \\ 0 & \frac{1}{2}\mu_3 l_3 l_2 A_3 \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}} & 0 & \dots & 0 \\ 0 & \vdots & 0 & \dots & 0 \\ 0 & \frac{1}{2}\mu_N l_N l_2 A_N \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}} & 0 & \dots & 0 \end{bmatrix}. \tag{B.39}$$

Observe that the diagonal term has a scalar plus its transpose, so we can drop the one half factor and one of the summands for a total contribution to $\partial L/\partial \theta_2$ of just

$$\mu_2 {l_2}^2 \dot{\mathbf{\Theta}}_2^{\mathrm{T}} \frac{\partial A_2}{\partial \theta_2} A_2^{\mathrm{T}} \dot{\mathbf{\Theta}}_2. \tag{B.40}$$

Now consider one of the pairs of off diagonal terms. Adding these contributions to $\partial L/\partial \theta_2$ of

$$\begin{aligned} &\frac{1}{2}\mu_2 l_1 l_2 \dot{\mathbf{\Theta}}_1^{\mathrm{T}} A_1 \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}} \dot{\mathbf{\Theta}}_2 + \frac{1}{2}\mu_2 l_2 l_1 \dot{\mathbf{\Theta}}_2^{\mathrm{T}} \frac{\partial A_2}{\partial \theta_2} A_1^{\mathrm{T}} \dot{\mathbf{\Theta}}_1 \\ &\quad = \frac{1}{2}\mu_2 l_1 l_2 \dot{\mathbf{\Theta}}_1^{\mathrm{T}} \left(A_1 \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}} + A_1 \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}}\right) \dot{\mathbf{\Theta}}_2 \\ &\quad = \mu_2 l_1 l_2 \dot{\mathbf{\Theta}}_1^{\mathrm{T}} A_1 \frac{\partial A_2}{\partial \theta_2}^{\mathrm{T}} \dot{\mathbf{\Theta}}_2. \end{aligned} \tag{B.41}$$

This has exactly the same form as the diagonal term, so summing over all terms we get for the position gradient components of the Euler-Lagrange equation just

$$\frac{\partial L}{\partial \theta_a} = \sum_k \mu_{\max(k,a)} l_k l_a \dot{\mathbf{\Theta}}_k^{\mathrm{T}} A_k \frac{\partial A_a}{\partial \theta_a}^{\mathrm{T}} \dot{\mathbf{\Theta}}_a - g\mu_a l_a \sin\theta_a, \tag{B.42}$$

and

$$\frac{\partial L}{\partial \phi_a} = \sum_k \mu_{\max(k,a)} l_k l_a \dot{\mathbf{\Theta}}_k^{\mathrm{T}} A_k \frac{\partial A_a}{\partial \phi_a}^{\mathrm{T}} \dot{\mathbf{\Theta}}_a. \tag{B.43}$$

The only thing that remains to do is evaluate the $A_k \partial A_a / \partial \phi_a{}^{\mathrm{T}}$ matrices. Utilizing eq. (B.27), one obtains easily

$$A_k \frac{\partial A_r}{\partial \theta_r}^{\mathrm{T}} = \begin{bmatrix} S_{\theta_k} C_{\theta_r} - C_{\theta_k} S_{\theta_r} C_{\phi_k - \phi_r} & C_{\theta_k} C_{\theta_r} S_{\phi_k - \phi_r} \\ S_{\theta_k} S_{\theta_r} S_{\phi_k - \phi_r} & S_{\theta_k} C_{\theta_r} C_{\phi_k - \phi_r} \end{bmatrix}, \tag{B.44}$$

and

$$A_k \frac{\partial A_r}{\partial \phi_r}^{\mathrm{T}} = \begin{bmatrix} C_{\theta_k} C_{\theta_r} S_{\phi_k - \phi_r} & -C_{\theta_k} S_{\theta_r} C_{\phi_k - \phi_r} \\ S_{\theta_k} C_{\theta_r} C_{\phi_k - \phi_r} & S_{\theta_k} S_{\theta_r} S_{\phi_k - \phi_r} \end{bmatrix}. \tag{B.45}$$

The right hand side of the Euler-Lagrange equations now becomes

$$\nabla_{\boldsymbol{\Theta}} L = \sum_k \left[\begin{bmatrix} \mu_{\max(k,r)} l_k l_r \dot{\boldsymbol{\Theta}}_k^{\mathrm{T}} A_k \frac{\partial A_r}{\partial \theta_r}^{\mathrm{T}} \dot{\boldsymbol{\Theta}}_r \\ \mu_{\max(k,r)} l_k l_r \dot{\boldsymbol{\Theta}}_k^{\mathrm{T}} A_k \frac{\partial A_r}{\partial \phi_r}^{\mathrm{T}} \dot{\boldsymbol{\Theta}}_r \end{bmatrix} \right]_r - g \left[\mu_r l_r \sin \theta_r \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right]_r . \tag{B.46}$$

Between eq. (B.46), eq. (B.33), and a few other auxiliary relations, all above we have completed the task of evaluating the Euler-Lagrange equations for this multiple particle distinct mass system. Unfortunately, just as the simple planar pendulum is a non-linear system, so is this. Possible options for solution are numerical methods or solution restricted to a linear approximation in a small neighborhood of a particular phase space point.

B.4 SUMMARY.

Looking back it is hard to tell the trees from the forest. Here is a summary of the results and definitions of importance. First the Langrangian itself

$$
\begin{aligned}
\mu_k &= \sum_{j=k}^{N} m_j \\
\boldsymbol{\Theta}_k &= \begin{bmatrix} \theta_k \\ \phi_k \end{bmatrix} \\
\boldsymbol{\Theta}^{\mathrm{T}} &= \begin{bmatrix} \boldsymbol{\Theta}_1^{\mathrm{T}} & \boldsymbol{\Theta}_2^{\mathrm{T}} & \dots & \boldsymbol{\Theta}_N^{\mathrm{T}} \end{bmatrix} \\
A_k &= \begin{bmatrix} C_{\phi_k} C_{\theta_k} & S_{\phi_k} C_{\theta_k} & -S_{\theta_k} \\ -S_{\phi_k} S_{\theta_k} & C_{\phi_k} S_{\theta_k} & 0 \end{bmatrix} \\
Q &= \left[\mu_{\max(r,c)} l_r l_c A_r A_c^{\mathrm{T}} \right]_{rc} \\
K &= \frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} Q \dot{\boldsymbol{\Theta}} \\
\Phi &= g \sum_{k=1}^{N} \mu_k l_k \cos\theta_k \\
L &= K - \Phi.
\end{aligned}
\tag{B.47a}
$$

Evaluating the Euler-Lagrange equations for the system, we get

$$
\begin{aligned}
0 &= \nabla_{\boldsymbol{\Theta}} L - \frac{d}{dt} \left(\nabla_{\dot{\boldsymbol{\Theta}}} L \right) \\
&= \sum_k \left[\begin{bmatrix} \mu_{\max(k,r)} l_k l_r \dot{\boldsymbol{\Theta}}_k^{\mathrm{T}} A_k \frac{\partial A_r}{\partial \theta_r}^{\mathrm{T}} \dot{\boldsymbol{\Theta}}_r \\ \mu_{\max(k,r)} l_k l_r \dot{\boldsymbol{\Theta}}_k^{\mathrm{T}} A_k \frac{\partial A_r}{\partial \phi_r}^{\mathrm{T}} \dot{\boldsymbol{\Theta}}_r \end{bmatrix} \right]_r - g \left[\mu_r l_r \sin\theta_r \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right]_r - \frac{d}{dt} \left(Q \dot{\boldsymbol{\Theta}} \right).
\end{aligned}
\tag{B.48}
$$

Making this explicit requires evaluation of some of the matrix products. With verification in multisphericalPendulum.nb, those are

$$
\begin{aligned}
A_r A_c^{\mathrm{T}} &= \begin{bmatrix} C_{\phi_c - \phi_r} C_{\theta_r} C_{\theta_c} + S_{\theta_r} S_{\theta_c} & -S_{\phi_c - \phi_r} C_{\theta_r} S_{\theta_c} \\ S_{\phi_c - \phi_r} C_{\theta_c} S_{\theta_r} & C_{\phi_c - \phi_r} S_{\theta_r} S_{\theta_c} \end{bmatrix} \\
A_k \frac{\partial A_r}{\partial \theta_r}^{\mathrm{T}} &= \begin{bmatrix} S_{\theta_k} C_{\theta_r} - C_{\theta_k} S_{\theta_r} C_{\phi_k - \phi_r} & C_{\theta_k} C_{\theta_r} S_{\phi_k - \phi_r} \\ S_{\theta_k} S_{\theta_r} S_{\phi_k - \phi_r} & S_{\theta_k} C_{\theta_r} C_{\phi_k - \phi_r} \end{bmatrix} \\
A_k \frac{\partial A_r}{\partial \phi_r}^{\mathrm{T}} &= \begin{bmatrix} C_{\theta_k} C_{\theta_r} S_{\phi_k - \phi_r} & -C_{\theta_k} S_{\theta_r} C_{\phi_k - \phi_r} \\ S_{\theta_k} C_{\theta_r} C_{\phi_k - \phi_r} & S_{\theta_k} S_{\theta_r} S_{\phi_k - \phi_r} \end{bmatrix}.
\end{aligned}
\tag{B.49}
$$

C DIRECT VARIATION OF MAXWELL EQUATIONS.

C.1 MOTIVATION, DEFINITIONS AND SETUP.

This document will attempt to calculate Maxwell's equation, which in multivector form is

$$\nabla F = J/\epsilon_0 c. \tag{C.1}$$

using a Lagrangian energy density variational approach.

C.1.1 *Tensor form of the field.*

Explicit expansion of the field bivector in terms of coordinates one has

$$\begin{aligned} F &= \mathbf{E} + Ic\mathbf{B} \\ &= E^k \gamma_{k0} + \gamma_{0123k0} c B^k \\ &= E^k \gamma_{k0} + (\gamma_0)^2 (\gamma_k)^2 \epsilon^{ij}{}_k c \gamma_{ij} B^k. \end{aligned} \tag{C.2}$$

The complete coordinate expansion of the field is

$$F = E^k \gamma_{k0} - c\epsilon^{ij}{}_k B^k \gamma_{ij}. \tag{C.3}$$

When this bivector is expressed in terms of basis bivectors $\gamma_{\mu\nu}$ we have

$$F = \sum_{\mu<\nu} (F \cdot \gamma^{\nu\mu}) \gamma_{\mu\nu} = \frac{1}{2} (F \cdot \gamma^{\nu\mu}) \gamma_{\mu\nu}. \tag{C.4}$$

As shorthand for the coordinates the field can be expressed with respect to various bivector basis sets in tensor form

$$\begin{aligned} F^{\mu\nu} &= F \cdot \gamma^{\nu\mu} & F &= (1/2) F^{\mu\nu} \gamma_{\mu\nu} \\ F_{\mu\nu} &= F \cdot \gamma_{\nu\mu} & F &= (1/2) F_{\mu\nu} \gamma^{\mu\nu} \\ F_\mu{}^\nu &= F \cdot \gamma_\nu{}^\mu & F &= (1/2) F_\mu{}^\nu \gamma^\mu{}_\nu \\ F^\mu{}_\nu &= F \cdot \gamma^\nu{}_\mu & F &= (1/2) F^\mu{}_\nu \gamma_\mu{}^\nu. \end{aligned} \tag{C.5}$$

In particular, we can extract the electric field components by dotting with a spacetime mix of indices

$$F^{i0} = E^k \gamma_{k0} \cdot \gamma^{0i} = E^i = -F_{i0}. \tag{C.6}$$

and the magnetic field components by dotting with the bivectors having a pure spatial mix of indices

$$F^{ij} = -c\epsilon^{ab}{}_k B^k \gamma_{ab} \cdot \gamma^{ji} = -c\epsilon^{ij}{}_k B^k = F_{ij}. \tag{C.7}$$

It is customary to summarize these tensors in matrix form

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -cB^3 & cB^2 \\ E^2 & cB^3 & 0 & -cB^1 \\ E^3 & -cB^2 & cB^1 & 0 \end{bmatrix}, \tag{C.8a}$$

$$F_{\mu\nu} = \begin{bmatrix} 0 & E^1 & E^2 & E^3 \\ -E^1 & 0 & -cB^3 & cB^2 \\ -E^2 & cB^3 & 0 & -cB^1 \\ -E^3 & -cB^2 & cB^1 & 0 \end{bmatrix}. \tag{C.8b}$$

Neither of these matrices will be needed explicitly, but are included for comparison since there is some variation in the sign conventions and units used for the field tensor. Observe that these matrix representations are both sparse and filled with redudancy, and are not a particularily great representation of the field.

c.1.1.1 *Potential form.*

With the assumption that the field can be expressed in terms of the curl of a potential vector

$$F = \nabla \wedge A, \tag{C.9}$$

the tensor expression of the field becomes

$$\begin{aligned} F^{\mu\nu} &= F \cdot (\gamma^\nu \wedge \gamma^\mu) = \partial^\mu A^\nu - \partial^\nu A^\mu \\ F_{\mu\nu} &= F \cdot (\gamma_\nu \wedge \gamma_\mu) = \partial_\mu A_\nu - \partial_\nu A_\mu \\ F^\mu{}_\nu &= F \cdot (\gamma^\nu \wedge \gamma_\mu) = \partial^\mu A_\nu - \partial_\nu A^\mu \\ F_\mu{}^\nu &= F \cdot (\gamma_\nu \wedge \gamma^\mu) = \partial_\mu A^\nu - \partial^\nu A_\mu. \end{aligned} \tag{C.10}$$

These field bivector coordinates will be used in the Lagrangian calculations.

D

LORENTZ INVARIANCE OF MAXWELL LAGRANGIAN.

D.1 WORKING IN MULTIVECTOR FORM.

D.1.1 *Lorentz boost of field Lagrangian.*

The multivector form of the field Lagrangian is

$$\begin{aligned}\mathcal{L} &= \kappa(\nabla \wedge A)^2 + A \cdot J \\ \kappa &= -\frac{\epsilon_0 c}{2}.\end{aligned} \tag{D.1}$$

Write the boosting transformation on a four vector in exponential form

$$L(X) = \exp(\alpha \hat{\mathbf{a}}/2) X \exp(-\alpha \hat{\mathbf{a}}/2) = \Lambda X \Lambda^\dagger, \tag{D.2}$$

where $\hat{\mathbf{a}} = a^i \gamma_i \wedge \gamma_0$ is any unit spacetime bivector, and α represents the rapidity angle.

Consider first the transformation of the interaction term with $A' = L(A)$, and $J' = L(J)$

$$\begin{aligned}A' \cdot J' &= \langle L(A) L(J) \rangle \\ &= \left\langle \Lambda A \Lambda^\dagger \Lambda J \Lambda^\dagger \right\rangle \\ &= \left\langle \Lambda A J \Lambda^\dagger \right\rangle \\ &= \left\langle \Lambda^\dagger \Lambda A J \right\rangle \\ &= \langle A J \rangle \\ &= A \cdot J.\end{aligned} \tag{D.3}$$

Now consider the boost applied to the field bivector $F = \mathbf{E} + Ic\mathbf{B} = \nabla \wedge A$, by boosting both the gradient and the potential

$$\begin{aligned}
\nabla' \wedge A' &= L(\nabla) \wedge L(A) \\
&= \Lambda\nabla) \wedge L(A) \\
&= (\Lambda\nabla\Lambda^\dagger) \wedge (\Lambda A\Lambda^\dagger) \\
&= \frac{1}{2}\left((\Lambda\nabla\Lambda^\dagger)(\Lambda A\Lambda^\dagger) - (\Lambda A\Lambda^\dagger)(\Lambda\nabla\Lambda^\dagger)\right) \\
&= \frac{1}{2}\left(\Lambda\nabla A\Lambda^\dagger - \Lambda A\nabla\Lambda^\dagger\right) \\
&= \Lambda(\nabla \wedge A)\Lambda^\dagger.
\end{aligned} \tag{D.4}$$

The boosted squared field bivector in the Lagrangian is thus

$$\begin{aligned}
(\nabla' \wedge A')^2 &= \Lambda(\nabla \wedge A)^2\Lambda^\dagger \\
&= \Lambda(\mathbf{E} + Ic\mathbf{B})^2\Lambda^\dagger \\
&= \Lambda((\mathbf{E}^2 - c^2\mathbf{B}^2) + 2Ic\mathbf{E}\cdot\mathbf{B})\Lambda^\dagger \\
&= ((\mathbf{E}^2 - c^2\mathbf{B}^2)\Lambda\Lambda^\dagger + 2(\Lambda I\Lambda^\dagger)c\mathbf{E}\cdot\mathbf{B}) \\
&= ((\mathbf{E}^2 - c^2\mathbf{B}^2) + 2I\Lambda\Lambda^\dagger c\mathbf{E}\cdot\mathbf{B}) \\
&= ((\mathbf{E}^2 - c^2\mathbf{B}^2) + 2Ic\mathbf{E}\cdot\mathbf{B}) \\
&= (\mathbf{E} + Ic\mathbf{B})^2 \\
&= (\nabla \wedge A)^2.
\end{aligned} \tag{D.5}$$

The commutation of the pseudoscalar I with the boost exponential $\Lambda = \exp(\alpha\hat{\mathbf{a}}/2) = \cosh(\alpha/2) + \hat{\mathbf{a}}\sinh(\alpha/2)$ is possible since I anticommutes with all four vectors and thus commutes with bivectors, such as $\hat{\mathbf{a}}$. I also necessarily commutes with the scalar components of this exponential, and thus commutes with any even grade multivector.

Putting all the pieces together this shows that the Lagrangian in its entirety is a Lorentz invariant

$$\mathcal{L}' = \kappa(\nabla' \wedge A')^2 + A' \cdot J' = \kappa(\nabla \wedge A)^2 + A \cdot J = \mathcal{L}. \tag{D.6}$$

FIXME: what is the conserved quantity associated with this? There should be one according to Noether's theorem? Is it the gauge condition $\nabla \cdot A = 0$?

D.1.1.1 *Maxwell equation invariance.*

Somewhat related, having calculated the Lorentz transform of $F = \nabla \wedge A$, is an aside showing that the Maxwell equation is unsurprisingly also is a Lorentz invariant.

$$\begin{aligned} \nabla'(\nabla' \wedge A') &= J' \\ \Lambda\nabla\Lambda^\dagger\Lambda(\nabla \wedge A)\Lambda^\dagger &= \Lambda J\Lambda^\dagger \\ \Lambda\nabla(\nabla \wedge A)\Lambda^\dagger &= \Lambda J\Lambda^\dagger. \end{aligned} \tag{D.7}$$

Pre and post multiplying with $\Lambda^\dagger$, and Λ respectively returns the unboosted equation

$$\nabla(\nabla \wedge A) = J. \tag{D.8}$$

D.1.2 *Lorentz boost applied to the Lorentz force Lagrangian.*

Next interesting case is the Lorentz force, which for a time positive metric signature is:

$$L = qA \cdot v/c + \frac{1}{2}mv \cdot v. \tag{D.9}$$

The boost invariance of the $A \cdot J$ dot product demonstrated above demonstrates the general invariance property for any four vector dot product, and this Lagrangian has nothing but dot products in it. It thus follows directly that the Lorentz force Lagrangian is also a Lorentz invariant.

D.2 REPEAT IN TENSOR FORM.

Now, I can follow the above, but presented with the same sort of calculation in tensor form I am hopeless to understand it. To attempt translating this into tensor form, it appears the first step is putting the Lorentz transform itself into tensor or matrix form.

D.2.1 *Translating versors to matrix form.*

To get the feeling for how this will work, assume $\hat{\mathbf{a}} = \sigma_1$, so that the boost is along the x-axis. In that case we have

$$L(X) = (\cosh(\alpha/2) + \gamma_{10}\sinh(\alpha/2))x^\mu\gamma_\mu(\cosh(\alpha/2) + \gamma_{01}\sinh(\alpha/2)).$$

(D.10)

Writing $C = \cosh(\alpha/2)$, and $S = \sinh(\alpha/2)$, and observing that the exponentials commute with the γ_2, and γ_3 directions so the exponential action on those directions cancel.

$$L(X) = x^2\gamma_2 + x^3\gamma_3 + (C + \gamma_{10}S)(x^0\gamma_0 + x^1\gamma_1)(C + \gamma_{01}S). \tag{D.11}$$

Expanding just the non-perpendicular parts of the above

$$\begin{aligned}
&(C + \gamma_{10}S)(x^0\gamma_0 + x^1\gamma_1)(C + \gamma_{01}S) \\
&= x^0(C^2\gamma_0 + \gamma_{10001}S^2) + x^0SC(\gamma_{001} + \gamma_{100}) \\
&\quad + x^1(C^2\gamma_1 + \gamma_{10101}S^2) + x^1SC(\gamma_{101} + \gamma_{101}) \\
&= x^0(C^2\gamma_0 - \gamma_{01100}S^2) + 2x^0SC\gamma_{001} + x^1(C^2\gamma_1 - \gamma_{11001}S^2) - 2x^1SC\gamma_{011} \\
&= (x^0\gamma_0 + x^1\gamma_1)(C^2 + S^2) + 2(\gamma_0)^2SC(x^0\gamma_1 + x^1\gamma_0) \\
&= (x^0\gamma_0 + x^1\gamma_1)\cosh(\alpha) + (\gamma_0)^2\sinh(\alpha)(x^0\gamma_1 + x^1\gamma_0) \\
&= \gamma_0(x^0\cosh(\alpha) + x^1\sinh((\gamma_0)^2\alpha)) + \gamma_1(x^1\cosh(\alpha) + x^0\sinh((\gamma_0)^2\alpha)).
\end{aligned} \tag{D.12}$$

In matrix form the complete transformation is thus

$$\begin{aligned}
\begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}' &= \begin{bmatrix} \cosh(\alpha(\gamma_0)^2) & \sinh(\alpha(\gamma_0)^2) & 0 & 0 \\ \sinh(\alpha(\gamma_0)^2) & \cosh(\alpha(\gamma_0)^2) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \\
&= \cosh(\alpha(\gamma_0)^2) \begin{bmatrix} 1 & \tanh(\alpha(\gamma_0)^2) & 0 & 0 \\ \tanh(\alpha(\gamma_0)^2) & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}.
\end{aligned} \tag{D.13}$$

This supplies the specific meaning for the α factor in the exponential form, namely:

$$\begin{aligned}
\alpha &= -\tanh^{-1}(\beta(\gamma_0)^2) \\
&= -\tanh^{-1}(|\mathbf{v}|/c(\gamma_0)^2).
\end{aligned} \tag{D.14}$$

Or

$$\begin{aligned}
\alpha\hat{\mathbf{a}} &= -\tanh^{-1}(\hat{\mathbf{a}}|\mathbf{v}|/c(\gamma_0)^2) \\
&= -\tanh^{-1}(\mathbf{v}/c(\gamma_0)^2).
\end{aligned} \tag{D.15}$$

Putting this back into the original Lorentz boost equation to tidy it up, and writing $\tanh(\mathbf{A}) = \mathbf{v}/c$, the Lorentz boost is

$$L(X) = \begin{cases} \exp(-\mathbf{A}/2)X\exp(\mathbf{A}/2) & \text{for } (\gamma_0)^2 = 1 \\ \exp(\mathbf{A}/2)X\exp(-\mathbf{A}/2) & \text{for } (\gamma_0)^2 = -1 \end{cases}. \tag{D.16}$$

Both of the metric signature options are indicated here for future reference and comparison with results using the alternate signature.

D.2.1.1 *Revisit the expansion to matrix form above.*

Looking back, multiplying out all the half angle terms as done above is this is the long dumb hard way to do it. A more sensible way would be to note that $\exp(\alpha\hat{\mathbf{a}}/2)$ anticommutes with both γ_0 and γ_1 thus

$$\begin{aligned}\exp(\alpha\hat{\mathbf{a}}/2)(x^0\gamma_0 + x^1\gamma_1)\exp(-\alpha\hat{\mathbf{a}}/2) &= \exp(\alpha\hat{\mathbf{a}})(x^0\gamma_0 + x^1\gamma_1) \\ &= (\cosh(\alpha) + \hat{\mathbf{a}}\sinh(\alpha))(x^0\gamma_0 + x^1\gamma_1).\end{aligned} \tag{D.17}$$

The matrix form thus follows directly.

D.3 TRANSLATING VERSORS TENSOR FORM.

After this temporary digression back to the multivector form of the Lorentz transformation lets dispose of the specifics of the boost direction and magnitude, and also the metric signature. Instead encode all of these in a single versor variable Λ, again writing

$$L(X) = \Lambda X \Lambda^\dagger. \tag{D.18}$$

D.3.1 *Tensor form of vector Lorentz transform.*

What is the general way to encode this linear transformation in tensor/-matrix form? The transformed vector is just that a vector, and thus can be written in terms of coordinates for some basis

$$\begin{aligned}L(X) &= (L(X)\cdot e^\mu)e_\mu \\ &= ((\Lambda(x^\nu\gamma_\nu)\Lambda^\dagger)\cdot e^\mu)e_\mu \\ &= x^\nu((\Lambda\gamma_\nu\Lambda^\dagger)\cdot e^\mu)e_\mu.\end{aligned} \tag{D.19}$$

The inner term is just the tensor that we want. Write

$$\begin{aligned}\Lambda_\nu{}^\mu &= (\Lambda\gamma_\nu\Lambda^\dagger)\cdot e^\mu \\ \Lambda^\nu{}_\mu &= (\Lambda\gamma^\nu\Lambda^\dagger)\cdot e_\mu,\end{aligned} \tag{D.20}$$

for

$$\begin{aligned}L(X) &= x^\nu\Lambda_\nu{}^\mu e_\mu \\ &= x_\nu\Lambda^\nu{}_\mu e^\mu.\end{aligned} \tag{D.21}$$

Completely eliminating the basis, working in just the coordinates $X = x'^\mu e_\mu = x'_\mu e^\mu$ this is

$$\begin{aligned}x'^\mu &= x^\nu\Lambda_\nu{}^\mu \\ x'_\mu &= x_\nu\Lambda^\nu{}_\mu.\end{aligned} \tag{D.22}$$

Now, in particular, having observed that the dot product is a Lorentz invariant this should supply the index manipulation rule for operating with the Lorentz boost tensor in a dot product context.

Write

$$\begin{aligned}L(X)\cdot L(Y) &= (x^\nu\Lambda_\nu{}^\mu e_\mu)\cdot(y_\alpha\Lambda^\alpha{}_\beta e^\beta) \\ &= x^\nu y_\alpha\Lambda_\nu{}^\mu\Lambda^\alpha{}_\beta e_\mu\cdot e^\beta \\ &= x^\nu y_\alpha\Lambda_\nu{}^\mu\Lambda^\alpha{}_\mu.\end{aligned} \tag{D.23}$$

Since this equals $x^\nu y_\nu$, the tensor rule must therefore be

$$\Lambda_\mu{}^\sigma\Lambda^\nu{}_\sigma = \delta_\mu{}^\nu. \tag{D.24}$$

After a somewhat long path, the core idea behind the Lorentz boost tensor is that it is the "matrix" of a linear transformation that leaves the four vector dot product unchanged. There is no need to consider any Clifford algebra formulations to express just that idea.

D.3.2 *Misc notes.*

FIXME: To complete the expression of this in tensor form enumerating exactly how to express the dot product in tensor form would also be reasonable. ie: how to compute the reciprocal coordinates without describing the basis. Doing this will introduce the metric tensor into the mix.

Looks like the result eq. (D.24) is consistent with [20] and that doc starts making a bit more sense now. I do see that he uses primes to distinguish the boost tensor from its inverse (using the inverse tensor (primed index down) to transform the covariant (down) coordinates). Is there a convention for keeping free vs. varied indices close to the body of the operator? For the boost tensor he puts the free index closer to Λ, but for the inverse tensor for a covariant coordinate transformation puts the free index further out? This also appears to be notational consistent with [22].

D.3.3 *Expressing bivector Lorentz transform in tensor form.*

Having translated a vector Lorentz transform into tensor form, the next step is to do the same for a bivector. In particular for the field bivector $F = \nabla \wedge A$.

Write

$$\begin{aligned}\nabla' &= \Lambda \gamma_\mu \partial^\mu \Lambda^\dagger \\ A' &= \Lambda A^\nu \gamma_\nu \Lambda^\dagger.\end{aligned} \tag{D.25}$$

$$\begin{aligned}\nabla' \cdot e^\beta &= (\Lambda \gamma_\mu \Lambda^\dagger) \cdot e^\beta \partial^\mu = \Lambda_\mu{}^\beta \partial^\mu \\ A' \cdot e^\beta &= (\Lambda \gamma_\nu \Lambda^\dagger) \cdot e^\beta A^\nu = \Lambda_\nu{}^\beta A^\nu.\end{aligned} \tag{D.26}$$

Then the transformed bivector is

$$\begin{aligned}F' = \nabla' \wedge A' &= ((\nabla' \cdot e^\alpha) e_\alpha) \wedge ((A' \cdot e^\beta) e_\beta) \\ &= (e_\alpha \wedge e_\beta) \Lambda_\mu{}^\alpha \Lambda_\nu{}^\beta \partial^\mu A^\nu,\end{aligned} \tag{D.27}$$

and finally the transformed tensor is thus

$$\begin{aligned}F^{ab\prime} &= F' \cdot (e^b \wedge e^a) \\ &= (e_\alpha \wedge e_\beta) \cdot (e^b \wedge e^a) \Lambda_\mu{}^\alpha \Lambda_\nu{}^\beta \partial^\mu A^\nu \\ &= (\delta_\alpha{}^a \delta_\beta{}^b - \delta_\beta{}^a \delta_\alpha{}^b) \Lambda_\mu{}^\alpha \Lambda_\nu{}^\beta \partial^\mu A^\nu \\ &= \Lambda_\mu{}^a \Lambda_\nu{}^b \partial^\mu A^\nu - \Lambda_\mu{}^b \Lambda_\nu{}^a \partial^\mu A^\nu \\ &= \Lambda_\mu{}^a \Lambda_\nu{}^b (\partial^\mu A^\nu - \partial^\nu A^\mu).\end{aligned} \tag{D.28}$$

Which gives the final transformation rule for the field bivector in tensor form

$$F^{ab\prime} = \Lambda_\mu{}^a \Lambda_\nu{}^b F^{\mu\nu}. \tag{D.29}$$

Returning to the original problem of field Lagrangian invariance, we want to examine how $F^{ab'}F_{ab'}$ transforms. That is

$$\begin{aligned} F^{ab'}F_{ab'} &= \Lambda_\mu{}^a \Lambda_\nu{}^b F^{\mu\nu} \Lambda^\alpha{}_a \Lambda^\beta{}_b F_{\alpha\beta} \\ &= (\Lambda_\mu{}^a \Lambda^\alpha{}_a)(\Lambda_\nu{}^b \Lambda^\beta{}_b) F^{\mu\nu} F_{\alpha\beta} \\ &= \delta_\mu{}^\alpha \delta_\nu{}^\beta F^{\mu\nu} F_{\alpha\beta} \\ &= F^{\mu\nu} F_{\mu\nu}. \end{aligned} \tag{D.30}$$

which is the desired result. Since the dot product remainder of the Lagrangian eq. (D.1) has already been shown to be Lorentz invariant this is sufficient to prove the Lagrangian boost or rotational invariance using tensor algebra.

Working this way is fairly compact and efficient, and required a few less steps than the multivector equivalent. To compare apples to apples, for the algebraic tools, it should be noted that if only the scalar part of $(\nabla \wedge A)^2$ was considered as implicitly done in the tensor argument above, the multivector approach would likely have been as compact as well.

E

LORENTZ TRANSFORM NOETHER CURRENT (INTERACTION LAGRANGIAN).

E.1 MOTIVATION.

Here we consider Noether's theorem applied to the covariant form of the Lorentz force Lagrangian. Boost under rotation or boost or a combination of the two will be considered.

E.2 COVARIANT RESULT.

For proper velocity v, four potential A, and positive time metric signature $(\gamma_0)^2 = 1$, the Lorentz for Lagrangian is

$$L = \frac{1}{2} m v \cdot v + q A \cdot v / c. \tag{E.1}$$

Let us see if Noether's can be used to extract an invariant from the Lorentz force Lagrangian eq. (E.1) under a Lorentz boost or a spatial rotational transformation. Four vector dot products are Lorentz invariants. This can be thought of as the definition of a Lorentz transform (ie: the transformations that leave the four vector dot products unchanged). Alternatively, this can be shown using the exponential form of the boost

$$L(x) = \exp(-\alpha \hat{\mathbf{a}}/2) x \exp(\alpha \hat{\mathbf{a}}/2). \tag{E.2}$$

The dot product of two such transformed vectors is

$$\begin{aligned} L(x) \cdot L(y) &= \langle \exp(-\alpha \hat{\mathbf{a}}/2) x \exp(\alpha \hat{\mathbf{a}}/2) \exp(-\alpha \hat{\mathbf{a}}/2) y \exp(\alpha \hat{\mathbf{a}}/2) \rangle \\ &= \langle \exp(-\alpha \hat{\mathbf{a}}/2) x y \exp(\alpha \hat{\mathbf{a}}/2) \rangle \\ &= x \cdot y \langle \exp(-\alpha \hat{\mathbf{a}}/2) \exp(\alpha \hat{\mathbf{a}}/2) \rangle \\ &= x \cdot y. \end{aligned} \tag{E.3}$$

Using the exponential form of the boost operation, boosting v, A leaves the Lagrangian unchanged. Therefore there is a conserved quantity according to Noether's, but what is it?

Also observe that the spacetime nature of the bivector $\hat{\mathbf{a}}$ has not actually been specified, which means that all the subsequent results apply to spatial rotation as well. Due to the negative spatial signature ($(\gamma_i)^2 = -1$) used here, for a spatial rotation α will represent a rotation in the negative sense in the oriented plane specified by the unit bivector $\hat{\mathbf{a}}$.

Consider change with respect to the rapidity factor (or rotational angle) α

$$\frac{\partial L'}{\partial \alpha} = \frac{d}{d\tau}\left(\frac{\partial x'}{\partial \alpha} \cdot \nabla_{v'} L\right). \tag{E.4}$$

The boost spacetime plane (or rotational plane) $\hat{\mathbf{a}}$ could also be considered a parameter in the transformation, but to use that or the combination of the two we need the multivector form of Noether's. These notes were in fact originally part of an attempt 8 to get a feeling for the scalar case as lead up to that so this is an exercise for later.

As for the derivatives in eq. (E.4) we have

$$\begin{aligned}\frac{\partial x'}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \exp(-\alpha\hat{\mathbf{a}}/2) x \exp(\alpha\hat{\mathbf{a}}/2) \\ &= -\frac{1}{2}\left(\hat{\mathbf{a}}x' - x'\hat{\mathbf{a}}\right) \\ &= -\hat{\mathbf{a}} \cdot x'.\end{aligned} \tag{E.5}$$

$$\nabla_{v'} L = p' + qA'/c. \tag{E.6}$$

So the conserved quantity is

$$\begin{aligned}-(\hat{\mathbf{a}} \cdot x') \cdot \left(p' + qA'/c\right) &= -\hat{\mathbf{a}} \cdot (x' \wedge (p' + qA'/c)) \\ &= -\hat{\mathbf{a}} \cdot \kappa.\end{aligned} \tag{E.7}$$

So we have a conserved quantity

$$x \wedge (p + qA/c) = \kappa. \tag{E.8}$$

This has the looks of the three dimensional angular momentum conservation expression (with an added term due to non-radial potential), but does not look like any quantity from relativistic texts that I have seen (not that I have really seen too much).

As an example to get a feeling for this take x to be a rest frame world-line. Then we have

$$ct\gamma_0 \wedge (m\dot{t}\gamma_0 + qA/c) = -qt\mathbf{A} = \kappa. \tag{E.9}$$

Which indicates that the product of observer time and the observers' three vector potential is a constant of motion. Curious. Not a familiar result.

Assuming these calculations are correct, then if this holds for all time for then $\kappa = 0$ due to the origin time of x. I would interpret this to mean that for the charged mass to be at rest, the vector potential must also be zero. So while $x = ct\gamma_0$ is simple for calculations, it does not appear to be a terribly interesting case.

FIXME: try plugging in specific solutions to the Lorentz force equation here to validate or invalidate this calculation.

One further thing that can be observed about this is that if we take derivatives of

$$x \wedge (p + qA/c) = \kappa, \tag{E.10}$$

we have

$$v \wedge (p + qA/c) + x \wedge (\dot{p} + q\dot{A}/c) = 0. \tag{E.11}$$

Or

$$\begin{aligned} x \wedge \dot{p} &= \frac{d}{d\tau}\left(q/cA \wedge x\right) \\ &= qA \wedge v/c + q/c\dot{A} \wedge x. \end{aligned} \tag{E.12}$$

So we have a relativistic torque expressed in terms of the potential, proper velocity and the variation of the potential.

E.3 EXPANSION IN OBSERVER FRAME.

This still is not familiar looking, but lets expand this in terms of a particular observable, and see what falls out. First the LHS, with $dt/d\tau = \gamma$

$$x \wedge \dot{p} = (ct\gamma_0 + x^i\gamma_i) \wedge \left(\gamma \frac{d}{dt}\left(m\gamma(c\gamma_0 + \frac{dx^j}{dt}\gamma_j)\right)\right). \tag{E.13}$$

So

$$\frac{1}{\gamma}(x \wedge \dot{p}) = -ct\frac{d(\gamma\mathbf{p})}{dt} + \mathbf{x}\frac{d(mc\gamma)}{dt} + x^i\gamma_i \wedge \frac{d}{dt}\left(m\gamma\frac{dx^j}{dt}\gamma_j\right). \tag{E.14}$$

But

$$\begin{aligned}\sigma_i \wedge \sigma_j &= \frac{1}{2}(\gamma_i\gamma_0\gamma_j\gamma_0 - \gamma_j\gamma_0\gamma_i\gamma_0) \\ &= -\frac{(\gamma_0)^2}{2}(\gamma_i\gamma_j - \gamma_j\gamma_i) \\ &= -\gamma_i \wedge \gamma_j.\end{aligned} \tag{E.15}$$

for

$$\frac{1}{\gamma}(x \wedge \dot{p}) = -ct\frac{d(\gamma\mathbf{p})}{dt} + \mathbf{x}\frac{d(mc\gamma)}{dt} - \mathbf{x} \wedge \frac{d(\gamma\mathbf{p})}{dt}. \tag{E.16}$$

Now, for the RHS of eq. (E.12), with $A^0 = \phi$

$$\begin{aligned}\frac{q}{c}\gamma\frac{d(x \wedge A)}{dt} &= \frac{q}{c}\gamma\frac{d}{dt}(ct\gamma_0 + x^i\gamma_i) \wedge (\phi\gamma_0 + A^j\gamma_j) \\ &= \frac{q}{c}\gamma\frac{d}{dt}\left(-ct\mathbf{A} + \phi\mathbf{x} - \mathbf{x} \wedge \mathbf{A}\right).\end{aligned} \tag{E.17}$$

Equating the vector and bivector parts, and employing a duality transformation for the bivector parts leaves two vector relationships

$$ct\frac{d(\gamma\mathbf{p})}{dt} - \mathbf{x}\frac{d(mc\gamma)}{dt} = \frac{q}{c}\frac{d\left(ct\mathbf{A} - \phi\mathbf{x}\right)}{dt}, \tag{E.18}$$

$$\mathbf{x} \times \frac{d(\gamma\mathbf{p})}{dt} = \frac{q}{c}\frac{d}{dt}\left(\mathbf{x} \times \mathbf{A}\right). \tag{E.19}$$

FIXME: the first equation looks like it could also be expressed in some sort more symmetric form. Perhaps a grade two (commutator) product between the multivectors $(mc\gamma, \mathbf{p}) = p\gamma_0$, and $(\phi, \mathbf{A}) = A\gamma_0$?

E.4 IN TENSOR FORM.

As can be seen above, the four vector form of eq. (E.12) is much more symmetric. What does it look like in tensor form? After first re-consolidating the proper time derivatives we can read the coordinate form off by inspection

$$x \wedge \dot{p} = \frac{d}{d\tau}\left(q/cA \wedge x\right). \tag{E.20}$$

$$\gamma_\mu \wedge \gamma_\nu x^\mu m v^\nu = \frac{d}{d\tau}\left(q/cA^\alpha x^\beta\right)\gamma_\alpha \wedge \gamma_\beta. \tag{E.21}$$

Which gives the tensor expression

$$\epsilon_{\mu\nu}\left(x^\mu v^\nu - \frac{d}{d\tau}\left(\frac{q}{mc}A^\mu x^\nu\right)\right) = 0. \tag{E.22}$$

This in turn implies the following six equations in μ, and ν

$$x^\mu v^\nu - x^\nu v^\mu = \frac{q}{mc}\frac{d}{d\tau}\left(A^\mu x^\nu - A^\nu x^\mu\right). \tag{E.23}$$

Looking to see if I got the right result, I asked on PF, and was pointed to [1]. That ASCII thread is hard to read but at least my result is similar. I will have to massage things to match them up more closely.

What I did not realize until I read that is that my rotation was not fixed as either hyperbolic or euclidean since I did not actually specify the specific nature of the bivector for the rotational plane. So I ended up with results for both the spatial invariance and the boost invariance at the same time. Have adjusted things above, but that is why the spatial rotation references all appear as afterthoughts.

Of the six equations in eq. (E.23), taking space time indices yields the vector eq. (E.18) as the conserved quantity for a boost. Similarly the second vector result in eq. (E.19) for purely spatial indices is the conserved quantity for spatial rotation. That makes my result seem more reasonable since I did not expect to get so much only considering boost.

F

CANONICAL ENERGY MOMENTUM TENSOR AND TRANSLATION.

F.1 MOTIVATION AND DIRECTION.

In [11] we saw that it was possible to express the Lorentz force equation for the charge per unit volume in terms of the energy momentum tensor.

Repeating

$$\begin{aligned}\nabla \cdot T(\gamma_\mu) &= \frac{1}{c}\left\langle F\gamma_\mu J\right\rangle \\ T(a) &= \frac{\epsilon_0}{2} F a \tilde{F}.\end{aligned} \tag{F.1}$$

While these may not appear too much like the Lorentz force equation as we are used to seeing it, with some manipulation we found

$$\begin{aligned}\frac{1}{c}\langle F\gamma_0 J\rangle &= -\mathbf{j}\cdot\mathbf{E} \\ \frac{1}{c}\langle F\gamma_k J\rangle &= (\rho\mathbf{E} + \mathbf{j}\times\mathbf{B})\cdot\sigma_k,\end{aligned} \tag{F.2}$$

where we now have an energy momentum pair of equations, the second of which if integrated over a volume is the Lorentz force for the charge in that volume. We have also seen that we can express the Lorentz force equation in GA form

$$m\ddot{x} = qF\cdot\dot{x}/c. \tag{F.3}$$

This was expressed in tensor form, toggling indices that was

$$m\ddot{x}_\mu = qF_{\mu\alpha}\dot{x}^\alpha. \tag{F.4}$$

We then saw in [10] that the covariant form of the energy momentum tensor relation was

$$\begin{aligned}T^{\mu\nu} &= \epsilon_0\left(F^{\alpha\mu}F^\nu{}_\alpha + \frac{1}{4}F^{\alpha\beta}F_{\alpha\beta}\eta^{\mu\nu}\right) \\ \partial_\nu T^{\mu\nu} &= F^{\alpha\mu}J_\alpha/c.\end{aligned} \tag{F.5}$$

This has identical structure (FIXME: sign error here?) to the covariant Lorentz force equation.

Now the energy momentum conservation equations above did not require the Lorentz force equations at all for their derivation, nor have we used the Lorentz force interaction Lagrangian to arrive at them. With Maxwell's equation and the Lorentz force equation together (or the equivalent field and interaction Lagrangians) we have the complete specification of classical electrodynamics. Curiously it appears that we have most of the structure of the Lorentz force equation (except for the association with mass) all in embedded in Maxwell's equation or the Maxwell field Lagrangian.

Now, a proper treatment of the field and charged mass interaction likely requires the Dirac Lagrangian, and hiding in there if one could extract it, is probably everything that could be said on the topic. It will be a long journey to get to that point, but how much can we do considering just the field Lagrangian?

For these reasons it seems desirable to understand the background behind the energy momentum tensor much better. In particular, it is natural to then expect that these conservation relations may also be found as a consequence of a symmetry and an associated Noether current (see 18.1). What is that symmetry? That symmetry should leave the field equations as calculated by the field Euler-Lagrange equations Given that symmetry how would one go about actually showing that this is the case? These are the questions to tackle here.

F.2 ON TRANSLATION AND DIVERGENCE SYMMETRIES.

F.2.1 *Symmetry due to total derivative addition to the Lagrangian.*

In [2] the energy momentum tensor is treated by considering spacetime translation, but I have unfortunately not understood much more than vague direction in that treatment.

In [23] it is also stated that the energy momentum tensor is the result of a Lagrangian spacetime translation, but I did not find details there.

There are examples of the canonical energy momentum tensor (in the simpler non-GA tensor form) and the symmetric energy momentum tensor in [7]. However, that treatment relies on analogy with mechanical form of Noether's theorem, and I had rather see it developed explicitly.

Finally, in an unexpected place (since I am not studying QFT but was merely curious), the clue required to understand the details of how this spacetime translation results in the energy momentum tensor was found in [25].

In Tong's treatment it is pointed out there is a symmetry for the Lagrangian if it is altered by a divergence

$$\mathcal{L} \to \mathcal{L} + \partial_\mu F^\mu. \tag{F.6}$$

It took me a while to figure out how this was a symmetry, but after a nice refreshing motorcycle ride, the answer suddenly surfaced. One can add a derivative to a mechanical Lagrangian and not change the resulting equations of motion. While tackling problem 5 of Tong's mechanics in 4.2, such an invariance was considered in detail in one of the problems for Tong's classical mechanics notes 5 . If one has altered the Lagrangian by adding an arbitrary function f to it.

$$\mathcal{L}' = \mathcal{L} + f. \tag{F.7}$$

Assuming to start a Lagrangian that is a function of a single field variable $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi)$, then the variation of the Lagrangian for the field equations yields

$$\begin{aligned}\frac{\delta \mathcal{L}'}{\delta \phi} &= \frac{\partial \mathcal{L}'}{\partial \phi} - \partial_\sigma \frac{\partial \mathcal{L}'}{\partial(\partial_\sigma \phi)} \\ &= \underbrace{\boxed{\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\sigma \phi)}}}_{=0} + \frac{\partial f}{\partial \phi} - \partial_\sigma \frac{\partial f}{\partial(\partial_\sigma \phi)}.\end{aligned} \tag{F.8}$$

So, if this transformed Lagrangian is a symmetry, it is sufficient to find the conditions for the variation of additional part to be zero

$$\frac{\delta f}{\delta \phi} = 0. \tag{F.9}$$

F.2.2 *Some examples adding a divergence.*

To validate the fact that we can add a divergence to the Lagrangian without changing the field equations lets work out a few concrete examples of eq. (F.9) of for Lagrangian alterations by a divergence $f = \partial_\mu F^\mu$.

Each of these examples will be for a single field variable Lagrangian with generalized coordinates $x^1 = x$, and $x^1 = y$.

F.2.2.1 *Simplest case. No partials.*

Let

$$\begin{aligned} F^1 &= \phi \\ F^2 &= 0. \end{aligned} \tag{F.10}$$

With this the divergence is

$$\begin{aligned} f &= \partial_x F^x + \partial_y F^y \\ &= \frac{\partial \phi}{\partial x}. \end{aligned} \tag{F.11}$$

Now the variation is

$$\begin{aligned} \frac{\delta f}{\delta \phi} &= \left(\frac{\partial}{\partial \phi} - \frac{\partial}{\partial x} \frac{\partial}{\partial(\partial \phi/\partial x)} - \frac{\partial}{\partial y} \frac{\partial}{\partial(\partial \phi/\partial y)} \right) \frac{\partial \phi}{\partial x} \\ &= \frac{\partial}{\partial x} \frac{\partial \phi}{\partial \phi} - \frac{\partial 1}{\partial x} \\ &= 0. \end{aligned} \tag{F.12}$$

Okay, so far so good.

F.2.2.2 *One partial.*

Now, let

$$\begin{aligned} F^1 &= \frac{\partial \phi}{\partial x} \\ F^2 &= 0. \end{aligned} \tag{F.13}$$

With this the divergence is

$$\begin{aligned} f &= \partial_x F^x + \partial_y F^y \\ &= \frac{\partial}{\partial x} \frac{\partial \phi}{\partial x}. \end{aligned} \tag{F.14}$$

And the variation is

$$\begin{aligned} \frac{\delta f}{\delta \phi} &= \left(\frac{\partial}{\partial \phi} - \frac{\partial}{\partial x} \frac{\partial}{\partial(\partial \phi/\partial x)} - \frac{\partial}{\partial y} \frac{\partial}{\partial(\partial \phi/\partial y)} \right) \frac{\partial}{\partial x} \frac{\partial \phi}{\partial x} \\ &= \frac{\partial}{\partial \phi} \frac{\partial}{\partial x} \frac{\partial \phi}{\partial x} \\ &= \frac{\partial}{\partial x} \frac{\partial}{\partial x} \frac{\partial \phi}{\partial \phi} \\ &= \frac{\partial}{\partial x} \frac{\partial 1}{\partial x} \\ &= 0. \end{aligned} \tag{F.15}$$

Again, assuming I am okay to switch the differentiation order, we have zero.

F.2.2.3 *Another partial.*

For the last concrete example before going on to the general case, try

$$\begin{aligned} F^1 &= \frac{\partial \phi}{\partial y} \\ F^2 &= 0. \end{aligned} \tag{F.16}$$

The divergence is

$$\begin{aligned} f &= \partial_x F^x + \partial_y F^y \\ &= \frac{\partial}{\partial x}\frac{\partial \phi}{\partial y}. \end{aligned} \tag{F.17}$$

And the variation is

$$\begin{aligned} \frac{\delta f}{\delta \phi} &= \left(\frac{\partial}{\partial \phi} - \frac{\partial}{\partial x}\frac{\partial}{\partial(\partial\phi/\partial x)} - \frac{\partial}{\partial y}\frac{\partial}{\partial(\partial\phi/\partial y)} \right) \frac{\partial}{\partial x}\frac{\partial \phi}{\partial y} \\ &= -\frac{\partial}{\partial y}\frac{\partial 1}{\partial x} \\ &= 0. \end{aligned} \tag{F.18}$$

F.2.2.4 *The general case.*

Because of linearity we have now seen that we can construct functions with any linear combinations of first and second derivatives

$$F^\mu = a^\mu \phi + \sum_\sigma b_\sigma{}^\mu \frac{\partial \phi}{\partial x^\sigma}. \tag{F.19}$$

and for such a function we will have

$$\frac{\delta(\partial_\mu F^\mu)}{\delta \phi} = 0. \tag{F.20}$$

How general can the function $F^\mu = F^\mu(\phi, \partial_\sigma \phi)$ be made and still yield a zero variational derivative?

To answer this, let us compute the derivative for a general divergence added to a single field variable Lagrangian. This is

$$\begin{aligned}
\frac{\delta(\partial_\mu F^\mu)}{\delta\phi} &= \sum_\mu \left(\frac{\partial}{\partial\phi} - \sum_\sigma \frac{\partial}{\partial x^\sigma} \frac{\partial}{\partial(\partial\phi/\partial x^\sigma)} \right) \frac{\partial F^\mu}{\partial x^\mu} \\
&= \sum_\mu \frac{\partial}{\partial x^\mu} \frac{\partial F^\mu}{\partial\phi} \\
&\quad - \sum_{\mu,\sigma} \frac{\partial}{\partial x^\sigma} \frac{\partial}{\partial(\partial\phi/\partial x^\sigma)} \left(\frac{\partial F^\mu}{\partial\phi} \frac{\partial\phi}{\partial x^\mu} + \sum_\alpha \frac{\partial F^\mu}{\partial(\partial\phi/\partial x^\alpha)} \frac{\partial(\partial\phi/\partial x^\alpha)}{\partial x^\mu} \right) \\
&= \partial_\mu \frac{\partial F^\mu}{\partial\phi} - \partial_\sigma \frac{\partial}{\partial(\partial_\sigma\phi)} \left(\frac{\partial F^\mu}{\partial\phi} \partial_\mu\phi + \frac{\partial F^\mu}{\partial(\partial_\alpha\phi)} \partial_{\mu\alpha}\phi \right).
\end{aligned} \tag{F.21}$$

For tractability in this last line the shorthand for the partials has been injected. Sums over α, μ, and σ are also now implied (this was made explicit prior to this in all cases where upper and lower indices were matched).

Treating these two last derivatives separately, we have for the first

$$\begin{aligned}
\partial_\sigma \frac{\partial}{\partial(\partial_\sigma\phi)} \frac{\partial F^\mu}{\partial\phi} \partial_\mu\phi &= \partial_\sigma \left(\frac{\partial}{\partial(\partial_\sigma\phi)} \frac{\partial F^\mu}{\partial\phi} \right) \partial_\mu\phi + \partial_\sigma \frac{\partial F^\mu}{\partial\phi} \frac{\partial}{\partial(\partial_\sigma\phi)} \partial_\mu\phi \\
&= \partial_\sigma \left(\frac{\partial}{\partial(\partial_\sigma\phi)} \frac{\partial F^\mu}{\partial\phi} \right) \partial_\mu\phi + \partial_\mu \frac{\partial F^\mu}{\partial\phi}.
\end{aligned} \tag{F.22}$$

So our $\partial F^\mu/\partial\phi$'s cancel out, and we are left with

$$\begin{aligned}
\frac{\delta(\partial_\mu F^\mu)}{\delta\phi} &= -\partial_\sigma \left(\left(\frac{\partial}{\partial(\partial_\sigma\phi)} \frac{\partial F^\mu}{\partial\phi} \right) \partial_\mu\phi + \frac{\partial}{\partial(\partial_\sigma\phi)} \left(\frac{\partial F^\mu}{\partial(\partial_\alpha\phi)} \partial_{\mu\alpha}\phi \right) \right) \\
&= -\partial_\sigma \left(\partial_\mu\phi \left(\frac{\partial}{\partial(\partial_\sigma\phi)} \frac{\partial F^\mu}{\partial\phi} \right) + \partial_{\mu\alpha}\phi \frac{\partial}{\partial(\partial_\sigma\phi)} \left(\frac{\partial F^\mu}{\partial(\partial_\alpha\phi)} \right) \right) \\
&= -\partial_\sigma \left((\partial_\mu\phi) \frac{\partial}{\partial\phi} \frac{\partial}{\partial(\partial_\sigma\phi)} F^\mu + \left(\partial_\mu \frac{\partial\phi}{\partial x^\alpha} \right) \frac{\partial}{\partial(\partial_\alpha\phi)} \frac{\partial}{\partial(\partial_\sigma\phi)} F^\mu \right).
\end{aligned} \tag{F.23}$$

Now there is a lot of indices and derivatives floating around. Writing $g^\mu = \partial F^\mu/\partial(\partial_\sigma\phi)$, we have something a bit easier to look at

$$\frac{\delta(\partial_\mu F^\mu)}{\delta\phi} = -\partial_\sigma \left((\partial_\mu\phi) \frac{\partial g^\mu}{\partial\phi} + \left(\partial_\mu \frac{\partial\phi}{\partial x^\alpha} \right) \frac{\partial g^\mu}{\partial(\partial_\alpha\phi)} \right). \tag{F.24}$$

But this is a chain rule expansion of the derivative $\partial_\mu g^\mu$

$$\frac{\partial g^\mu}{\partial x^\mu} = \frac{\partial \phi}{\partial x^\mu}\frac{\partial g^\mu}{\partial \phi} + \frac{\partial \partial_\beta \phi}{\partial x^\mu}\frac{\partial g^\mu}{\partial \partial_\beta \phi}. \tag{F.25}$$

So, we finally have

$$\frac{\delta(\partial_\mu F^\mu)}{\delta \phi} = -\partial_{\sigma\mu} g^\mu. \tag{F.26}$$

This is

$$\frac{\delta(\partial_\mu F^\mu)}{\delta \phi} = -\partial_{\sigma\mu}\frac{\partial F^\mu}{\partial(\partial_\sigma \phi)}. \tag{F.27}$$

I do not think we have any right asserting that this is zero for arbitrary F^μ. However if the Taylor expansion of F^μ with respect to variables ϕ, and $\partial_\sigma \phi$ has no higher than first order terms in the field variables $\partial_\sigma \phi$, we will certainly have a zero variational derivative and a corresponding symmetry.

F.2.2.5 *More examples to confirm the symmetry requirements.*

As a confirmation that a zero in eq. (F.27) requires linear field derivatives, lets try two more example calculations.

First with non-linear powers of ϕ to show that we have more freedom to construct the function first powers. Let

$$\begin{aligned} F^1 &= \phi^2 \\ F^2 &= 0. \end{aligned} \tag{F.28}$$

We have

$$\begin{aligned} \frac{\delta(\partial_\mu F^\mu)}{\delta \phi} &= \left(\frac{\partial}{\partial \phi} - \partial_\sigma \frac{\partial}{\partial(\partial_\sigma \phi)}\right) 2\phi\phi_x \\ &= 2\phi_x - \partial_x(2\phi) \\ &= 0. \end{aligned} \tag{F.29}$$

Zero as expected. Generalizing the function to include arbitrary polynomial powers is no harder. Let

$$\begin{aligned} F^1 &= \phi^k \\ F^2 &= 0 \\ \partial_\mu F^\mu &= k\phi^{k-1}\phi_x. \end{aligned} \tag{F.30}$$

So we have

$$\begin{aligned}\frac{\delta(\partial_\mu F^\mu)}{\delta\phi} &= k(k-1)\phi^{k-2}\phi_x - \partial_x(k\phi^{k-1}) \\ &= 0.\end{aligned} \tag{F.31}$$

Okay, now moving on to the derivatives. Picking a divergence that should not will not generate a symmetry, something with a non-linear derivative should do the trick. Let us Try

$$\begin{aligned}F^1 &= (\phi_x)^2 \\ F^2 &= 0.\end{aligned} \tag{F.32}$$

$$\begin{aligned}\frac{\delta(\partial_\mu F^\mu)}{\delta\phi} &= \left(\frac{\partial}{\partial\phi} - \partial_\sigma \frac{\partial}{\partial(\partial_\sigma\phi)}\right) 2\phi_x\phi_{xx} \\ &= -2\partial_x\phi_{xx} \\ &= -2\phi_{xxx}.\end{aligned} \tag{F.33}$$

So, sure enough, unless additional conditions can be imposed on ϕ, such a transformation will not be a symmetry.

F.2.3 *Symmetry for Wave equation under spacetime translation.*

The Lagrangian for a one dimensional wave equation is

$$\mathcal{L} = \frac{1}{2v^2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2. \tag{F.34}$$

Under a transformation of variables

$$\begin{aligned}x &\to x' = x + a \\ t &\to t' = t + \tau.\end{aligned} \tag{F.35}$$

Employing a multivariable Taylor expansion (see [8]) for our Lagrangian having no explicit dependence on t and x, we have

$$\mathcal{L}' = \mathcal{L} + \overset{*}{\boxed{(a\partial_x + \tau\partial_t)\mathcal{L}}} + \cdots \tag{F.36}$$

That first order term of the Taylor expansion $*$, can be written as a divergence $\partial_\mu F^\mu$, with $F^1 = a\mathcal{L}$, and $F^2 = \tau\mathcal{L}$, however both of these are

quadratic in ϕ_x, and ϕ_t, which is not linear. That linearity in the derivatives was required for eq. (F.27) to be definitively zero for the transformation to be a symmetry. So after all that goofing around with derivatives and algebra it is defeated by the simplest field Lagrangian.

Now, if we continue we find that we do in fact still have a symmetry by introducing a linearized spacetime translation. This follows from direct expansion

$$\begin{aligned}(*) &= (a\partial_x + \tau\partial_t)\mathcal{L} \\ &= a\left(\frac{1}{v^2}\phi_t\partial_x\phi_t - \phi_x\partial_x\phi_x\right) + \tau\left(\frac{1}{v^2}\phi_t\partial_t\phi_t - \phi_x\partial_t\phi_x\right).\end{aligned} \tag{F.37}$$

Next, calculation of the variational derivative we have

$$\begin{aligned}\frac{\delta(*)}{\delta\phi} &= \left(\frac{\partial}{\partial\phi} - \partial_x\frac{\partial}{\partial\phi_x} - \partial_t\frac{\partial}{\partial\phi_t}\right)(*) \\ &= -\partial_x\left(-a\partial_{xx}\phi - \tau\partial_{tx}\phi\right) - \frac{1}{v^2}\partial_t\left(a\partial_{xt}\phi + \tau\partial_{tt}\phi\right) \\ &= a\left(\partial_x\left(\phi_{xx} - \frac{1}{v^2}\phi_{tt}\right)\right) + \tau\left(\partial_t\left(\phi_{xx} - \frac{1}{v^2}\phi_{tt}\right)\right).\end{aligned} \tag{F.38}$$

Since we have $\phi_{xx} = \frac{1}{v^2}\phi_{tt}$ by variation of eq. (F.34). So we do in fact have a symmetry from the linearized spacetime translation for any shift $(t, x) \to (t + \tau, x + a)$.

F.2.4 *Symmetry condition for arbitrary linearized spacetime translation.*

If we want to be able to alter the Lagrangian with a linearized vector translation of the generalized coordinates by some arbitrary shift, since we do not have the linear derivatives for many Lagrangians of interest (wave equations, Maxwell equation, ...) then can we find a general condition that is responsible for the translation symmetry that we have observed must exist for the simple wave equation.

For a general Lagrangian $\mathcal{L} = \mathcal{L}(\phi(x), \partial_\mu\phi(x))$ under shift by some vector a

$$x \to x' = x + a, \tag{F.39}$$

we have

$$\mathcal{L}' = \left(e^{a\cdot\nabla}\right)\mathcal{L} = \mathcal{L} + (a\cdot\nabla)\mathcal{L} + \frac{1}{2!}(a\cdot\nabla)^2\mathcal{L} + \cdots \tag{F.40}$$

Now, if we have

$$\frac{\delta((a\cdot\nabla)\mathcal{L})}{\delta\phi} \stackrel{?}{=} (a\cdot\nabla)\overbrace{\boxed{\frac{\delta\mathcal{L}}{\delta\phi}}}^{=0}. \tag{F.41}$$

then this would explain the fact that we have a symmetry under linearized translation for the wave equation Lagrangian. Can this interchange of differentiation order be justified?

Writing out this variational derivative in full we have

$$\begin{aligned}\frac{\delta((a\cdot\nabla)\mathcal{L})}{\delta\phi} &= \left(\frac{\partial}{\partial\phi} - \partial_\sigma\frac{\partial}{\partial\phi_\sigma}\right)a^\mu\partial_\mu\mathcal{L} \\ &= a^\mu\left(\frac{\partial}{\partial\phi}\frac{\partial}{\partial x^\mu} - \frac{\partial}{\partial x^\sigma}\frac{\partial}{\partial\phi_\sigma}\frac{\partial}{\partial x^\mu}\right)\mathcal{L}.\end{aligned} \tag{F.42}$$

Now, one can impose continuity conditions on the field variables and Lagrangian sufficient to allow the commutation of the coordinate partials. Namely

$$\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}f(\phi,\partial_\sigma\phi) = \frac{\partial}{\partial x^\nu}\frac{\partial}{\partial x^\mu}f(\phi,\partial_\sigma\phi). \tag{F.43}$$

However, we have a dependence between the field variables and the coordinates

$$\frac{\partial}{\partial x^\mu} = \frac{\partial\phi}{\partial x^\mu}\frac{\partial}{\partial\phi} + \sum_\sigma\frac{\partial\phi_\sigma}{\partial x^\mu}\frac{\partial}{\partial\phi_\sigma}. \tag{F.44}$$

Given this, can we commute the field partials and the coordinate partials like so

$$\begin{aligned}\frac{\partial}{\partial\phi}\frac{\partial}{\partial x^\mu} &\stackrel{?}{=} \frac{\partial}{\partial x^\mu}\frac{\partial}{\partial\phi} \\ \frac{\partial}{\partial\phi_\sigma}\frac{\partial}{\partial x^\mu} &\stackrel{?}{=} \frac{\partial}{\partial x^\mu}\frac{\partial}{\partial\phi_\sigma}.\end{aligned} \tag{F.45}$$

This is not obvious to me due to the dependence between the two.

If that is a reasonable thing to do, then the variational derivative of this directional derivative is zero

$$\begin{aligned}\frac{\delta((a\cdot\nabla)\mathcal{L})}{\delta\phi} &= a^\mu\frac{\partial}{\partial x^\mu}\left(\frac{\partial}{\partial\phi} - \frac{\partial}{\partial x^\sigma}\frac{\partial}{\partial\phi_\sigma}\right)\mathcal{L} \\ &= (a\cdot\nabla)\frac{\delta\mathcal{L}}{\delta\phi} \\ &= 0.\end{aligned} \tag{F.46}$$

To make any progress below I had to assume that this is justifiable. With this assumption or requirement we therefore have a symmetry for any Lagrangian altered by the addition of a directional derivative, as is required for the first order Taylor series approximation associated with a spacetime (or spatial or timelike) translation.

F.2.4.1 *An error above to revisit.*

In an email discussing what I initially thought was a typo in [25], he says that while it is correct to transform the Lagrangian using a Taylor expansion in $\phi(x+a)$ as I have done, this actually results from $x \to x - a$, as opposed to the positive shift given in eq. (F.39). There was discussion of this in the context of Lorentz transformations around (1.26) of his QFT course notes, also applicable to translations. The subtlety is apparently due to differences between passive and active transformations. I am sure he is right, and I think this is actually consistent with the treatment of [2] where they include an inverse operation in the transformed Lagrangian (that minus is surely associated with the inverse of the translation transformation). It will take further study for me to completely understand this point, but provided the starting point is really considered the Taylor series expansion based on $\phi(x) \to \phi(x+a)$ and not based on eq. (F.39) then nothing else I have done here is wrong. Also note that in the end our Noether current can be adjusted by an arbitrary multiplicative constant so the direction of the translation will also not change the final result.

F.3 NOETHER CURRENT.

F.3.1 *Vector parametrized Noether current.*

In **??** the derivation of Noether's theorem given a single variable parametrized alteration of the Lagrangian was seen to essentially be an exercise in the application of the chain rule.

How to extend that argument to the multiple variable case is not immediately obvious. In GA we can divide by vectors but attempting to formulate a derivative this way gives us left and right sided derivatives. How do we overcome this to examine change of the Lagrangian with respect to a vector parametrization? One possibility is a scalar parametrization of

the magnitude of the translation vector. If the translation is along $a = \alpha u$, where u is a unit vector we can write

$$\begin{aligned}\mathcal{L}' &= \mathcal{L} + \delta\mathcal{L} \\ &= \mathcal{L} + (a \cdot \nabla)\mathcal{L} \\ &= \mathcal{L} + \alpha(u \cdot \nabla)\mathcal{L}.\end{aligned} \tag{F.47}$$

So we have

$$\frac{d\mathcal{L}'}{d\alpha} = (u \cdot \nabla)\mathcal{L}. \tag{F.48}$$

Now our previous Noether's current was derived by considering just the sort of derivative on the LHS above, but on the RHS we are back to working with a directional derivative. The key is finding a logical starting point for the chain rule like expansion that we expect to produce the conservation current.

$$\begin{aligned}\delta\mathcal{L} &= (a \cdot \nabla)\mathcal{L} \\ &= a^\mu \partial_\mu \mathcal{L} \\ &= a^\mu \left(\frac{\partial \phi}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi} + \sum_\sigma \frac{\partial \phi_\sigma}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_\sigma} \right) \\ &= \frac{\partial \mathcal{L}}{\partial \phi}(a \cdot \nabla)\phi + \sum_\sigma \frac{\partial \mathcal{L}}{\partial \phi_\sigma}(a \cdot \nabla)\phi_\sigma \\ &= \left(\sum_\sigma \partial_\sigma \frac{\partial \mathcal{L}}{\partial \phi_\sigma} \right)(a \cdot \nabla)\phi + \sum_\sigma \frac{\partial \mathcal{L}}{\partial \phi_\sigma}(a \cdot \nabla)\phi_\sigma \\ &= \left(\sum_\sigma \partial_\sigma \frac{\partial \mathcal{L}}{\partial \phi_\sigma} \right)(a \cdot \nabla)\phi + \sum_\sigma \frac{\partial \mathcal{L}}{\partial \phi_\sigma}\partial_\sigma((a \cdot \nabla)\phi) \\ &= \sum_\sigma \partial_\sigma \left(\frac{\partial \mathcal{L}}{\partial \phi_\sigma}(a \cdot \nabla)\phi \right).\end{aligned} \tag{F.49}$$

So far so good, but where to go from here? The trick (again from Tong) is that the difference with itself is zero. With a switch of dummy indices $\sigma \to \mu$, we have

$$\begin{aligned}0 &= \delta\mathcal{L} - \delta\mathcal{L} \\ &= \sum_\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \phi_\mu}(a \cdot \nabla)\phi \right) - a^\mu \partial_\mu \mathcal{L} \\ &= \sum_\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \phi_\mu}(a \cdot \nabla)\phi - a^\mu \mathcal{L} \right).\end{aligned} \tag{F.50}$$

Now we have a quantity that is zero for any vector a, and can say we have a conserved current $T(a)$ with coordinates

$$T^\mu(a) = \frac{\partial \mathcal{L}}{\partial \phi_\mu}(a \cdot \nabla)\phi - a^\mu \mathcal{L}. \tag{F.51}$$

Finally, putting this back into vector form

$$\begin{aligned} T(a) &= \gamma_\mu T^\mu(a) \\ &= \left(\gamma_\mu \frac{\partial \mathcal{L}}{\partial \phi_\mu}\right)(a \cdot \nabla)\phi - \gamma_\mu a^\mu \mathcal{L}. \end{aligned} \tag{F.52}$$

So we have

$$\begin{aligned} T(a) &= \left(\left(\gamma_\mu \frac{\partial}{\partial \phi_\mu}\right)\mathcal{L}\right)(a \cdot \nabla)\phi - a\mathcal{L} \\ \nabla \cdot T(a) &= 0. \end{aligned} \tag{F.53}$$

So after a long journey, I have in eq. (F.53) a derivation of a conservation current associated with a linearized vector displacement of the generalized coordinates. I recalled that the treatment in [2] somehow eliminated the a. That argument is still tricky involving their linear operator theory, but I have at least obtained their equation (13.15). They treat a multivector displacement whereas I only looked at vector displacement. They also do it in three lines, whereas building up to this (or even understanding it) based on what I know required 13 pages.

F.3.2 *Comment on the operator above.*

We have something above that is gradient like in eq. (F.53). Our spacetime gradient operator is

$$\nabla = \gamma^\mu \frac{\partial}{\partial x^\mu}. \tag{F.54}$$

Whereas this unknown field variable derivative operator

$$\text{something} = \gamma_\mu \frac{\partial}{\partial \phi_\mu}. \tag{F.55}$$

is somewhat like a velocity gradient with respect to the field variable. It would be reasonable to expect that this will have a role in the field canonical momentum.

F.3.3 *In tensor form.*

The conserved current of eq. (F.53) can be put into tensor form by considering the action on each of the basis vectors.

$$T(\gamma_\nu)\cdot\gamma^\mu = \left(\left(\frac{\partial}{\partial\phi_\mu}\right)\mathcal{L}\right)(\gamma_\nu\cdot(\gamma^\sigma\partial_\sigma))\phi - \gamma_\nu\cdot\gamma^\mu\mathcal{L}. \tag{F.56}$$

Thus writing $T^\mu{}_\nu = T(\gamma_\nu)\cdot\gamma^\mu$ we have

$$T^\mu{}_\nu = \frac{\partial\mathcal{L}}{\partial\phi_\mu}\partial_\nu\phi - \delta_\nu{}^\mu\mathcal{L}. \tag{F.57}$$

F.3.4 *Multiple field variables.*

In order to deal with the Maxwell Lagrangian a generalization to multiple field variables is required. Suppose now that we have a Lagrangian density $\mathcal{L} = \mathcal{L}(\phi^\alpha, \partial_\beta\phi^\alpha)$. Proceeding with the chain rule application again we have after some latex search and replace adding in indices in all the right places (proof by regular expressions)

$$\begin{aligned}
\delta\mathcal{L} &= (a\cdot\nabla)\mathcal{L} \\
&= a^\mu\partial_\mu\mathcal{L} \\
&= a^\mu\left(\frac{\partial\phi^\alpha}{\partial x^\mu}\frac{\partial\mathcal{L}}{\partial\phi^\alpha} + \frac{\partial\partial_\sigma\phi^\alpha}{\partial x^\mu}\frac{\partial\mathcal{L}}{\partial\partial_\sigma\phi^\alpha}\right) \\
&= \frac{\partial\mathcal{L}}{\partial\phi^\alpha}(a\cdot\nabla)\phi^\alpha + \frac{\partial\mathcal{L}}{\partial\partial_\sigma\phi^\alpha}(a\cdot\nabla)\partial_\sigma\phi^\alpha \\
&= \left(\partial_\sigma\frac{\partial\mathcal{L}}{\partial\partial_\sigma\phi^\alpha}\right)(a\cdot\nabla)\phi^\alpha + \frac{\partial\mathcal{L}}{\partial\partial_\sigma\phi^\alpha}(a\cdot\nabla)\partial_\sigma\phi^\alpha \\
&= \left(\partial_\sigma\frac{\partial\mathcal{L}}{\partial\partial_\sigma\phi^\alpha}\right)(a\cdot\nabla)\phi^\alpha + \frac{\partial\mathcal{L}}{\partial\partial_\sigma\phi^\alpha}\partial_\sigma((a\cdot\nabla)\phi^\alpha) \\
&= \partial_\sigma\left(\frac{\partial\mathcal{L}}{\partial\partial_\sigma\phi^\alpha}(a\cdot\nabla)\phi^\alpha\right).
\end{aligned} \tag{F.58}$$

In the above manipulations (and those below), any repeated index, regardless of whether upper and lower indices are matched implies summation.

Using this we have a multiple field generalization of eq. (F.51). The Noether current and its conservation law in coordinate form is

$$\begin{aligned}
T^\mu(a) &= \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^\alpha}(a\cdot\nabla)\phi^\alpha - a^\mu\mathcal{L} \\
\partial_\mu T^\mu(a) &= 0.
\end{aligned} \tag{F.59}$$

Or in vector form, corresponding to eq. (F.53)

$$\begin{aligned} T(a) &= \left(\left(\gamma_\mu \frac{\partial}{\partial \partial_\mu \phi^\alpha}\right)\mathcal{L}\right)(a\cdot\nabla)\phi^\alpha - a\mathcal{L} \\ \nabla\cdot T(a) &= 0. \end{aligned} \tag{F.60}$$

And finally in tensor form, as in eq. (F.57)

$$\begin{aligned} T^\mu{}_\nu &= \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^\alpha}\partial_\nu \phi^\alpha - \delta_\nu{}^\mu \mathcal{L} \\ \partial_\mu T^\mu{}_\nu &= 0. \end{aligned} \tag{F.61}$$

F.3.5 *Spatial Noether current.*

The conservation arguments above have been expressed with the assumption that the Lagrangian density is a function of both spatial and time coordinates, and this was made explicit with the use of the Dirac basis to express the Noether current.

It should be pointed out that for a purely spatial Lagrangian density, such as that of electrostatics

$$\mathcal{L} = -\frac{\epsilon_0}{2}(\boldsymbol{\nabla}\phi)^2 + \rho\phi. \tag{F.62}$$

the same results apply. In this case it would be reasonable to summarize the conservation under translation using the Pauli basis and write

$$\begin{aligned} T(\mathbf{a}) &= \sigma_k \frac{\partial \mathcal{L}}{\partial \partial_k \phi}\mathbf{a}\cdot\boldsymbol{\nabla}\phi - \mathbf{a}\mathcal{L} \\ \boldsymbol{\nabla}\cdot T(\mathbf{a}) &= 0. \end{aligned} \tag{F.63}$$

Without the time translation, calling the vector Noether current the energy momentum tensor is not likely appropriate. Perhaps just the canonical energy momentum tensor? Working with such a spatial Lagrangian density later should help clarify how to label things.

F.4 FIELD HAMILTONIAN.

A special case of eq. (F.57) is for time translation of the Lagrangian.

For that, our Noether current, writing $\mathcal{H}^\mu = T^\mu{}_0$ is

$$\begin{aligned} \mathcal{H}^0 &= \frac{\partial \mathcal{L}}{\partial \dot{\phi}}\dot{\phi} - \mathcal{L} \\ \mathcal{H}^k &= \frac{\partial \mathcal{L}}{\partial \phi_k}\dot{\phi}. \end{aligned} \tag{F.64}$$

These are expected to have a role associated with field energy and momentum respectively.

For the Maxwell Lagrangian we will need the multiple field current

$$\begin{aligned}\mathcal{H}^0 &= \frac{\partial \mathcal{L}}{\partial \partial_0 \phi^\alpha}\partial_0\phi^\alpha - \mathcal{L}\\ \mathcal{H}^k &= \frac{\partial \mathcal{L}}{\partial \partial_k \phi^\alpha}\partial_0\phi^\alpha.\end{aligned} \tag{F.65}$$

F.5 WAVE EQUATION.

The energy momentum tensor has been computed for some general field Lagrangians. Now let's consider some specific concrete examples. The Lagrangian for the relativistic wave equation is an obvious first candidate due to simplicity.

F.5.1 *Tensor components and energy term.*

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\partial_\mu\phi\partial^\mu\phi\\ &= \frac{1}{2}\phi_\mu\phi^\mu\\ &= \frac{1}{2}\left(\nabla\phi\right)^2\\ &= \frac{1}{2}\left(\dot{\phi}^2 - (\boldsymbol{\nabla}\phi)^2\right).\end{aligned} \tag{F.66}$$

In the explicit spacetime split above we have a split into terms that appear to correspond to kinetic and potential terms

$$\mathcal{L} = K - V. \tag{F.67}$$

To compute the tensor, we first need $\partial\mathcal{L}/\partial\phi_\mu = \phi^\mu$, which gives us

$$T^\mu{}_\nu = \phi^\mu\phi_\nu - \delta_\nu{}^\mu\mathcal{L}. \tag{F.68}$$

Writing this out in matrix form (with rows μ, and columns ν), we have

$$\begin{bmatrix} \frac{1}{2}(\dot{\phi}^2+\phi_x^2+\phi_y^2+\phi_z^2) & \dot{\phi}\phi_x & \dot{\phi}\phi_y & \dot{\phi}\phi_z \\ -\phi_x\dot{\phi} & \frac{1}{2}(-\dot{\phi}^2-\phi_x^2+\phi_y^2+\phi_z^2) & -\phi_x\phi_y & -\phi_x\phi_z \\ -\phi_y\dot{\phi} & -\phi_y\phi_x & \frac{1}{2}(-\dot{\phi}^2+\phi_x^2-\phi_y^2+\phi_z^2) & -\phi_y\phi_z \\ -\phi_z\dot{\phi} & -\phi_z\phi_x & -\phi_z\phi_y & \frac{1}{2}(-\dot{\phi}^2+\phi_x^2+\phi_y^2-\phi_z^2) \end{bmatrix}. \tag{F.69}$$

As mentioned by Jackson, the canonical energy momentum tensor is not necessarily symmetric, and we see that here. We have what is expected for the wave energy in the $0, 0$ element

$$\begin{aligned} T^0{}_0 &= K + V \\ &= \frac{1}{2}(\dot{\phi}^2 + (\boldsymbol{\nabla}\phi)^2). \end{aligned} \tag{F.70}$$

F.5.2 *Conservation equations.*

How about the conservation equations when written in full. The first is

$$\begin{aligned} 0 &= \partial_\mu T^\mu{}_0 \\ &= \frac{1}{2}\partial_t\left(\dot{\phi}^2 + \phi_x^2 + \phi_y^2 + \phi_z^2\right) - \partial_x(\phi_x\dot{\phi}) - \partial_y(\phi_y\dot{\phi}) - \partial_z(\phi_z\dot{\phi}) \\ &= \dot{\phi}\ddot{\phi} + \phi_x\phi_{xt} + \phi_y\phi_{yt} + \phi_z\phi_{zt} - \phi_{xx}\dot{\phi} - \phi_{yy}\dot{\phi} - \phi_{zz}\dot{\phi} - \phi_x\phi_{tx} - \phi_y\phi_{ty} - \phi_z\phi_{tz} \\ &= \dot{\phi}(\ddot{\phi} - \phi_{xx} - \phi_{yy} - \phi_{zz}). \end{aligned} \tag{F.71}$$

So our first conservation equation is

$$0 = \dot{\phi}(\nabla^2\phi). \tag{F.72}$$

But $\nabla^2\phi = 0$ is just our wave equation, the result of the variation of the Lagrangian itself. So curiously the divergence of energy-momentum four vector $T^\mu{}_0$ ends up as another method of supplying the wave equation!

How about one of the other conservation equations? The pattern will all be the same, so calculating one is sufficient.

$$\begin{aligned} 0 &= \partial_\mu T^\mu{}_1 \\ &= \partial_t(\dot{\phi}\phi_x) + \frac{1}{2}\partial_x(-\dot{\phi}^2 - \phi_x^2 + \phi_y^2 + \phi_z^2) - \partial_y(\phi_y\phi_x) - \partial_z(\phi_z\phi_x) \\ &= \ddot{\phi}\phi_x + \dot{\phi}\phi_{xt} - \dot{\phi}\phi_{tx} - \phi_x\phi_{xx} + \phi_y\phi_{yx} + \phi_z\phi_{zx} - \phi_{yy}\phi_x - \phi_y\phi_{xy} - \phi_{zz}\phi_x - \phi_z\phi_{xz} \\ &= \phi_x(\ddot{\phi} - \phi_{xx} - \phi_{yy} - \phi_{zz}). \end{aligned} \tag{F.73}$$

It should probably not be surprising that we have such a symmetric relation between space and time for the wave equations and we can summarize the spacetime translation conservation equations by

$$\begin{aligned} 0 &= \partial_\mu T^\mu{}_\nu \\ &= \phi_\nu(\nabla^2\phi). \end{aligned} \tag{F.74}$$

F.5.3 *Invariant length.*

It has been assumed that $T(\gamma_\mu)$ are four vectors. If that is the cast we ought to have an invariant length.

Let us calculate the vector square of $T(\gamma_0)$. Picking off first column of our tensor in eq. (F.69), we have

$$\begin{aligned}
(T(\gamma_0))^2 &= (\gamma_\mu T^\mu{}_0)\cdot(\gamma_\nu T^\nu{}_0) \\
&= (T^0{}_0)^2 - (T^1{}_0)^2 - (T^2{}_0)^2 - (T^3{}_0)^2 \\
&= \frac{1}{4}\left(\dot\phi^2 + \phi_x^2 + \phi_y^2 + \phi_z^2\right)^2 - \phi_x^2\dot\phi^2 - \phi_y^2\dot\phi^2 - \phi_z^2\dot\phi^2 \\
&= \frac{1}{4}\left(\dot\phi^4 + \phi_x^4 + \phi_y^4 + \phi_z^4\right) - \frac{1}{2}\left(\dot\phi^2\phi_x^2 + \dot\phi^2\phi_y^2 + \dot\phi^2\phi_z^2\right) \\
&\quad + \frac{1}{2}\left(+\phi_x^2\phi_y^2 + \phi_y^2\phi_z^2 + \phi_z^2\phi_x^2\right) \\
&= \frac{1}{4}\left(\dot\phi^2 - \phi_x^2 - \phi_y^2 - \phi_z^2\right)^2 .
\end{aligned} \tag{F.75}$$

But this is just our (squared) Lagrangian density, and we therefore have

$$(T(\gamma_0))^2 = \mathcal{L}^2. \tag{F.76}$$

Doing the same calculation for the second column, which is representative of the other two by symmetry, we have

$$(T(\gamma_k))^2 = -\mathcal{L}^2. \tag{F.77}$$

Summarizing all four squares we have

$$(T(\gamma_\mu))^2 = (\gamma_\mu)^2\mathcal{L}^2. \tag{F.78}$$

All of these conservation current four vectors have the same length up to a sign, where $T(\gamma_0)$ is timelike (positive square), whereas $T(\gamma_k)$ is spacelike (negative square).

Now, is $\mathcal{L}^2$ a Lorentz invariant? If so we can justify calling $T(\gamma_\mu)$ four vectors. Reflection shows that this is in fact the case, since $\mathcal{L}$ is a Lorentz

invariant. The transformation properties of $\mathcal{L}$ go with the gradient. Writing $\nabla' = R\nabla\tilde{R}$, we have

$$
\begin{aligned}
\mathcal{L}' &= \frac{1}{2}\nabla'\phi\cdot\nabla'\phi \\
&= \frac{1}{2}\left\langle R\nabla\tilde{R}\phi R\nabla\tilde{R}\phi\right\rangle \\
&= \frac{1}{2}\left\langle R\nabla\phi\nabla\tilde{R}\phi\right\rangle \\
&= \frac{1}{2}\left\langle \underbrace{\boxed{\tilde{R}R}}_{=1}\nabla\phi\nabla\phi\right\rangle \\
&= \frac{1}{2}\nabla\phi\cdot\nabla\phi \\
&= \mathcal{L}.
\end{aligned}
\tag{F.79}
$$

F.5.4 *Diagonal terms of the tensor.*

There is a conjugate structure evident in the diagonal terms of the matrix for the tensor. In particular, the $T^0{}_0$ can be expressed using the Hermitian conjugate from QM. For a multivector F, this was defined as

$$
F^\dagger = \gamma_0\tilde{F}\gamma_0. \tag{F.80}
$$

We have for $T^0{}_0$

$$
\begin{aligned}
T^0{}_0 &= \frac{1}{2}(\nabla\phi)^\dagger\cdot(\nabla\phi) \\
&= \frac{1}{2}\langle\gamma_0\nabla\gamma_0\phi\nabla\phi\rangle \\
&= \frac{1}{2}\left\langle(\gamma_0\nabla\phi)^2\right\rangle \\
&= \frac{1}{2}\left\langle(\gamma_0(\gamma^0\partial_0+\gamma^k\partial_k)\phi)^2\right\rangle \\
&= \frac{1}{2}\left\langle((\partial_0-\gamma^k\gamma_0\partial_k)\phi)^2\right\rangle \\
&= \frac{1}{2}\left\langle((\partial_0+\boldsymbol{\nabla})\phi)^2\right\rangle \\
&= \frac{1}{2}\left(\dot{\phi}^2+(\boldsymbol{\nabla}\phi)^2\right).
\end{aligned}
\tag{F.81}
$$

Now conjugation with respect to the time basis vector should not be special in any way, and should be equally justified defining a conjugation operation along any of the spatial directions too. Is there a symbol for this? Let us write for now

$$F^{\dagger_\mu} \equiv \gamma_\mu \tilde{F} \gamma^\mu. \tag{F.82}$$

There is a possibility that the sign picked here is not appropriate for all purposes. It is hard to tell for now since we have a vector F that equals its reverse, and in fact after a computation with both μ indices down I have raised an index altering an initial choice of $F^{\dagger_\mu} = \gamma_\mu \tilde{F} \gamma_\mu$.

Applying this, for $\mu \neq 0$ we have

$$\begin{aligned}
(\nabla\phi)^{\dagger_\mu} \cdot (\nabla\phi) &= -\left\langle \gamma_\mu \nabla \gamma_\mu \phi \nabla \phi \right\rangle \\
&= -\left\langle ((\partial_\mu + \gamma_\mu \sum_{\nu\neq\mu} \gamma^\nu \partial_\nu)\phi)^2 \right\rangle \\
&= -((\partial_\mu\phi)^2 + \sum_{\nu\neq\mu} (\gamma_\mu\gamma^\nu)^2 (\partial_\nu\phi)^2) \\
&= -((\partial_\mu\phi)^2 - \sum_{\nu\neq\mu} (\gamma_\mu)^2 (\gamma^\nu)^2 (\partial_\nu\phi)^2) \\
&= -((\partial_\mu\phi)^2 + \sum_{\nu\neq\mu} (\gamma^\nu)^2 (\partial_\nu\phi)^2) \\
&= -(\partial_\mu\phi)^2 - (\partial_0\phi)^2 + \sum_{k\neq\mu, k\neq 0} (\partial_k\phi)^2.
\end{aligned} \tag{F.83}$$

This recovers the diagonal terms, and allows us to write (no sum)

$$T^\mu{}_\mu = \frac{1}{2} (\nabla\phi)^{\dagger_\mu} \cdot (\nabla\phi). \tag{F.84}$$

F.5.4.1 *As a projection?*

As a vector (a projection of $T(\gamma_\mu)$ onto the γ_μ direction) this is (again no sum)

$$\begin{aligned}
\gamma_\mu T^\mu{}_\mu &= \frac{1}{2} \gamma_\mu (\nabla\phi)^{\dagger_\mu} \cdot (\nabla\phi) \\
&= \frac{1}{2} \gamma_\mu \left\langle \gamma^\mu \nabla\phi \gamma_\mu \nabla\phi \right\rangle \\
&= \frac{1}{4} \gamma_\mu (\gamma^\mu \nabla\phi \gamma_\mu \nabla\phi + \nabla\phi \gamma_\mu \nabla\phi \gamma^\mu) \\
&= \frac{1}{4} ((\nabla\phi \gamma_\mu \nabla\phi) + \gamma_\mu (\nabla\phi \gamma_\mu \nabla\phi) \gamma^\mu).
\end{aligned} \tag{F.85}$$

Intuition says this may have a use when assembling a complete vector representation of $T(\gamma_\mu)$ in terms of the gradient, but what that is now is not clear.

F.5.5 *Momentum.*

Now, let us look at the four vector $T(\gamma_0) = \gamma_\mu T^\mu{}_0$ more carefully. We have seen the energy term of this, but have not looked at the spatial part (momentum).

We can calculate the spatial component by wedging with the observer unit velocity γ_0, and get

$$\begin{aligned} T(\gamma_0) \wedge \gamma_0 &= \gamma_k \gamma_0 T^k{}_0 \\ &= -\sigma_k \dot{\phi} \phi_k \\ &= -\dot{\phi} \boldsymbol{\nabla} \phi. \end{aligned} \tag{F.86}$$

Right away we have something interesting! The wave momentum is related to the gradient operator, exactly as we have in quantum physics, despite the fact that we are only looking at the classical wave equation (for light or some other massless field effect).

F.6 WAVE EQUATION. GA FORM FOR THE ENERGY MOMENTUM TENSOR.

Some of the playing around above was attempting to find more structure for the terms of the energy momentum tensor. For the diagonal terms this was done successfully. However, doing so for the remainder is harder when working backwards from the tensor in coordinate form.

F.6.1 *Calculate GA form.*

Let us step back to the defining relation eq. (F.60), from which we see that we wish to calculate

$$\begin{aligned} \gamma_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^\alpha} &= \gamma_\mu \partial^\mu \phi \\ &= \gamma^\mu \partial_\mu \phi \\ &= \nabla \phi. \end{aligned} \tag{F.87}$$

This completely removes the indices from the tensor, leaving us with

$$\begin{aligned} T(a) &= (\nabla\phi)a\cdot\nabla\phi - \frac{a}{2}(\nabla\phi)^2 \\ &= (\nabla\phi)\left(\frac{1}{2}(a\nabla\phi + \nabla\phi a) - \nabla\phi\frac{a}{2}\right). \end{aligned} \tag{F.88}$$

Thus we have

$$T(a) = \frac{1}{2}(\nabla\phi)a(\nabla\phi). \tag{F.89}$$

This meets the intuitive expectation that the energy momentum tensor for the wave equation could be expressed completely in terms of the gradient.

F.6.2 *Verify against tensor expression.*

There is in fact a surprising simplicity to the result of eq. (F.89). It is somewhat hard to believe that it summarizes the messy matrix we have calculated above. To verify this let us derive the tensor relation of eq. (F.68).

$$\begin{aligned} T^\mu{}_\nu &= T(\gamma_\nu)\cdot\gamma^\mu \\ &= \frac{1}{2}\langle(\nabla\phi)\gamma_\nu(\nabla\phi)\gamma^\mu\rangle \\ &= \frac{1}{2}\left\langle\gamma^\alpha\partial_\alpha\phi\gamma_\nu\gamma_\beta\partial^\beta\phi\gamma^\mu\right\rangle \\ &= \frac{1}{2}\partial_\alpha\phi\partial^\beta\phi\left\langle\gamma^\alpha\gamma_\nu\gamma_\beta\gamma^\mu\right\rangle \\ &= \frac{1}{2}\partial_\alpha\phi\partial^\beta\phi\left(\delta^\alpha{}_\nu\delta_\beta{}^\mu + (\gamma^\alpha\wedge\gamma_\nu)\cdot(\gamma_\beta\wedge\gamma^\mu)\right) \\ &= \frac{1}{2}\left(\partial_\nu\phi\partial^\mu\phi + \partial^\alpha\phi\partial_\beta\phi(\gamma_\alpha\wedge\gamma_\nu)\cdot(\gamma^\beta\wedge\gamma^\mu)\right) \\ &= \frac{1}{2}\left(\partial_\nu\phi\partial^\mu\phi + (\partial^\alpha\phi\partial_\beta\phi)\gamma_\alpha\cdot(\underbrace{\boxed{\gamma_\nu\cdot(\gamma^\beta\wedge\gamma^\mu)}}_{=\delta_\nu{}^\beta\gamma^\mu - \delta_\nu{}^\mu\gamma^\beta})\right) \\ &= \frac{1}{2}\left(\partial_\nu\phi\partial^\mu\phi + (\partial^\alpha\phi\partial_\beta\phi)(\delta_\nu{}^\beta\delta_\alpha{}^\mu - \delta_\nu{}^\mu\delta_\alpha{}^\beta)\right) \\ &= \frac{1}{2}\left(\partial_\nu\phi\partial^\mu\phi + \partial^\mu\phi\partial_\nu\phi - \delta_\nu{}^\mu(\partial^\alpha\phi\partial_\alpha\phi)\right) \\ &= \partial_\nu\phi\partial^\mu\phi - \delta_\nu{}^\mu\mathcal{L}. \qquad \square \end{aligned} \tag{F.90}$$

F.6.3 *Invariant length.*

Putting the energy momentum tensor in GA form makes the demonstration of the invariant length almost trivial. We have for any a

$$\begin{aligned}
(T(a))^2 &= \frac{1}{4}\nabla\phi a\nabla\phi\nabla\phi a\nabla\phi \\
&= \frac{1}{4}(\nabla\phi)^2\nabla\phi a^2\nabla\phi \\
&= \frac{1}{4}(\nabla\phi)^4 a^2 \\
&= \mathcal{L}^2 a^2.
\end{aligned} \tag{F.91}$$

This recovers eq. (F.78), which came at considerably higher cost in terms of guesswork.

F.6.4 *Energy and Momentum split (again).*

By wedging with γ_0 we can extract the momentum terms of $T(\gamma_0)$. That is

$$\begin{aligned}
T(\gamma_0)\wedge\gamma_0 &= \left((\gamma_0\cdot\nabla\phi)\nabla\phi - \frac{1}{2}(\nabla\phi)^2\gamma_0\right)\wedge\gamma_0 \\
&= (\gamma_0\cdot\nabla\phi)(\nabla\phi\wedge\gamma_0) - \frac{1}{2}(\nabla\phi)^2\overbrace{\boxed{(\gamma_0\wedge\gamma_0)}}^{=0} \\
&= \dot{\phi}(\gamma^k\gamma_0\partial_k\phi) \\
&= -\dot{\phi}\boldsymbol{\nabla}\phi.
\end{aligned} \tag{F.92}$$

For the energy term, dotting with γ_0 we have

$$\begin{aligned}
T(\gamma_0)\cdot\gamma_0 &= \left((\gamma_0\cdot\nabla\phi)\nabla\phi - \frac{1}{2}(\nabla\phi)^2\gamma_0\right)\cdot\gamma_0 \\
&= (\gamma_0\cdot\nabla\phi)^2 - \frac{1}{2}(\nabla\phi)^2 \\
&= \dot{\phi}^2 - \frac{1}{2}(\dot{\phi}^2 - (\boldsymbol{\nabla}\phi)^2) \\
&= \frac{1}{2}\left(\dot{\phi}^2 + (\boldsymbol{\nabla}\phi)^2\right).
\end{aligned} \tag{F.93}$$

Wedging with γ_0 itself does not provide us with a relative spatial vector. For example, consider the proper time velocity four vector (still working with $c = 1$)

$$\begin{aligned} v &= \frac{dt}{d\tau}\frac{d}{dt}\left(t\gamma_0 + \gamma_k x^k\right) \\ &= \frac{dt}{d\tau}\left(\gamma_0 + \gamma_k \frac{dx^k}{dt}\right). \end{aligned} \tag{F.94}$$

We have

$$v \cdot \gamma_0 = \frac{dt}{d\tau} = \gamma, \tag{F.95}$$

and

$$v \wedge \gamma_0 = \frac{dt}{d\tau}\sigma_k \frac{dx^k}{dt}. \tag{F.96}$$

Or

$$\begin{aligned} \mathbf{v} &\equiv \sigma_k \frac{dx^k}{dt} \\ &= \frac{v \wedge \gamma_0}{v \cdot \gamma_0}. \end{aligned} \tag{F.97}$$

This suggests that the form for the relative momentum (spatial) vector for the field should therefore be

$$\begin{aligned} \mathbf{p} &\equiv \frac{T(\gamma_0) \wedge \gamma_0}{T(\gamma_0) \cdot \gamma_0} \\ &= -\frac{\dot{\phi}}{\frac{1}{2}(\dot{\phi}^2 + (\boldsymbol{\nabla}\phi)^2)}\boldsymbol{\nabla}\phi \\ &= -\frac{2}{1 + \left(\frac{\boldsymbol{\nabla}\phi}{\dot{\phi}}\right)^2}\frac{\boldsymbol{\nabla}\phi}{\dot{\phi}} \\ &= -\frac{2}{\frac{\dot{\phi}}{\boldsymbol{\nabla}\phi} + \frac{\boldsymbol{\nabla}\phi}{\dot{\phi}}}. \end{aligned} \tag{F.98}$$

This has been written in a few different ways, looking for something familiar, and not really finding it. It would be useful to revisit this after considering in detail wave momentum in a mechanical sense, perhaps with a limiting argument as given in [4] (ie: one dimensional Lagrangian density considering infinite sequence of springs in a line).

F.7 SCALAR KLEIN GORDON.

A number of details have been extracted considering the scalar wave equation. Now lets move to a two field variable Lagrangian.

$$\mathcal{L} = \frac{1}{2}\partial_\mu\psi\partial^\mu\psi - \frac{m^2c^2}{2\hbar^2}\psi^2. \tag{F.99}$$

This forced wave equation will have almost the same energy momentum tensor. The exception will be the diagonal terms for which we have an additional factor of $m^2c^2\psi^2/2\hbar^2$.

This also means that the conservation equations will be altered slightly

$$\begin{aligned} 0 &= \partial_\mu T^\mu{}_\nu \\ &= \phi_\nu\left(\nabla^2\phi + \frac{m^2c^2}{\hbar^2}\phi\right). \end{aligned} \tag{F.100}$$

Again the divergence of the individual canonical energy momentum tensor four vectors reproduces the field equations that we also obtain from the variation.

F.8 COMPLEX KLEIN GORDON.

F.8.1 *Tensor in GA form.*

$$\mathcal{L} = \partial_\mu\psi\partial^\mu\psi^* - \frac{m^2c^2}{\hbar^2}\psi\psi^*. \tag{F.101}$$

We first want to calculate what perhaps could be called the field velocity gradient

$$\begin{aligned} \gamma_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)} &= \gamma_\mu\partial^\mu\psi \\ &= \nabla\psi. \end{aligned} \tag{F.102}$$

Similarly

$$\begin{aligned} \gamma_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi^*)} &= \gamma_\mu\partial^\mu\psi^* \\ &= \nabla\psi^*. \end{aligned} \tag{F.103}$$

Assembling results into an application of eq. (F.60), we have

$$\begin{aligned}
T(a) &= \nabla\psi(a\cdot\nabla)\psi^* + \nabla\psi^*(a\cdot\nabla)\psi - a\mathcal{L} \\
&= \nabla\psi(a\cdot\nabla)\psi^* + \nabla\psi^*(a\cdot\nabla)\psi - a\frac{1}{2}(\nabla\psi\nabla\psi^* + \nabla\psi^*\nabla\psi) + a\frac{m^2c^2}{\hbar^2}\psi\psi^* \\
&= \nabla\psi(a\cdot\nabla\psi^* - \frac{1}{2}\nabla\psi^* a) + \nabla\psi^*(a\cdot\nabla\psi - \frac{1}{2}\nabla\psi a) + a\frac{m^2c^2}{\hbar^2}\psi\psi^* \\
&= \frac{1}{2}\left((\nabla\psi)a(\nabla\psi^*) + (\nabla\psi^*)a(\nabla\psi)\right) + a\frac{m^2c^2}{\hbar^2}\psi\psi^*.
\end{aligned} \tag{F.104}$$

Since vectors equal their own reverse this is just

$$T(a) = (\nabla\psi)a(\nabla\psi^*) + a\frac{m^2c^2}{\hbar^2}\psi\psi^*. \tag{F.105}$$

F.8.2 *Tensor in index form.*

Expanding the energy momentum tensor in index notation we have

$$\begin{aligned}
T^\mu{}_\nu &= T(\gamma_\nu)\cdot\gamma^\mu \\
&= \partial_\alpha\psi\partial_\beta\psi^*\left\langle\gamma^\alpha\gamma_\nu\gamma^\beta\gamma^\mu\right\rangle + \delta_\nu{}^\mu\frac{m^2c^2}{\hbar^2}\psi\psi^* \\
&= \partial_\nu\psi\partial^\mu\psi^* + \partial^\mu\psi\partial_\nu\psi^* - \partial^\alpha\psi\partial_\alpha\psi^*\delta_\nu{}^\mu + \delta_\nu{}^\mu\frac{m^2c^2}{\hbar^2}\psi\psi^*.
\end{aligned} \tag{F.106}$$

So we have

$$T^\mu{}_\nu = \partial^\mu\psi\partial_\nu\psi^* + \partial^\mu\psi^*\partial_\nu\psi - \delta_\nu{}^\mu\mathcal{L}. \tag{F.107}$$

This index representation also has a nice compact elegance.

F.8.3 *Invariant Length?*

Writing for short $b = \nabla\psi$, and working in natural units $m^2c^2 = \hbar^2$, we have

$$\begin{aligned}
(T(a))^2 &= (bab^* + a\psi\psi^*)^2 \\
&= \langle ab^*bab^*b\rangle + a^2\psi^2(\psi^*)^2 + 2a\cdot(bab^*).
\end{aligned} \tag{F.108}$$

Unlike the light wave equation this does not (obviously) appear to have a natural split into something times a^2. Is there a way to do it?

F.8.4 *Divergence relation.*

Borrowing notation from above to calculate the divergence we want

$$\begin{aligned}\nabla\cdot(bab^*) &= \langle\nabla(bab^*)\rangle \\ &= \left\langle (b^* \overleftrightarrow{\nabla} b)a\right\rangle \\ &= a\cdot\left\langle b^* \overleftrightarrow{\nabla} b\right\rangle_1 .\end{aligned} \tag{F.109}$$

Here cyclic reordering of factors within the scalar product was used. In order for that to be a meaningful operation the gradient must be allowed to operate bidirectionally, so this is really just shorthand for

$$b^* \overleftrightarrow{\nabla} b \equiv \dot{b}^*\dot{\nabla}b + b^*\dot{\nabla}\dot{b}, \tag{F.110}$$

where the more conventional overdot notation is used to indicate the scope of the operation. In particular, for $b = \nabla\psi$, we have

$$\left\langle b^* \overleftrightarrow{\nabla} b\right\rangle_1 = (\nabla^2\psi^*)(\nabla\psi) + (\nabla\psi^*)(\nabla^2\psi). \tag{F.111}$$

Our tensor also has a vector scalar product that we need the divergence of. That is

$$\begin{aligned}\nabla\cdot(a\psi\psi^*) &= \langle\nabla(a\psi\psi^*)\rangle \\ &= a\cdot\nabla(\psi\psi^*).\end{aligned} \tag{F.112}$$

Putting things back together we have

$$\nabla\cdot T(a) = a\cdot\left(\left\langle(\nabla\psi^*)\overleftrightarrow{\nabla}(\nabla\psi)\right\rangle_1 + \frac{m^2c^2}{\hbar^2}\nabla(\psi\psi^*)\right). \tag{F.113}$$

This is

$$0 = \nabla\cdot T(a) = a\cdot\left((\nabla^2\psi^*)(\nabla\psi) + (\nabla\psi^*)(\nabla^2\psi) + \frac{m^2c^2}{\hbar^2}\nabla(\psi\psi^*)\right). \tag{F.114}$$

Again, we see that the divergence of the canonical energy momentum tensor produces the field equations that we get by direct variation! Put explicitly we have zero for all displacements a, so must also have

$$0 = (\nabla\psi)\left(\nabla^2\psi^* + \frac{m^2c^2}{\hbar^2}\psi^*\right) + (\nabla\psi^*)\left(\nabla^2\psi + \frac{m^2c^2}{\hbar^2}\psi\right). \tag{F.115}$$

Also noteworthy above is the adjoint relationship. The adjoint $\overline{F}$ of a an operator F was defined via the dot product

$$a \cdot F(b) \equiv b \cdot \overline{F}(a). \tag{F.116}$$

So we have a concrete example of the adjoint applied to the gradient, and for this energy momentum tensor we have

$$\overline{T}(\nabla) = \left\langle (\nabla\psi^*)\nabla(\nabla\psi)\right\rangle_1 + \frac{m^2c^2}{\hbar^2}\nabla(\psi\psi^*). \tag{F.117}$$

Here the arrows notation has been dropped, where it is implied that this derivative acts on all neighboring vectors either unidirectionally or bidirectionally as appropriate.

Now, this adjoint tensor is a curious beastie. Intuition says this one will have a Lorentz invariant length. A moment of reflection shows that this is in fact the case since the adjoint was fully expanded in eq. (F.115). That vector is zero, and the length is therefore also necessarily invariant.

F.8.5 *TODO.*

How about the energy and momentum split in this adjoint form? Could also write out adjoint in index notation for comparison to non-adjoint tensor in index form.

F.9 ELECTROSTATICS POISSON EQUATION.

F.9.1 *Lagrangian and spatial Noether current.*

$$\mathcal{L} = -\frac{\epsilon_0}{2}(\boldsymbol{\nabla}\phi)^2 + \rho\phi. \tag{F.118}$$

Evaluating this yields the desired $\boldsymbol{\nabla}^2\phi = -\rho/\epsilon_0$, or $\boldsymbol{\nabla}\cdot\mathbf{E} = \rho/\epsilon_0$.

F.9.2 *Energy momentum tensor.*

In this particular case we then have

$$\begin{aligned}
T(\mathbf{a}) &= \sigma_k(-\epsilon_0\partial_k\phi)\mathbf{a}\cdot\boldsymbol{\nabla}\phi - \mathbf{a}\mathcal{L} \\
&= -\epsilon_0(\boldsymbol{\nabla}\phi)\mathbf{a}\cdot\boldsymbol{\nabla}\phi - \mathbf{a}(-\frac{\epsilon_0}{2}(\boldsymbol{\nabla}\phi)^2 + \rho\phi) \\
&= -\epsilon_0(\boldsymbol{\nabla}\phi)\mathbf{a}\cdot\boldsymbol{\nabla}\phi + (\boldsymbol{\nabla}\phi)^2\mathbf{a}\frac{\epsilon_0}{2} - \mathbf{a}\rho\phi \\
&= \frac{\epsilon_0}{2}(\boldsymbol{\nabla}\phi)\left(-2\mathbf{a}\cdot\boldsymbol{\nabla}\phi + \boldsymbol{\nabla}\phi\mathbf{a}\right) - \mathbf{a}\rho\phi \\
&= \frac{\epsilon_0}{2}(\boldsymbol{\nabla}\phi)\left(-\mathbf{a}\boldsymbol{\nabla}\phi - \boldsymbol{\nabla}\phi\mathbf{a} + \boldsymbol{\nabla}\phi\mathbf{a}\right) - \mathbf{a}\rho\phi \\
&= -\frac{\epsilon_0}{2}(\boldsymbol{\nabla}\phi)\mathbf{a}\boldsymbol{\nabla}\phi - \mathbf{a}\rho\phi.
\end{aligned} \tag{F.119}$$

It in terms of $\mathbf{E} = -\boldsymbol{\nabla}\phi$ this is

$$T(\mathbf{a}) = -\frac{\epsilon_0}{2}\mathbf{E}\mathbf{a}\mathbf{E} - \mathbf{a}\rho\phi. \tag{F.120}$$

This is not immediately recognizable (at least to me), and also does not appear to be easily separable into something times $\mathbf{a}$.

F.9.3 *Divergence and adjoint tensor.*

What will we get with the divergence calculation?

$$\begin{aligned}
\boldsymbol{\nabla}\cdot(\mathbf{E}\mathbf{a}\mathbf{E}) &= \langle\boldsymbol{\nabla}(\mathbf{E}\mathbf{a}\mathbf{E})\rangle \\
&= \mathbf{a}\cdot\left\langle\mathbf{E}\overleftrightarrow{\boldsymbol{\nabla}}\mathbf{E}\right\rangle_1.
\end{aligned} \tag{F.121}$$

Also want

$$\begin{aligned}
\boldsymbol{\nabla}\cdot(\mathbf{a}\rho\phi) &= \langle\boldsymbol{\nabla}(\mathbf{a}\rho\phi)\rangle \\
&= \mathbf{a}\cdot\boldsymbol{\nabla}(\rho\phi).
\end{aligned} \tag{F.122}$$

Assembling these we have

$$\boldsymbol{\nabla}\cdot T(\mathbf{a}) = -\mathbf{a}\cdot\left(\left\langle\frac{\epsilon_0}{2}\mathbf{E}\overleftrightarrow{\boldsymbol{\nabla}}\mathbf{E}\right\rangle_1 + \boldsymbol{\nabla}(\rho\phi)\right). \tag{F.123}$$

From this we can pick off the adjoint

$$\begin{aligned}
\bar{T}(\boldsymbol{\nabla}) &= -\frac{\epsilon_0}{2}\left\langle \mathbf{E} \overleftrightarrow{\boldsymbol{\nabla}} \mathbf{E} \right\rangle_1 - \boldsymbol{\nabla}(\rho\phi) \\
&= -\frac{\epsilon_0}{2}\left((\dot{\mathbf{E}} \cdot \dot{\boldsymbol{\nabla}})\mathbf{E}\mathbf{E}(\boldsymbol{\nabla} \cdot \mathbf{E})\right) - \boldsymbol{\nabla}(\rho\phi) \\
&= -\epsilon_0(\boldsymbol{\nabla}^2\phi)\boldsymbol{\nabla}\phi - \boldsymbol{\nabla}(\rho\phi) \\
&= -\epsilon_0\boldsymbol{\nabla}(\boldsymbol{\nabla}\phi)^2 - \boldsymbol{\nabla}(\rho\phi) \\
&= \boldsymbol{\nabla}\left(-\epsilon_0(\boldsymbol{\nabla}\phi)^2 - \rho\phi\right).
\end{aligned} \tag{F.124}$$

If we write $\mathcal{L} = K - V$, then we have in this case

$$\bar{T}(\boldsymbol{\nabla}) = \boldsymbol{\nabla}(K + V) = 0. \tag{F.125}$$

Since the gradient of this quantity is zero everywhere it must be constant

$$K + V = \text{constant}. \tag{F.126}$$

We did not have any time dependence in the Lagrangian, and blindly following the math to calculate the associated symmetry with the field translation, we end up with a conservation statement that appears to be about energy.

TODO: am used to (as in [3]) seeing electrostatic energy written

$$U = \frac{1}{2}\epsilon_0 \int \mathbf{E}^2 dV = \frac{1}{2}\int \rho\phi dV. \tag{F.127}$$

Reconcile this with eq. (F.126).

F.10 SCHRÖDINGER EQUATION

While not a Lorentz invariant Lagrangian, we do not have a dependence on that, and can still calculate a Noether current on spatial translation.

$$\mathcal{L} = \frac{\hbar^2}{2m}(\boldsymbol{\nabla}\psi) \cdot (\boldsymbol{\nabla}\psi^*) + V\psi\psi^* + i\hbar\left(\psi\partial_t\psi^* - \psi^*\partial_t\psi\right). \tag{F.128}$$

For this Lagrangian density it is worth noting that the action is in fact

$$S = \int d^3x\mathcal{L}. \tag{F.129}$$

... ie: $\partial_t \psi$ is not a field variable in the variation (this is why there is no factor of $1/2$ in the probability current term). Calculating the Noether current for a vector translation $\mathbf{a}$ we have

$$T(\mathbf{a}) = \frac{\hbar^2}{2m}\nabla\psi\mathbf{a}\cdot\nabla\psi^* + \frac{\hbar^2}{2m}\nabla\psi^*\mathbf{a}\cdot\nabla\psi - \mathbf{a}\mathcal{L}. \tag{F.130}$$

Expanding the divergence is messy but straightforward

$$\begin{aligned}
&\nabla\cdot T(\mathbf{a}) \\
&= \frac{\hbar^2}{2m}\langle\nabla\left(\nabla\psi\mathbf{a}\cdot\nabla\psi^* + \nabla\psi^*\mathbf{a}\cdot\nabla\psi\right) - \nabla(\nabla\psi\cdot\nabla\psi^*)\mathbf{a}\rangle \\
&\quad - \mathbf{a}\cdot\nabla\left(V\psi\psi^* + i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi})\right) \\
&= \frac{\hbar^2}{4m}\langle\nabla\left(\nabla\psi(\mathbf{a}\nabla\psi^* + \nabla\psi^*\mathbf{a}) + \nabla\psi^*(\mathbf{a}\nabla\psi + \nabla\psi\mathbf{a})\right) - 2\nabla(\nabla\psi\cdot\nabla\psi^*)\mathbf{a}\rangle \\
&\quad - \mathbf{a}\cdot\nabla\left(V\psi\psi^* + i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi})\right) \\
&= \frac{\hbar^2}{4m}\mathbf{a}\cdot\left\langle\nabla\psi^*\overleftrightarrow{\nabla}\nabla\psi + \nabla\psi\overleftrightarrow{\nabla}\nabla\psi^*\right\rangle_1 \\
&\quad + \frac{\hbar^2}{4m}\mathbf{a}\cdot\langle\nabla(\nabla\psi\nabla\psi^*) + \nabla(\nabla\psi^*\nabla\psi) - 2\nabla(\nabla\psi\cdot\nabla\psi^*)\rangle_1 \\
&\quad - \mathbf{a}\cdot\nabla\left(V\psi\psi^* + i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi})\right) \\
&= \frac{\hbar^2}{4m}\mathbf{a}\cdot\left\langle\nabla\psi^*\overleftrightarrow{\nabla}\nabla\psi + \nabla\psi\overleftrightarrow{\nabla}\nabla\psi^*\right\rangle_1 \\
&\quad - \mathbf{a}\cdot\nabla\left(V\psi\psi^* + i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi})\right) \\
&= \frac{\hbar^2}{4m}\mathbf{a}\cdot 2\left(\nabla\psi^*\nabla^2\psi + \nabla\psi\nabla^2\psi^*\right) - \mathbf{a}\cdot\nabla\left(V\psi\psi^* + i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi})\right) \\
&= \frac{\hbar^2}{2m}\mathbf{a}\cdot\nabla\left(\nabla\psi^*\cdot\nabla\psi\right) - \mathbf{a}\cdot\nabla\left(V\psi\psi^* + i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi})\right).
\end{aligned} \tag{F.131}$$

Which is, finally,

$$\nabla\cdot T(\mathbf{a}) = \mathbf{a}\cdot\nabla\left(\frac{\hbar^2}{2m}\nabla\psi^*\cdot\nabla\psi - V\psi\psi^* - i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi})\right). \tag{F.132}$$

Picking off the adjoint we have

$$\overline{T}(\nabla) = \frac{\hbar^2}{2m}\nabla\psi^*\cdot\nabla\psi - V\psi\psi^* - i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi}). \tag{F.133}$$

Just like the electrostatics equation, it appears that we can make an association with Kinetic (K) and Potential (ϕ) energies with the adjoint stress tensor.

$$\begin{aligned}
K &= \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^* \cdot \boldsymbol{\nabla}\psi \\
\phi &= V\psi\psi^* + i\hbar(\psi\dot{\psi}^* - \psi^*\dot{\psi}) \\
\mathcal{L} &= K - \phi \\
\overline{T}(\boldsymbol{\nabla}) &= K + \phi.
\end{aligned} \tag{F.134}$$

FIXME: Unlike the electrostatics case however, there is no conserved scalar quantity that is obvious. The association in this case with energy is by analogy, not connected to anything reasonably physical seeming. How to connect this with actual physical concepts? Can this be written as the gradient of something? Because of the time derivatives perhaps the space time gradient would be required, however, because of the non-Lorentz invariant nature I had expect that terms may have to be added or subtracted to make that possible.

F.11 MAXWELL EQUATION.

Wanting to see some of the connections between the Maxwell equation and the Lorentz force was the original reason for examining this canonical energy momentum tensor concept in detail.

F.11.1 *Lagrangian.*

Recall that the Lagrangian for the vector grades of Maxwell's equation

$$\nabla F = J/\epsilon_0 c, \tag{F.135}$$

is of the form

$$\begin{aligned}
\mathcal{L} &= \kappa(\nabla \wedge A)\cdot(\nabla \wedge A) + J\cdot A \\
&= \kappa(\gamma^\mu \wedge \gamma^\nu)\cdot(\gamma_\alpha \wedge \gamma_\beta)\partial_\mu A_\nu \partial^\alpha A^\beta + J^\sigma A_\sigma.
\end{aligned} \tag{F.136}$$

We can fix the constant κ by taking variational derivatives and comparing with eq. (F.135)

$$\begin{aligned}
0 &= \frac{\partial \mathcal{L}}{\partial A_\sigma} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\sigma)} \\
&= J^\sigma - 2\kappa(\gamma^\mu \wedge \gamma^\sigma)\cdot(\gamma_\alpha \wedge \gamma_\beta)\partial_\mu \partial^\alpha A^\beta.
\end{aligned} \tag{F.137}$$

Taking γ^σ dot products with eq. (F.135) we have

$$\begin{aligned} 0 &= \gamma^\sigma \cdot (J - \epsilon_0 c \nabla \cdot F) \\ &= J^\sigma - \epsilon_0 c \gamma^\sigma \cdot (\gamma^\mu \cdot (\gamma_\alpha \wedge \gamma_\beta)) \partial_\mu \partial^\alpha A^\beta . \end{aligned} \tag{F.138}$$

So we have $2\kappa = -\epsilon_0 c$, and can write our Lagrangian density as

$$\begin{aligned} \mathcal{L} &= -\frac{\epsilon_0}{2} (\nabla \wedge A) \cdot (\nabla \wedge A) + \frac{J}{c} \cdot A \\ &= -\frac{\epsilon_0}{2} (\gamma^\mu \wedge \gamma^\nu) \cdot (\gamma_\alpha \wedge \gamma_\beta) \partial_\mu A_\nu \partial^\alpha A^\beta + \frac{J^\sigma}{c} A_\sigma . \end{aligned} \tag{F.139}$$

F.11.2 *Energy momentum tensor.*

For the Lagrangian density we have

$$\begin{aligned} \gamma_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -\epsilon_0 \gamma_\mu (\gamma^\mu \wedge \gamma^\nu) \cdot (\gamma_\alpha \wedge \gamma_\beta) \partial^\alpha A^\beta \\ &= -\epsilon_0 \gamma_\mu (\delta^\mu{}_\beta \delta^\nu{}_\alpha - \delta^\mu{}_\alpha \delta^\nu{}_\beta) \partial^\alpha A^\beta \\ &= -\epsilon_0 \gamma_\mu (\partial^\nu A^\mu - \partial^\mu A^\nu) \\ &= \epsilon_0 \gamma_\mu F^{\mu\nu} . \end{aligned} \tag{F.140}$$

One can guess that the vector contraction of $F^{\mu\nu}$ above is an expression of a dot product with our bivector field. This is in fact the case

$$\begin{aligned} F \cdot \gamma^\nu &= (\gamma_\alpha \wedge \gamma_\beta) \cdot \gamma^\nu \partial^\alpha A^\beta \\ &= (\gamma_\alpha \delta_\beta{}^\nu - \gamma_\beta \delta_\alpha{}^\nu) \partial^\alpha A^\beta \\ &= \gamma_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) \\ &= \gamma_\mu F^{\mu\nu} . \end{aligned} \tag{F.141}$$

We therefore have

$$\begin{aligned} T(a) &= \epsilon_0 (F \cdot \gamma^\nu) a \cdot \nabla A_\nu - a \mathcal{L} \\ &= \epsilon_0 \left((F \cdot \gamma^\nu) a \cdot \nabla A_\nu + \frac{a}{2} F \cdot F \right) - a \left(A \cdot J/c \right) . \end{aligned} \tag{F.142}$$

F.11.3 *Index form of tensor.*

Before trying to factor out a, let us expand the tensor in abstract index form. This is

$$\begin{aligned} T_\nu{}^\mu &= T(\gamma_\nu)\cdot\gamma^\mu \\ &= \epsilon_0\left(F^{\mu\beta}\partial_\nu A_\beta + \frac{\delta_\nu{}^\mu}{2}F\cdot F\right) - \delta_\nu{}^\mu A^\sigma J_\sigma/c \\ &= \epsilon_0\left(F^{\mu\beta}\partial_\nu A_\beta - \frac{\delta_\nu{}^\mu}{4}F^{\alpha\beta}F_{\alpha\beta}\right) - \delta_\nu{}^\mu A^\sigma J_\sigma/c. \end{aligned} \tag{F.143}$$

In particular, note that this is not the familiar symmetric tensor from the Poynting relations.

F.11.4 *Adjoint.*

Now, we want to move on to a computation of the adjoint so that a can essentially be factored out. Doing so is resisting initial attempts. As an aid, introduce a few vector valued helper variables

$$\begin{aligned} F^\mu &= F\cdot\gamma^\mu \\ G_\mu &= \nabla A_\nu. \end{aligned} \tag{F.144}$$

Then we have

$$\begin{aligned} \nabla\cdot T(a) &= \frac{\epsilon_0}{2}\left(\langle\nabla(F^\nu(aG_\nu + G_\nu a)\rangle + a\cdot\left\langle\nabla(F^2)\right\rangle_1\right) - a\cdot\nabla\left(A\cdot J/c\right) \\ &= \frac{\epsilon_0}{2}a\cdot\left\langle G_\nu\overleftrightarrow{\nabla}F^\nu + \nabla(F^\nu G_\nu) + \nabla(F^2)\right\rangle_1 - a\cdot\nabla\left(A\cdot J/c\right). \end{aligned} \tag{F.145}$$

This provides the adjoint energy momentum tensor, albeit in a form that looks like it can be reduced further

$$0 = \bar{T}(\nabla) = \frac{\epsilon_0}{2}\left\langle G_\nu\overleftrightarrow{\nabla}F^\nu + \nabla(F^\nu G_\nu) + \nabla(F^2)\right\rangle_1 - \nabla\left(A\cdot J/c\right). \tag{F.146}$$

We want to write this as a gradient of something, to determine the conserved quantity. Getting part way is not too hard.

$$\begin{aligned} \bar{T}(\nabla) = \frac{\epsilon_0}{2}&\overset{*}{\boxed{\left(\left\langle G_\nu\overleftrightarrow{\nabla}F^\nu\right\rangle_1 + \nabla\cdot(F^\nu\wedge G_\nu)\right)}} \\ &+ \nabla\left(\frac{\epsilon_0}{2}(F^\nu\cdot G_\nu + F\cdot F) - A\cdot J/c\right). \end{aligned} \tag{F.147}$$

It would be nice if these first two terms $*$ cancel. Can we be so lucky?

$$
\begin{aligned}
(*) &= \left\langle G_\nu \overset{\leftrightarrow}{\nabla} F^\nu \right\rangle_1 + \nabla \cdot (F^\nu \wedge G_\nu) \\
&= \left\langle (G_\nu \overset{\leftarrow}{\nabla}) F^\nu + G_\nu (\overset{\rightarrow}{\nabla} F^\nu) \right\rangle_1 + (\nabla \cdot F^\nu) G_\nu - F^\nu (\nabla \cdot G_\nu) \\
&= (\nabla \cdot G_\nu) F^\nu + F^\nu \cdot (\nabla \wedge G_\nu) + G_\nu (\nabla \cdot F^\nu) \\
&\quad + G_\nu \cdot (\nabla \wedge F^\nu) + (\nabla \cdot F^\nu) G_\nu - F^\nu (\nabla \cdot G_\nu) \\
&= 2 G_\nu (\nabla \cdot F^\nu) + G_\nu \cdot (\nabla \wedge F^\nu)
\end{aligned}
\tag{F.148}
$$

This is not obviously zero. How about $F^\nu \cdot G_\nu$?

$$
\begin{aligned}
F^\nu \cdot G_\nu &= \left\langle ((\gamma_\alpha \wedge \gamma_\beta) \cdot \gamma^\nu) \gamma^\sigma \right\rangle \partial^\alpha A^\beta \partial_\sigma A_\nu \\
&= (\delta_\alpha{}^\sigma \delta_\beta{}^\nu - \delta_\beta{}^\sigma \delta_\alpha{}^\nu) \partial^\alpha A^\beta \partial_\sigma A_\nu \\
&= \partial^\alpha A^\beta (\partial_\alpha A_\beta - \partial_\beta A_\alpha) \\
&= \partial^\alpha A^\beta F_{\alpha\beta} \\
&= \frac{1}{2} F_{\alpha\beta} F^{\alpha\beta}.
\end{aligned}
\tag{F.149}
$$

Ah. Up to a sign, this was $F \cdot F$. What is the sign?

$$
\begin{aligned}
F \cdot F &= (\gamma_\alpha \wedge \gamma_\beta) \cdot (\gamma^\mu \wedge \gamma^\nu) \partial^\alpha A^\beta \partial_\mu A_\nu \\
&= (\delta_\alpha{}^\nu \delta_\beta{}^\mu - \delta_\beta{}^\nu \delta_\alpha{}^\mu) \partial^\alpha A^\beta \partial_\mu A_\nu \\
&= \partial^\alpha A^\beta (\partial_\beta A_\alpha - \partial_\alpha A_\beta) \\
&= \partial^\alpha A^\beta F_{\beta\alpha} \\
&= \frac{1}{2} F_{\beta\alpha} (\partial^\alpha A^\beta - \partial^\beta A^\alpha) \\
&= \frac{1}{2} F_{\beta\alpha} F^{\alpha\beta} \\
&= -F^\nu \cdot G_\nu.
\end{aligned}
\tag{F.150}
$$

Bad first guess. It is the second two terms that cancel, not the first, leaving us with

$$
\overline{T}(\nabla) = \frac{\epsilon_0}{2} \left(\left\langle G_\nu \overset{\leftrightarrow}{\nabla} F^\nu \right\rangle_1 + \nabla \cdot (F^\nu \wedge G_\nu) \right) - \nabla \left(A \cdot J/c \right). \tag{F.151}
$$

Now, intuition tells me that it ought to be possible to simplify this further, in particular, eliminating the ν indices.

Think I will take a break from this for a while, and come back to it later.

F.12 NOMENCLATURE. LINEARIZED SPACETIME TRANSLATION.

Applying the translation $x^\mu \to x^\mu + e^\mu$, is what I thought would be called "spacetime translation". But to do so we need higher order powers of the exponential vector translation operator (ie: multivariable Taylor series operator)

$$\sum_k (1/k!)(e^\mu \partial_\mu)^k. \tag{F.152}$$

The transformation that appears to result in the canonical energy momentum tensor has only the linear term of this operator, so I called it "linearized spacetime translation operator", which seemed like a better name (to me). That is all. My guess is that what is typically referred to as the spacetime translation that generates the canonical energy momentum tensor is really just the first order term of the translation operation, and not truly a complete translation. If that is the case, then dropping the linearized adjective would probably be reasonable.

It is somewhat odd that the derived conditions for a divergence added to the Lagrangian are immediately busted by the wave equation. I think the saving grace is the fact that an arbitrary $\partial_\mu F^\mu$ is not necessarily a symmetry is the fact the translation of the coordinates is not an arbitrary divergence. This directional derivative operator is applied to the Lagrangian itself and not to an arbitrary function. This builds in the required symmetry (you could also add in or subtract out additional divergence terms that meet the derived conditions and not change anything).

Now, if the first order term of the Taylor expansion is a symmetry because we can commute the field partials and the coordinate partials then the higher order terms should also be symmetries. This would mean that a true translation $\mathcal{L} \to \exp(e^\mu \partial_\mu)\mathcal{L}$ would also be a symmetry. What conservation current would we get from that? Would it be the symmetric energy momentum tensor?

INDEX

BIBLIOGRAPHY

[1] John Baez. *Symmetry under boosts gives what conserved quantity* , 2006. URL http://math.ucr.edu/home/baez/boosts.html. [Online; accessed 22-May-2014]. (Cited on page 321.)

[2] C. Doran and A.N. Lasenby. *Geometric algebra for physicists*. Cambridge University Press New York, Cambridge, UK, 1st edition, 2003. (Cited on pages 77, 147, 148, 150, 153, 161, 163, 255, 269, 284, 324, 333, and 335.)

[3] R.P. Feynman, R.B. Leighton, and M.L. Sands. *Feynman lectures on physics.[Lectures on physics]*. Addison-Wesley Publishing Company. Reading, Massachusetts, 1963. (Cited on pages 15 and 352.)

[4] H. Goldstein. *Classical mechanics*. Cambridge: Addison-Wesley Press, Inc, 1st edition, 1951. (Cited on pages xi, 5, 16, 19, 155, 164, 209, 265, 267, 295, and 346.)

[5] D. Hestenes. *New Foundations for Classical Mechanics*. Kluwer Academic Publishers, 1999. (Cited on page 295.)

[6] E.A. Jackson. *Equilibrium statistical mechanics*. Dover Pubns, 2000. (Cited on page 21.)

[7] JD Jackson. *Classical Electrodynamics*. John Wiley and Sons, 2nd edition, 1975. (Cited on pages 20, 162, 203, 289, and 324.)

[8] Peeter Joot. *Exploring physics with Geometric Algebra, Part I.*, chapter Developing some intuition for Multivariable and Multivector Taylor Series. peeterjoot.com, 2009. URL https://peeterjoot.com/archives/math2015/gabookI.pdf. [Online; accessed 24-August-2023]. (Cited on page 330.)

[9] Peeter Joot. *Exploring physics with Geometric Algebra, Part I.*, chapter Generator of rotations in arbitrary dimensions. peeterjoot.com, 2009. URL https://peeterjoot.com/archives/math2015/gabookI.pdf. [Online; accessed 24-August-2023]. (Cited on page 277.)

[10] Peeter Joot. *Exploring physics with Geometric Algebra, Part II.*, chapter Energy momentum tensor relation to Lorentz force. peeterjoot.com, 2009. URL https://peeterjoot.com/archives/math2015/gabookII.pdf. [Online; accessed 24-August-2023]. (Cited on page 323.)

[11] Peeter Joot. *Exploring physics with Geometric Algebra, Part II.*, chapter Lorentz force relation to the energy momentum tensor. peeterjoot.com, 2009. URL https://peeterjoot.com/archives/math2015/gabookII.pdf. [Online; accessed 24-August-2023]. (Cited on page 323.)

[12] Peeter Joot. *Misc Physics and Math Play.*, chapter Worked calculus of variations problems from Byron and Fuller. peeterjoot.com, 2009. URL https://peeterjoot.com/archives/math2009/miscphysics.pdf. [Online; accessed 24-August-2023]. (Cited on page 10.)

[13] Peeter Joot. *Mathematica modules for Geometric Algebra's GA(2,0), GA(3,0), and GA(1,3)*, 2017. URL https://github.com/peeterjoot/gapauli. [Online; accessed 24-Oct-2020]. (Cited on page 127.)

[14] Peeter Joot. *Quantum field theory*. Kindle Direct Publishing, Toronto, 2018. (Cited on page 261.)

[15] Peeter Joot. *Geometric Algebra for Electrical Engineers*. Kindle Direct Publishing, Toronto, 2019. (Cited on pages 102 and 109.)

[16] LD Landau and EM Lifshitz. *Mechanics, vol. 1*. 1976. (Cited on pages xi and 61.)

[17] A. Lasenby, C. Doran, and S. Gull. A multivector derivative approach to Lagrangian field theory. *Foundations of Physics*, 23(10): 1295–1327, 1993. (Cited on page 15.)

[18] JA Lewis. *Mechanics: Berkeley Physics Course. vol. 1. Charles Kittel, Walter D. Knight, and Malvin A. Ruderman. McGraw-Hill, New York, 1965. xxii+ 480 pp. Illus.$ 5.50*. 1965 by the American Association for the Advancement of Science, 1965. (Cited on page 155.)

[19] A. Macdonald. *Vector and Geometric Calculus*. CreateSpace Independent Publishing Platform, 2012. (Cited on page 109.)

[20] Joseph A. Minahan. *Joe's Relatively Small Book of Special Relativity*, chapter Tensors. 2011. URL http://www.physics.uu.se/teorfys/sites/fysast.uu.se.teorfys/files/specrel.pdf. [Online; accessed 22-May-2014]. (Cited on page 315.)

[21] M. Montesinos and E. Flores. Symmetric energy-momentum tensor in Maxwell, Yang-Mills, and Proca theories obtained using only Noether's theorem. *Arxiv preprint hep-th/0602190*, 2006. (Cited on page 277.)

[22] Prof. William J Spence. *Special relativity, four vectors, .. covariant and contravariant vectors, tensors* , 2012. URL https://ph.qmul.ac.uk/sites/default/files/EMT7new.pdf. [Online; accessed 27-Dec-2022]. (Cited on page 315.)

[23] M.A. Srednicki. *Quantum Field Theory*. Cambridge University Press, 2007. (Cited on pages 266 and 324.)

[24] Dr. David Tong. *Classical Mechanics Lagrangian Problem Set 1.*, . URL http://www.damtp.cam.ac.uk/user/tong/dynamics/mf1.pdf. (Cited on pages 20, 61, and 205.)

[25] Dr. David Tong. *Quantum Field Theory.*, . URL http://www.damtp.cam.ac.uk/user/tong/qft.html. (Cited on pages 325 and 333.)

[26] Dr. David Tong. *Classical Mechanics.*, 2012. URL http://www.damtp.cam.ac.uk/user/tong/dynamics.htm. [Online; accessed 22-May-2014]. (Cited on pages 153 and 155.)