## Hamiltonian notes.

Originally appeared at:
http://sites.google.com/site/peeterjoot/math2009/hamiltonian.pdf?revision=6
Peeter Joot — peeter.joot@gmail.com
Sept 26, 2009 RCSfile : hamiltonian.tex, v Last Revision : 1.37 Date : 2009/10/2123: $43: 40$

## Contents

1 Motivation 1
2 Hamiltonian as a conserved quantity. 2
3 Some syntactic sugar. In vector form. 3
4 The Hamiltonian principle. 3
5 Examples. 5
5.1 Force free motion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
5.2 Linear potential (surface gravitation) . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
5.3 Harmonic oscillator (spring potential) . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
5.4 Harmonic oscillator (change of variables.) . . . . . . . . . . . . . . . . . . . . . . . . . 10
5.5 Force free system dependent on only differences. . . . . . . . . . . . . . . . . . . . . . 13
5.6 Gravitational potential . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
5.7 Pendulum . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
5.8 Spherical pendulum . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
5.9 Double and multiple pendulums, and general quadratic velocity dependence. . . . . 25
5.9.1 Single pendulum verification. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
5.9.2 Double pendulum explicitly. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
5.10 Dangling mass connected by string to another. . . . . . . . . . . . . . . . . . . . . . . 37
5.11 Non-covariant Lorentz force. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
5.11.1 Canonical momenta . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
5.11.2 Euler-Lagrange evaluation. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
5.11.3 Hamiltonian. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39
5.11.4 Hamiltonian equation evaluation. . . . . . . . . . . . . . . . . . . . . . . . . . 40
5.12 Covariant force free case. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41
5.13 Covariant Lorentz force. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41

1. Motivation

I've 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 [1].

Some notation will be invented along the way as reasonable, since I'd 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 2 N complex configuration space introduced in [2]. (NOT YET DONE).

## 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 $\mathcal{L}=\mathcal{L}\left(q^{i}, \dot{q}^{i}\right)$.

$$
\begin{aligned}
\frac{d \mathcal{L}}{d t} & =\frac{\partial \mathcal{L}}{\partial q^{i}} \frac{d q^{i}}{d t}+\frac{\partial \mathcal{L}}{\partial \dot{q}^{i}} \frac{d \dot{q}^{i}}{d t} \\
& =\dot{q}^{i} \frac{d}{d t} \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}}+\frac{\partial \mathcal{L}}{\partial \dot{q}^{i}} \frac{d \dot{q}^{i}}{d t} \\
& =\frac{d}{d t}\left(\dot{q}^{i} \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}}\right)
\end{aligned}
$$

We can therefore form the difference

$$
\begin{equation*}
\frac{d}{d t}\left(\dot{q}^{i} \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}}-\mathcal{L}\right)=0 \tag{1}
\end{equation*}
$$

and find that this quantity, labeled H , is a constant of motion for the system

$$
\begin{equation*}
H \equiv \dot{q}^{\dot{q}} \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}}-\mathcal{L}=\text { constant } \tag{2}
\end{equation*}
$$

We'll 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 $p^{i}=m q^{i}$ for the non-canonical momenta)

$$
\begin{equation*}
P_{i} \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}} \tag{3}
\end{equation*}
$$

and note that these are the coordinates of a sort of velocity gradient of the Lagrangian. We've 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

$$
\begin{equation*}
\nabla \mathcal{L}=\frac{d}{d t} \nabla_{v} \mathcal{L} \tag{4}
\end{equation*}
$$

where $\nabla_{v}=e^{i} \partial \mathcal{L} / \partial \dot{x}^{i}=e^{i} P_{i}, \nabla=e^{i} \partial / \partial x^{i}$, and $x=e_{i} x^{i}$.

## 3. Some syntactic sugar. In vector form.

Following Jackson [3] (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 $q^{i \prime} s$ is given its own symbol P. He writes

$$
\begin{equation*}
H=\mathbf{v} \cdot \mathbf{P}-\mathcal{L} \tag{5}
\end{equation*}
$$

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

$$
\begin{align*}
& q=e_{i} q^{i}  \tag{6}\\
& v=f_{i} \dot{q}^{i} \tag{7}
\end{align*}
$$

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

$$
\begin{align*}
e^{i} \cdot e_{j} & =\delta^{i}{ }_{j}  \tag{8}\\
f^{i} \cdot f_{j} & =\delta^{i}{ }_{j}  \tag{9}\\
e^{i} \cdot f_{j} & =\delta^{i}{ }_{j} \tag{10}
\end{align*}
$$

We can now define configuration space position and velocity gradients

$$
\begin{align*}
\nabla & \equiv e^{i} \frac{\partial}{\partial q^{i}}  \tag{11}\\
\nabla_{v} & \equiv f^{i} \frac{\partial}{\partial \dot{q}^{i}} \tag{12}
\end{align*}
$$

so the conjugate momenta in vector form is now

$$
\begin{equation*}
P \equiv \nabla_{v} \mathcal{L}=f^{i} \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}} \tag{13}
\end{equation*}
$$

Our Hamiltonian takes the form

$$
\begin{equation*}
H=v \cdot P-\mathcal{L} \tag{14}
\end{equation*}
$$

## 4. The Hamiltonian principle.

We want to take partials of 2 with respect to $P_{i}$ and $q^{i}$. In terms of the canonical momenta we want to differentiate

$$
\begin{equation*}
H \equiv \dot{q}^{i} P_{i}-\mathcal{L}\left(q^{i}, \dot{q}^{i}, t\right) \tag{15}
\end{equation*}
$$

for the $P_{i}$ partial we have

$$
\frac{\partial H}{\partial P_{i}}=\dot{q}^{i}
$$

and for the $q^{i}$ partial

$$
\begin{aligned}
\frac{\partial H}{\partial q^{i}} & =-\frac{\partial \mathcal{L}}{\partial q^{i}} \\
& =-\frac{d}{d t} \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}}
\end{aligned}
$$

These two results taken together form what I believe is called the Hamiltonian principle

$$
\begin{align*}
\frac{\partial H}{\partial P_{i}} & =\dot{q}^{i}  \tag{16}\\
\frac{\partial H}{\partial q^{i}} & =-\dot{P}_{i}  \tag{17}\\
P_{i} & =\frac{\partial \mathcal{L}}{\partial \dot{q}^{i}} \tag{18}
\end{align*}
$$

A set of 2 N first order equations equivalent to the second order Euler-Lagrange equations. These appear to follow straight from the definitions. Given that I'm curious why the more complex method of derivation is chosen in [4]. There the total differential of the Hamiltonian is computed

$$
\begin{aligned}
d H & =\dot{q}^{i} d P_{i}+d \dot{q}^{i} P_{i}-d q^{i} \frac{\partial \mathcal{L}}{\partial q^{i}}-d \dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}^{i}}-d t \frac{\partial \mathcal{L}}{\partial t} \\
& =\dot{q}^{i} d P_{i}+d \dot{q}^{i}\left(P_{i}-\frac{\partial \mathcal{L}}{\partial \dot{q}^{i}}\right)-d q^{i} \frac{\partial \mathcal{L}}{\partial q^{i}}-d t \frac{\partial \mathcal{L}}{\partial t} \\
& =\dot{q}^{i} d P_{i}-d q^{i} \underbrace{\frac{\partial \mathcal{L}}{\partial q^{i}}}_{=d / d t P_{i}}-d t \frac{\partial \mathcal{L}}{\partial t}
\end{aligned}
$$

A term by term comparison to the total differential written out explicitly

$$
\begin{equation*}
d H=\frac{\partial H}{\partial q^{i}} d q^{i}+\frac{\partial H}{\partial P_{i}} d P_{i}+\frac{\partial H}{\partial t} d t \tag{19}
\end{equation*}
$$

allows the Hamiltonian equations to be picked off.

$$
\begin{align*}
& \frac{\partial H}{\partial P_{i}}=\dot{q}^{i}  \tag{20}\\
& \frac{\partial H}{\partial q^{i}}=-\dot{P}_{i}  \tag{21}\\
& \frac{\partial H}{\partial t}=-\frac{\partial \mathcal{L}}{\partial t} \tag{22}
\end{align*}
$$

I guess that isn't that much more complicated and it does yield a relation between the Hamiltonian and Lagrangian time derivatives.

## 5. Examples.

Now, that's just about the most abstract way we can start things off isn't it? Getting some initial feel for this constant of motion can be had by considering a sequence of Lagrangians, starting with the very simplest.

### 5.1. Force free motion

Our very simplest Lagrangian is that of one dimensional purely kinetic motion

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m v^{2}=\frac{1}{2} m \dot{x}^{2} \tag{23}
\end{equation*}
$$

Our Hamiltonian is in this case just

$$
\begin{equation*}
H=\dot{x} m \dot{x}-\frac{1}{2} m \dot{x}=\frac{1}{2} m v^{2} \tag{24}
\end{equation*}
$$

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

$$
\begin{equation*}
H=\frac{1}{2 m} p^{2} \tag{25}
\end{equation*}
$$

We have then

$$
\begin{aligned}
& \frac{\partial H}{\partial p}=\frac{p}{m}=\dot{x} \\
& \frac{\partial H}{\partial x}=0=-\dot{p}
\end{aligned}
$$

Just for fun we can put this simple linear system in matrix form

$$
\frac{d}{d t}\left[\begin{array}{c}
p  \tag{26}\\
x
\end{array}\right]=\frac{1}{m}\left[\begin{array}{ll}
0 & 0 \\
1 & 0
\end{array}\right]\left[\begin{array}{c}
p \\
x
\end{array}\right]
$$

A linear system of this form $y^{\prime}=A y$ can be solved by exponentiation with solution

$$
\begin{equation*}
y=e^{A t} y_{0} \tag{27}
\end{equation*}
$$

In this case our matrix is nilpotent degree 2 so we can exponentiate only requiring up to the first order power

$$
\begin{equation*}
e^{A t}=I+A t \tag{28}
\end{equation*}
$$

specifically

$$
\left[\begin{array}{c}
p  \tag{29}\\
x
\end{array}\right]=\left[\begin{array}{ll}
1 & 0 \\
\frac{t}{m} & 1
\end{array}\right]\left[\begin{array}{l}
p_{0} \\
x_{0}
\end{array}\right]
$$

Written out in full this is just

$$
\begin{align*}
& p=p_{0}  \tag{30}\\
& x=\frac{p_{0}}{m} t+x_{0} \tag{31}
\end{align*}
$$

Since the canonical momentum is the regular momentum $p=m v$ 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'd 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

$$
\begin{gather*}
H=\frac{1}{m} \mathbf{p}^{2}  \tag{32}\\
\frac{d}{d t}\left[\begin{array}{c}
p_{x} \\
x \\
p_{y} \\
y \\
p_{z} \\
z
\end{array}\right]=\frac{1}{m}\left[\begin{array}{lllll}
0 & 0 & & & \\
1 & 0 & & & \\
& & 0 & 0 & \\
& & 1 & 0 & \\
& & & & 0
\end{array}\right]  \tag{33}\\
\\
\\
\\
\end{gather*}
$$

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 26, and we have

$$
\begin{align*}
& p_{i}(t)=p_{i}(0)  \tag{34}\\
& x_{i}(t)=\frac{p_{i}(0)}{m} t+x_{i}(0) \tag{35}
\end{align*}
$$

Or just

$$
\begin{align*}
& \mathbf{p}(t)=\mathbf{p}(0)  \tag{36}\\
& \mathbf{x}(t)=\frac{\mathbf{p}(0)}{m} t+\mathbf{x}(0) \tag{37}
\end{align*}
$$

### 5.2. Linear potential (surface gravitation)

For the gravitational force $F=-m g \hat{\mathbf{z}}=-\boldsymbol{\nabla} \phi$, we have $\phi=m g z$, and a Lagrangian of

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m \mathbf{v}^{2}-\phi=\frac{1}{2} m \mathbf{v}^{2}-m g z \tag{38}
\end{equation*}
$$

Without velocity dependence the canonical momentum is the momentum $m \mathbf{v}$, and our Hamiltonian is

$$
\begin{equation*}
H=\frac{1}{2 m} \mathbf{p}^{2}+m g z \tag{39}
\end{equation*}
$$

The Hamiltonian equations are

$$
\begin{gather*}
\frac{\partial H}{\partial p_{i}}=\dot{x}_{i}=\frac{1}{m} p_{i}  \tag{40}\\
\sigma_{i} \frac{\partial H}{\partial x_{i}}=-\sigma_{i} \dot{p}_{i}=\left[\begin{array}{c}
0 \\
0 \\
m g
\end{array}\right] \tag{41}
\end{gather*}
$$

In matrix form we have

$$
\frac{d}{d t}\left[\begin{array}{c}
p_{x}  \tag{42}\\
x \\
p_{y} \\
y \\
p_{z} \\
z
\end{array}\right]=\frac{1}{m}\left[\begin{array}{llllll}
0 & 0 & & & & \\
1 & 0 & & & & \\
& & 0 & 0 & & \\
& & 1 & 0 & & \\
& & & & 0 & 0 \\
& & & & 1 & 0
\end{array}\right]\left[\begin{array}{c}
p_{x} \\
x \\
p_{y} \\
y \\
p_{z} \\
z
\end{array}\right]+\left[\begin{array}{c}
0 \\
0 \\
0 \\
0 \\
-m g \\
0
\end{array}\right]
$$

So our problem is now reduced to solving a linear system of the form

$$
\begin{equation*}
y^{\prime}=A y+b \tag{43}
\end{equation*}
$$

That extra little term $b$ throws a wrench into things and I'm 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've solved the problem for the $x$ and $y$ coordinates in the force free case. That leaves

$$
\frac{d}{d t}\left[\begin{array}{c}
p_{z}  \tag{44}\\
z
\end{array}\right]=\frac{1}{m}\left[\begin{array}{ll}
0 & 0 \\
1 & 0
\end{array}\right]\left[\begin{array}{c}
p_{z} \\
z
\end{array}\right]+\left[\begin{array}{c}
-m g \\
0
\end{array}\right]
$$

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

$$
\begin{equation*}
p_{z}=-m g t+p_{z}(0) \tag{45}
\end{equation*}
$$

This can be substituted into the position equation for

$$
\begin{equation*}
\dot{z}=\frac{1}{m}\left(p_{z}(0)-m g t\right) \tag{46}
\end{equation*}
$$

Direct integration is now possible for the final solution

$$
\begin{aligned}
z & =\frac{1}{m}\left(p_{z}(0) t-m g t^{2} / 2\right)+z_{0} \\
& =\frac{p_{z}(0)}{m} t-\frac{g}{2} t^{2}+z_{0}
\end{aligned}
$$

Again something that we could have gotten in many easier ways. Using the result we see that the solution to 44 in matrix form, again with $A=\frac{1}{m}\left[\begin{array}{ll}0 & 0 \\ 1 & 0\end{array}\right]$ is

$$
\left[\begin{array}{c}
p_{z}  \tag{47}\\
z
\end{array}\right]=e^{A t}\left[\begin{array}{c}
p_{z}(0) \\
z(0)
\end{array}\right]-m g\left[\begin{array}{c}
t \\
\frac{1}{2 m} t^{2}
\end{array}\right]
$$

I thought if I wrote this out how to solve 43 may be more obvious, but that path is still unclear. If $A$ were invertible, which it isn't, then writing $b=A c$ would allow for a change of variables. Does this matter for consideration of a physical problem. Not really, so I'll fight the urge to play with the math for a while and perhaps revisit this later separately.

### 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 $F=-k x \hat{\mathbf{x}}=-\nabla \phi$, we have $\phi=k x^{2} / 2$, and a Lagrangian of

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m \mathbf{v}^{2}-\frac{1}{2} k \mathbf{x}^{2} \tag{48}
\end{equation*}
$$

Our Hamiltonian is again just the total energy

$$
\begin{equation*}
H=\frac{1}{2 m} \mathbf{p}^{2}+\frac{1}{2} k \mathbf{x}^{2} \tag{49}
\end{equation*}
$$

Evaluating the Hamiltonian equations we have

$$
\begin{align*}
& \frac{\partial H}{\partial p_{i}}=\dot{x}_{i}=p_{i} / m  \tag{50}\\
& \frac{\partial H}{\partial x_{i}}=-\dot{p}_{i}=k x_{i} \tag{51}
\end{align*}
$$

Considering just the $x$ dimension (the others have the free particle behavior), our matrix phase space representation is

$$
\frac{d}{d t}\left[\begin{array}{l}
p  \tag{52}\\
x
\end{array}\right]=\left[\begin{array}{cc}
0 & -k \\
1 / m & 0
\end{array}\right]\left[\begin{array}{l}
p \\
x
\end{array}\right]
$$

So with

$$
A=\left[\begin{array}{cc}
0 & -k  \tag{53}\\
1 / m & 0
\end{array}\right]
$$

Our solution is

$$
\left[\begin{array}{c}
p  \tag{54}\\
x
\end{array}\right]=e^{A t}\left[\begin{array}{l}
p_{0} \\
x_{0}
\end{array}\right]
$$

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

$$
\left[\begin{array}{c}
p  \tag{55}\\
x
\end{array}\right]_{i+1}=e^{A \Delta t_{i+1}}\left[\begin{array}{c}
p \\
x
\end{array}\right]_{i}
$$

Or

$$
\left[\begin{array}{c}
p  \tag{56}\\
x
\end{array}\right]_{i+1}=e^{A \Delta t_{i+1}} e^{A \Delta t_{i}} \cdots e^{A \Delta t_{1}}\left[\begin{array}{l}
p \\
x
\end{array}\right]_{0}
$$

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

$$
A^{2}=\left[\begin{array}{cc}
0 & -k  \tag{57}\\
1 / m & 0
\end{array}\right]\left[\begin{array}{cc}
0 & -k \\
1 / m & 0
\end{array}\right]=-\frac{k}{m} I
$$

The first few powers are

$$
\begin{align*}
A^{2} & =-\frac{k}{m} I  \tag{58}\\
A^{3} & =-\frac{k}{m} A  \tag{59}\\
A^{4} & =\left(\frac{k}{m}\right)^{2} I  \tag{60}\\
A^{5} & =\left(\frac{k}{m}\right)^{2} A \tag{61}
\end{align*}
$$

So exponentiating we can collect cosine and sine terms

$$
\begin{aligned}
e^{A t} & =I\left(1-\frac{k}{m} \frac{t^{2}}{2!}+\left(\frac{k}{m}\right)^{2} \frac{t^{4}}{4!}+\cdots\right)+A \sqrt{\frac{m}{k}}\left(\sqrt{\frac{k}{m}}-\left(\sqrt{\frac{k}{m}}\right)^{3} \frac{t^{3}}{3!}+\left(\sqrt{\frac{k}{m}}\right)^{5} \frac{t^{5}}{5!}\right) \\
& =I \cos \left(\sqrt{\frac{k}{m}} t\right)+\sqrt{\frac{m}{k}} A \sin \left(\sqrt{\frac{k}{m}} t\right)
\end{aligned}
$$

As a check it is readily verified that this satisfies the desired $d\left(e^{A t}\right) / d t=A e^{A t}$ property. The full solution in phase space representation is therefore

$$
\left[\begin{array}{c}
p  \tag{62}\\
x
\end{array}\right]=\left[\begin{array}{l}
p_{0} \\
x_{0}
\end{array}\right] \cos \left(\sqrt{\frac{k}{m}} t\right)+\sqrt{\frac{m}{k}}\left[\begin{array}{c}
-k x_{0} \\
p_{0} / m
\end{array}\right] \sin \left(\sqrt{\frac{k}{m}} t\right)
$$

Written out separately this is clearer

$$
\begin{align*}
& p=p_{0} \cos \left(\sqrt{\frac{k}{m}} t\right)-\sqrt{\frac{m}{k}} k x_{0} \sin \left(\sqrt{\frac{k}{m}} t\right)  \tag{63}\\
& x=x_{0} \cos \left(\sqrt{\frac{k}{m}} t\right)+\sqrt{\frac{m}{k}} \frac{p_{0}}{m} \sin \left(\sqrt{\frac{k}{m}} t\right) \tag{64}
\end{align*}
$$

One can readily verify that $m \dot{x}=p$, and $m \ddot{x}=-k x$ as expected.
Let's pause before leaving the harmonic oscillator to see if 63 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.

### 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

$$
\begin{align*}
& P=x \sqrt{\frac{k}{2}}+\frac{p}{\sqrt{2 m}}  \tag{65}\\
& Q=x \sqrt{\frac{k}{2}}-\frac{p}{\sqrt{2 m}} \tag{66}
\end{align*}
$$

In particular the Hamiltonian is then just

$$
\begin{equation*}
H=P^{2}+Q^{2} \tag{67}
\end{equation*}
$$

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 don't however, need the rotation to do that. Suppose we write for just the scaling change of variables

$$
\begin{align*}
p & =\sqrt{2 m} P_{s}  \tag{68}\\
x & =\sqrt{\frac{2}{k}} Q_{s} \tag{69}
\end{align*}
$$

or

$$
\left[\begin{array}{c}
p  \tag{70}\\
x
\end{array}\right]=\left[\begin{array}{cc}
\sqrt{2 m} & 0 \\
0 & \sqrt{\frac{2}{k}}
\end{array}\right]\left[\begin{array}{l}
P_{s} \\
Q_{s}
\end{array}\right]
$$

This also gives the Hamiltonian 67, and the Hamiltonian equations are transformed to

$$
\begin{aligned}
\frac{d}{d t}\left[\begin{array}{l}
P_{s} \\
Q_{s}
\end{array}\right] & =\left[\begin{array}{cc}
1 / \sqrt{2 m} & 0 \\
0 & \sqrt{\frac{k}{2}}
\end{array}\right]\left[\begin{array}{cc}
0 & -k \\
1 / m & 0
\end{array}\right]\left[\begin{array}{cc}
\sqrt{2 m} & 0 \\
0 & \sqrt{\frac{2}{k}}
\end{array}\right]\left[\begin{array}{l}
P_{s} \\
Q_{s}
\end{array}\right] \\
& =\left[\begin{array}{cc}
0 & -\sqrt{\frac{k}{m}} \\
\sqrt{\frac{k}{m}} & 0
\end{array}\right]\left[\begin{array}{c}
P_{s} \\
Q_{s}
\end{array}\right]
\end{aligned}
$$

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

$$
\begin{equation*}
\omega \equiv \sqrt{\frac{k}{m}} \tag{71}
\end{equation*}
$$

and our system is reduced to

$$
\frac{d}{d t}\left[\begin{array}{l}
P_{s}  \tag{72}\\
Q_{s}
\end{array}\right]=\omega\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]\left[\begin{array}{l}
P_{s} \\
Q_{s}
\end{array}\right]
$$

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 $\sigma_{2}$. That is

$$
\sigma_{2}=i\left[\begin{array}{cc}
0 & -1  \tag{73}\\
1 & 0
\end{array}\right]
$$

Inverting this, and labeling this matrix $\mathcal{I}$ we can write

$$
\mathcal{I} \equiv\left[\begin{array}{cc}
0 & -1  \tag{74}\\
1 & 0
\end{array}\right]=-i \sigma_{2}
$$

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

$$
\frac{d}{d t}\left[\begin{array}{l}
P_{s}  \tag{75}\\
Q_{s}
\end{array}\right]=\mathcal{I} \omega\left[\begin{array}{l}
P_{s} \\
Q_{s}
\end{array}\right]
$$

We can now solve the system directly. Writing $\mathbf{z}_{s}=\binom{P_{s}}{Q_{s}}$, this is just

$$
\begin{equation*}
\mathbf{z}_{s}(t)=e^{\mathcal{I} \omega t} \mathbf{z}_{s}(0)=(I \cos (\omega t)+\mathcal{I} \sin (\omega t)) \mathbf{z}_{s}(0) \tag{76}
\end{equation*}
$$

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

$$
\begin{align*}
P & =Q_{s}+P_{s}  \tag{77}\\
Q & =Q_{s}-P_{s} \tag{78}
\end{align*}
$$

Inverting this we have

$$
\left[\begin{array}{l}
P_{s}  \tag{79}\\
Q_{s}
\end{array}\right]=\frac{1}{2}\left[\begin{array}{cc}
1 & -1 \\
1 & 1
\end{array}\right]\left[\begin{array}{l}
P \\
Q
\end{array}\right]
$$

The Hamiltonian after this change of variables is now

$$
\frac{d}{d t}\left[\begin{array}{l}
P  \tag{80}\\
Q
\end{array}\right]=\frac{\omega}{2}\left[\begin{array}{cc}
1 & 1 \\
-1 & 1
\end{array}\right]\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]\left[\begin{array}{cc}
1 & -1 \\
1 & 1
\end{array}\right]\left[\begin{array}{l}
P \\
Q
\end{array}\right]
$$

But multiplying this out one finds that the equations of motion for the state space vector are unchanged by the rotation, and writing $\mathbf{z}=\binom{P}{Q}$ for the state vector, the Hamiltonian equations are

$$
\begin{equation*}
\mathbf{z}^{\prime}=\mathcal{I} \omega \mathbf{z} \tag{81}
\end{equation*}
$$

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.

### 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,

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m_{1} \dot{r}_{1}^{2}+\frac{1}{2} m_{2} \dot{r}_{2}^{2} \tag{82}
\end{equation*}
$$

With a change of variables

$$
\begin{align*}
& x=r_{2}-r_{1}  \tag{83}\\
& y=r_{2} \tag{84}
\end{align*}
$$

and elimination of $r_{1}$, and $r_{2}$ we have

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m_{1}(\dot{y}-\dot{x})^{2}+\frac{1}{2} m_{2} \dot{y}^{2} \tag{85}
\end{equation*}
$$

We now need the conjugate momenta in terms of $\dot{x}$ and $\dot{y}$. These are

$$
\begin{align*}
& P_{x}=\frac{\partial \mathcal{L}}{\partial \dot{x}}=-m_{1}(\dot{y}-\dot{x})  \tag{86}\\
& P_{y}=\frac{\partial \mathcal{L}}{\partial \dot{y}}=m_{1}(\dot{y}-\dot{x})+m_{2} \dot{y} \tag{87}
\end{align*}
$$

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 86 . That is

$$
\left[\begin{array}{c}
\dot{x}  \tag{88}\\
\dot{y}
\end{array}\right]=\left[\begin{array}{cc}
m_{1} & -m_{1} \\
-m_{1} & \left(m_{1}+m_{2}\right)
\end{array}\right]^{-1}\left[\begin{array}{l}
P_{x} \\
P_{y}
\end{array}\right]
$$

This is

$$
\left[\begin{array}{c}
\dot{x}  \tag{89}\\
\dot{y}
\end{array}\right]=\frac{1}{m_{1}} \frac{1}{m_{2}}\left[\begin{array}{cc}
m_{1}+m_{2} & m_{1} \\
m_{1} & m_{1}
\end{array}\right]\left[\begin{array}{c}
P_{x} \\
P_{y}
\end{array}\right]
$$

Of these only $\dot{y}$ and $\dot{y}-\dot{x}$ are of interest and after a bit of manipulation we find

$$
\begin{align*}
& \dot{y}=\frac{1}{m_{2}}\left(P_{x}+P_{y}\right)  \tag{90}\\
& \dot{x}=\frac{1}{m_{1}} \frac{1}{m_{2}}\left(\left(m_{1}+m_{2}\right) P_{x}+m_{1} P_{y}\right) \tag{91}
\end{align*}
$$

From this we find the Lagrangian in terms of the conjugate momenta

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2 m_{1}} P_{x}^{2}+\frac{1}{2 m_{2}}\left(P_{x}+P_{y}\right)^{2} \tag{92}
\end{equation*}
$$

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

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m_{1} \dot{\mathbf{r}}_{1}^{2}+\frac{1}{2} m_{2} \dot{\mathbf{r}}_{2}^{2}-\phi\left(\mathbf{x}_{1}-\mathbf{x}_{2}\right) \tag{93}
\end{equation*}
$$

With

$$
\begin{align*}
& \mathbf{x}=\mathbf{x}_{1}-\mathbf{x}_{2}  \tag{94}\\
& \mathbf{y}=\mathbf{x}_{2} \tag{95}
\end{align*}
$$

The 3D generalization of the above (following by adding indexes then summing) becomes

$$
\begin{gather*}
\mathbf{P}_{x}=\sigma_{j} \frac{\partial \mathcal{L}}{\partial \dot{x}^{j}}=-m_{1}(\dot{\mathbf{y}}-\dot{\mathbf{x}})  \tag{96}\\
\mathbf{P}_{y}=\sigma_{j} \frac{\partial \mathcal{L}}{\partial \dot{y}^{j}}=m_{1}(\dot{\mathbf{y}}-\dot{\mathbf{x}})+m_{2} \dot{\mathbf{y}}  \tag{97}\\
\mathcal{L}=\frac{1}{2 m_{1}} \mathbf{P}_{x}{ }^{2}+\frac{1}{2 m_{2}}\left(\mathbf{P}_{x}+\mathbf{P}_{y}\right)^{2}-\phi(\mathbf{x})  \tag{98}\\
H=\frac{1}{2 m_{1}} \mathbf{P}_{x}{ }^{2}+\frac{1}{2 m_{2}}\left(\mathbf{P}_{x}+\mathbf{P}_{y}\right)^{2}+\phi(\mathbf{x}) \tag{99}
\end{gather*}
$$

Finally, evaluation of the Hamiltonian equations we have

$$
\begin{aligned}
\sigma_{j} \frac{\partial H}{\partial P_{x}^{j}} & =\dot{\mathbf{x}} \\
& =\sigma_{j}\left(\frac{1}{m_{1}} P_{x}^{j}+\frac{1}{m_{2}}\left(P_{x}^{j}+P_{y}^{j}\right)\right) \\
& =\frac{1}{m_{1}} \mathbf{P}_{x}+\frac{1}{m_{2}}\left(\mathbf{P}_{x}+\mathbf{P}_{y}\right) \\
\sigma_{j} \frac{\partial H}{\partial P_{y}^{j}} & =\dot{\mathbf{y}} \\
& =\sigma_{j} \frac{1}{m_{2}}\left(P_{x}^{j}+P_{y}^{j}\right) \\
& =\frac{1}{m_{2}}\left(\mathbf{P}_{x}+\mathbf{P}_{y}\right)
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{j} \frac{\partial H}{\partial x^{j}} & =-\dot{\mathbf{P}}_{x} \\
& =-\sigma_{j} \frac{\partial \phi}{\partial x^{j}} \\
& =-\nabla_{\mathbf{x}} \phi(\mathbf{x}) \\
\sigma_{j} \frac{\partial H}{\partial y^{j}} & =-\dot{\mathbf{P}}_{y} \\
& =-\sigma_{j} \frac{\partial \phi}{\partial y^{j}} \\
& =0
\end{aligned}
$$

Summarizing we have four first order equations

$$
\begin{align*}
\dot{\mathbf{x}} & =\left(\frac{1}{m_{1}}+\frac{1}{m_{2}}\right) \mathbf{P}_{x}+\frac{1}{m_{2}} \mathbf{P}_{y}  \tag{100}\\
\dot{\mathbf{y}} & =\frac{1}{m_{2}}\left(\mathbf{P}_{x}+\mathbf{P}_{y}\right)  \tag{101}\\
\dot{\mathbf{P}}_{x} & =\nabla_{\mathbf{x}} \phi(\mathbf{x})  \tag{102}\\
\dot{\mathbf{P}}_{y} & =0 \tag{103}
\end{align*}
$$

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.

### 5.6. Gravitational potential

Next I'd 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'd 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

$$
\begin{align*}
& \text { Force on } 2=-G m_{1} m_{2} \frac{\left(\mathbf{r}_{2}-\mathbf{r}_{1}\right)}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|}  \tag{104}\\
& \text { Force on } 1=G m_{1} m_{2} \frac{\left(\mathbf{r}_{2}-\mathbf{r}_{1}\right)}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|} \tag{105}
\end{align*}
$$

Guessing that the Lagrangian has a single term for the interaction potential

$$
\begin{equation*}
\phi_{21}=\kappa \frac{1}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|} \tag{106}
\end{equation*}
$$

so that we have

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m \mathbf{v}_{1}{ }^{2}+\frac{1}{2} m \mathbf{v}_{2}^{2}-\phi_{21} \tag{107}
\end{equation*}
$$

We can evaluate the Euler-Lagrange equations and see if the result is consistent with the Newtonian force laws of 104 Suppose we write the coordinates of $\mathbf{r}_{i}$ as $x^{k}{ }_{i}$. There are then six EulerLagrange equations

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x^{j}} & =\frac{d}{d t} \frac{\partial \mathcal{L}}{\partial \dot{x}_{i}^{j}} \\
-\frac{\partial \phi_{21}}{\partial x^{j_{i}}} & =m_{i} \ddot{x}_{i}^{j}
\end{aligned}
$$

Evaluating the potential derivatives separately. Consider the $i=2$ derivative

$$
\begin{aligned}
\frac{\partial \phi_{21}}{\partial x^{j}{ }_{2}} & =\kappa \frac{\partial}{\partial x^{j_{2}}}\left(\sum_{k}\left(x^{k}{ }_{2}-x^{k}{ }_{1}\right)^{2}\right)^{-1 / 2} \\
& =-\kappa \frac{1}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|^{3}} \sum_{k}\left(x^{k}{ }_{2}-x^{k}{ }_{1}\right) \frac{\partial}{\partial x^{j}{ }_{2}}\left(x^{k}{ }_{2}-x^{k}{ }_{1}\right) \\
& =-\kappa \frac{1}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|^{3}}\left(x^{j}{ }_{2}-x^{j}{ }_{1}\right)
\end{aligned}
$$

Therefore the final result of the Euler-Lagrange equations is

$$
\begin{align*}
\kappa \frac{1}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|^{3}}\left(x^{j}{ }_{2}-x^{j}{ }_{1}\right) & =m_{2} \ddot{x}_{2}^{j}  \tag{108}\\
-\kappa \frac{1}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|^{3}}\left(x^{j}{ }_{2}-x^{j}{ }_{1}\right) & =m_{1} \ddot{x}_{1}^{j} \tag{109}
\end{align*}
$$

which confirms the Lagrangian and potential guess and fixes the constant $\kappa=-G m_{1} m_{2}$. With the sign fixed, our potential, Lagrangian, and Hamiltonian are respectively

$$
\begin{align*}
\phi_{21} & =-\frac{G m_{2} m_{1}}{\left|\mathbf{r}_{2}-\mathbf{r}_{1}\right|}  \tag{110}\\
\mathcal{L} & =\frac{1}{2} m_{1} \mathbf{v}_{1}^{2}+\frac{1}{2} m_{2} \mathbf{v}_{2}^{2}-\phi_{21}  \tag{111}\\
H & =\frac{1}{2 m_{1}} \mathbf{p}_{1}^{2}+\frac{1}{2 m_{2}} \mathbf{p}_{2}^{2}+\phi_{21} \tag{112}
\end{align*}
$$

There is however an undesirable asymmetry to this expression, in particular a formulation that extends to multiple particles seems desirable. Let's write instead a slight variation

$$
\begin{equation*}
\phi_{i j}=-\frac{G m_{i} m_{j}}{\left|\mathbf{r}_{i}-\mathbf{r}_{j}\right|} \tag{113}
\end{equation*}
$$

and form a scaled by two double summation over all pairs of potentials

$$
\begin{equation*}
\mathcal{L}=\sum_{i} \frac{1}{2} m_{i} \mathbf{v}_{i}^{2}-\frac{1}{2} \sum_{i \neq j} \phi_{i j} \tag{114}
\end{equation*}
$$

Having established what seems like an appropriate form for the Lagrangian, we can write the Hamiltonian for the multiparticle gravitational interaction by inspection

$$
\begin{equation*}
H=\sum_{i} \frac{1}{2 m_{i}} \mathbf{p}_{i}^{2}+\frac{1}{2} \sum_{i \neq j} \phi_{i j} \tag{115}
\end{equation*}
$$

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 $k$ th component of the $i$ th particle

$$
\begin{gather*}
\frac{\partial H}{\partial p_{i}^{k}}=\dot{x}_{i}^{k}=\frac{1}{m_{i}} p_{i}^{k}  \tag{116}\\
\frac{\partial H}{\partial x^{k}{ }_{i}}=-\dot{p}_{i}^{k}=G \sum_{j \neq i} m_{i} m_{j} \frac{x^{k}{ }_{i}-x^{k}{ }_{j}}{\left|\mathbf{r}_{i}-\mathbf{r}_{j}\right|^{3}} \tag{117}
\end{gather*}
$$

The state space vector for this system of equations is brutally ugly, and could be put into the following form for example

$$
\mathbf{z}=\left[\begin{array}{c}
p_{1}{ }_{1}{ }_{1}  \tag{118}\\
p^{2} \\
p_{1}{ }_{1} \\
x_{1}{ }_{1} \\
x_{1}{ }_{1}{ }_{1} \\
x^{3}{ }_{1} \\
p_{1}{ }_{2} \\
p^{2} \\
p_{2}{ }_{2} \\
x^{1}{ }_{2} \\
\vdots \\
\vdots
\end{array}\right]
$$

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

$$
\begin{equation*}
\frac{d \mathbf{z}}{d t}=A(\mathbf{z}) \tag{119}
\end{equation*}
$$

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

$$
\begin{equation*}
\mathbf{z}\left(t_{0}+\Delta t\right)=\mathbf{z}\left(t_{0}\right)+A\left(\mathbf{z}\left(t_{0}\right)\right) \Delta t \tag{120}
\end{equation*}
$$

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 didn't have the redundancy built into the Hamiltonian equations).

FIXME: should be able to solve this for an arbitrary $\Delta t$ 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?

### 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 $m g h=m g l(1-\cos \theta)$. The Lagrangian is therefore

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m l^{2} \dot{\theta}^{2}-m g l(1-\cos \theta) \tag{121}
\end{equation*}
$$

The constant $m g l$ term can be dropped, and our canonical momentum conjugate to $\dot{\theta}$ is $p_{\theta}=$ $m l^{2} \dot{\theta}$, so our Hamiltonian is

$$
\begin{equation*}
H=\frac{1}{2 m l^{2}} p_{\theta}{ }^{2}-m g l \cos \theta \tag{123}
\end{equation*}
$$

We can now compute the Hamiltonian equations

$$
\begin{align*}
& \frac{\partial H}{\partial p_{\theta}}=\dot{\theta}=\frac{1}{m l^{2}} p_{\theta}  \tag{125}\\
& \frac{\partial H}{\partial q}=-\dot{p}_{\theta}=m g l \sin \theta \tag{126}
\end{align*}
$$

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

$$
\begin{equation*}
\sin \theta \approx \sin \theta_{0}+\cos \theta_{0} \alpha \tag{127}
\end{equation*}
$$

The linear approximation of the Hamiltonian equations after a change of variables become

$$
\frac{d}{d t}\left[\begin{array}{c}
p_{\theta}  \tag{128}\\
\alpha
\end{array}\right]=\left[\begin{array}{cc}
0 & -m g l \cos \theta_{0} \\
1 / m l^{2} & 0
\end{array}\right]\left[\begin{array}{c}
p_{\theta} \\
\alpha
\end{array}\right]+\left[\begin{array}{c}
-m g l \sin \theta_{0} \\
\dot{\theta}_{0}
\end{array}\right]
$$

A change of variables that scales the factors in the matrix to have equal magnitude and equivalent dimensions is helpful. Writing

$$
\left[\begin{array}{c}
p_{\theta}  \tag{129}\\
\alpha
\end{array}\right]=\left[\begin{array}{ll}
a & 0 \\
0 & 1
\end{array}\right] \mathbf{z}
$$

one finds

$$
\frac{d \mathbf{z}}{d t}=\left[\begin{array}{cc}
0 & -m g l \cos \theta_{0} / a  \tag{130}\\
a / m l^{2} & 0
\end{array}\right] \mathbf{z}+\frac{1}{a}\left[\begin{array}{c}
-m g l \sin \theta_{0} \\
\dot{\theta}_{0}
\end{array}\right]
$$

To tidy this up, we want

$$
\begin{equation*}
\left|\frac{a}{m l^{2}}\right|=\left|\frac{m g l \cos \theta_{0}}{a}\right| \tag{131}
\end{equation*}
$$

Or

$$
\begin{equation*}
a=m l^{2} \sqrt{\frac{g}{l}\left|\cos \theta_{0}\right|} \tag{132}
\end{equation*}
$$

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

$$
\frac{d \mathbf{z}}{d t}=\sqrt{\frac{g}{l} \cos \theta_{0}}\left[\begin{array}{cc}
0 & -1  \tag{133}\\
1 & 0
\end{array}\right] \mathbf{z}+\frac{1}{m l^{2} \sqrt{\frac{g}{l} \cos \theta_{0}}}\left[\begin{array}{c}
-m g l \sin \theta_{0} \\
\dot{\theta}_{0}
\end{array}\right]
$$

and above the horizontal where $\theta_{0} \in(\pi / 2,3 \pi / 2)$ we get

$$
\frac{d \mathbf{z}}{d t}=\sqrt{-\frac{g}{l} \cos \theta_{0}}\left[\begin{array}{ll}
0 & 1  \tag{134}\\
1 & 0
\end{array}\right] \mathbf{z}+\frac{1}{m l^{2} \sqrt{-\frac{g}{l} \cos \theta_{0}}}\left[\begin{array}{c}
-m g l \sin \theta_{0} \\
\dot{\theta}_{0}
\end{array}\right]
$$

Since $\left(\begin{array}{cc}0 & -1 \\ 1 & 0\end{array}\right)$ 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 $\left(\begin{array}{ll}0 & 1 \\ 1 & 0\end{array}\right)$ squares to identity. And around $\pm \pi / 2$ something totally different (return to this later). The problem is now reduced to solving a nonhomogeneous first order matrix equation of the form

$$
\begin{equation*}
\mathbf{z}^{\prime}=\Omega \mathbf{z}+\mathbf{b} \tag{135}
\end{equation*}
$$

But we have the good fortune of being able to easily exponentiate and invert this matrix $\Omega$. The homogeneous problem

$$
\begin{equation*}
\mathbf{z}^{\prime}=\Omega \mathbf{z} \tag{136}
\end{equation*}
$$

has the solution

$$
\begin{equation*}
\mathbf{z}_{h}(t)=e^{\Omega t} \mathbf{z}_{t=0} \tag{137}
\end{equation*}
$$

Assuming a specific solution $z=e^{\Omega t} f(t)$ for the non-homogeneous equation, one finds $z=$ $\Omega^{-1}\left(e^{\Omega t}-I\right) \mathbf{b}$. The complete solution with both the homogeneous and non-homogeneous parts is thus

$$
\begin{equation*}
\mathbf{z}(t)=e^{\Omega t} \mathbf{z}_{0}+\Omega^{-1}\left(e^{\Omega t}-I\right) \mathbf{b} \tag{138}
\end{equation*}
$$

Going back to the pendulum problem, lets write

$$
\begin{equation*}
\omega=\sqrt{\frac{g}{l}\left|\cos \theta_{0}\right|} \tag{139}
\end{equation*}
$$

Below the horizontal we have

$$
\begin{align*}
\Omega & =\omega\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]  \tag{140}\\
\Omega^{-1} & =-\frac{1}{\omega}\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]  \tag{141}\\
e^{\Omega t} & =\cos (\omega t)\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+\sin (\omega t)\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right] \tag{142}
\end{align*}
$$

Whereas above the horizontal we have

$$
\begin{align*}
\Omega & =\omega\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right]  \tag{143}\\
\Omega^{-1} & =\frac{1}{\omega}\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right]  \tag{144}\\
e^{\Omega t} & =\cosh (\omega t)\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+\sinh (\omega t)\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right] \tag{145}
\end{align*}
$$

In both cases we have

$$
\begin{align*}
{\left[\begin{array}{c}
p_{\theta} \\
\alpha
\end{array}\right] } & =\left[\begin{array}{cc}
m l^{2} \omega & 0 \\
0 & 1
\end{array}\right] \mathbf{z}  \tag{146}\\
\mathbf{b} & =\frac{1}{\omega}\left[\begin{array}{c}
-\frac{g}{l} \sin \theta_{0} \\
\frac{\theta_{0}}{m l^{2}}
\end{array}\right] \tag{147}
\end{align*}
$$

(where the real angle was $\theta=\theta_{0}+\alpha$ ). Since in this case $\Omega^{-1}$ and $e^{\Omega t}$ commute, we have below the horizontal

$$
\begin{aligned}
\mathbf{z}(t) & =e^{\Omega t}\left(\mathbf{z}_{0}-\Omega^{-1} \mathbf{b}\right)-\Omega^{-1} \mathbf{b} \\
& =\left(\cos (\omega t)\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+\sin (\omega t)\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]\right)\left(\mathbf{z}_{0}+\frac{1}{\omega}\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right] \mathbf{b}\right)+\frac{1}{\omega}\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right] \mathbf{b}
\end{aligned}
$$

Expanding out the $\mathbf{b}$ terms and doing the same for above the horizontal we have respectively (below and above)

$$
\begin{align*}
& \mathbf{z}_{\text {low }}(t)=\left(\cos (\omega t)\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+\sin (\omega t)\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]\right)\left(\mathbf{z}_{0}-\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m l^{2}} \\
\frac{I}{l} \sin \theta_{0}
\end{array}\right]\right)-\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m l^{2}} \\
\frac{g}{T} \sin \theta_{0}
\end{array}\right]  \tag{148}\\
& \mathbf{z}_{\text {high }}(t)=\left(\cosh (\omega t)\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+\sinh (\omega t)\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right]\right)\left(\mathbf{z}_{0}+\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m l^{2}} \\
\frac{g}{T} \sin \theta_{0}
\end{array}\right]\right)+\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m l^{2}} \\
\frac{g}{T} \sin \theta_{0}
\end{array}\right] \tag{149}
\end{align*}
$$

The only thing that is really left is re-insertion of the original momentum and position variables using the inverse relation

$$
\mathbf{z}=\left[\begin{array}{cc}
1 /\left(m l^{2} \omega\right) & 0  \tag{150}\\
0 & 1
\end{array}\right]\left[\begin{array}{c}
p_{\theta} \\
\theta-\theta_{0}
\end{array}\right]
$$

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 $\theta$. Also recall that $p_{\theta}=m l^{2} \dot{\theta}$, so we have

$$
\begin{align*}
\mathbf{z} & =\left[\begin{array}{c}
\dot{\theta} / \omega \\
\theta-\theta_{0}
\end{array}\right]  \tag{151}\\
\mathbf{z}_{0} & =\left[\begin{array}{c}
\dot{\theta}_{t=0} / \omega \\
0
\end{array}\right] \tag{152}
\end{align*}
$$

If this is somehow mystically free of all math mistakes then we have the final solution

$$
\begin{align*}
& {\left[\begin{array}{c}
\dot{\theta}(t) / \omega \\
\theta(t)-\theta_{0}
\end{array}\right]_{\text {low }}=\left(\cos (\omega t)\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+\sin (\omega t)\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]\right)\left(\frac{\dot{\theta}(0)}{\omega}\left[\begin{array}{l}
1 \\
0
\end{array}\right]-\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m m^{2}} \\
\frac{g}{I} \sin \theta_{0}
\end{array}\right]\right)-\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m l^{2}} \\
\frac{g}{I} \sin \theta_{0}
\end{array}\right]}  \tag{154}\\
& {\left[\begin{array}{c}
\dot{\theta}(t) / \omega \\
\theta(t)-\theta_{0}
\end{array}\right]_{\text {high }}=\left(\cosh (\omega t)\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+\sinh (\omega t)\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right]\right)\left(\frac{\dot{\theta}(0)}{\omega}\left[\begin{array}{l}
1 \\
0
\end{array}\right]+\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m l^{2}} \\
\frac{g}{I} \sin \theta_{0}
\end{array}\right]\right)+\frac{1}{\omega^{2}}\left[\begin{array}{c}
\frac{\dot{\theta}_{0}}{m l^{2}} \\
\frac{g}{I} \sin \theta_{0}
\end{array}\right]} \tag{155}
\end{align*}
$$

A qualification is required to call this a solution since it is only a solution is the restricted range where $\theta$ is close enough to $\theta_{0}$ (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's probably also worth considering but I don't feel like trying now.

### 5.8. Spherical pendulum

For the spherical rigid pendulum of length $l$, we have for the distance above the lowest point

$$
\begin{equation*}
h=l(1+\cos \theta) \tag{156}
\end{equation*}
$$

(measuring $\theta$ down from the North pole as conventional). The Lagrangian is therefore

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m l^{2}\left(\dot{\theta}^{2}+\sin ^{2} \theta \dot{\phi}^{2}\right)-m g l(1+\cos \theta) \tag{157}
\end{equation*}
$$

We can drop the constant term, using the simpler Lagrangian

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} m l^{2}\left(\dot{\theta}^{2}+\sin ^{2} \theta \dot{\phi}^{2}\right)-m g l \cos \theta \tag{158}
\end{equation*}
$$

To express the Hamiltonian we need first the conjugate momenta, which are

$$
\begin{align*}
& P_{\theta}=\frac{\partial \mathcal{L}}{\partial \dot{\theta}}=m l^{2} \dot{\theta}  \tag{159}\\
& P_{\phi}=\frac{\partial \mathcal{L}}{\partial \dot{\phi}}=m l^{2} \sin ^{2} \theta \dot{\phi} \tag{160}
\end{align*}
$$

We can now write the Hamiltonian

$$
\begin{equation*}
H=\frac{1}{2 m l^{2}}\left(P_{\theta}^{2}+\frac{1}{\sin ^{2} \theta} P_{\phi}^{2}\right)+m g l \cos \theta \tag{161}
\end{equation*}
$$

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 $\phi$ is itself proportional to $\sin ^{2} \theta$. By expressing the energy in terms of this $P_{\phi}$ 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's avoid that region. For regions where $\sin \theta$ is nicely non-zero, we get for the Hamiltonian equations

$$
\begin{align*}
& \frac{\partial H}{\partial P_{\phi}}=\dot{\phi}=\frac{1}{m l^{2} \sin ^{2} \theta} P_{\phi}  \tag{162}\\
& \frac{\partial H}{\partial P_{\theta}}=\dot{\theta}=\frac{1}{m l^{2}} P_{\theta}  \tag{163}\\
& \frac{\partial H}{\partial \phi}=-\dot{P}_{\phi}=0  \tag{164}\\
& \frac{\partial H}{\partial \theta}=-\dot{P}_{\theta}=-\frac{\cos \theta}{m l^{2} \sin ^{3} \theta} P_{\phi}^{2}-m g l \sin \theta \tag{165}
\end{align*}
$$

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?

Lets write the state space vector for the system as

$$
\mathbf{x}=\left[\begin{array}{c}
P_{\theta}  \tag{166}\\
\theta \\
P_{\phi} \\
\phi
\end{array}\right]
$$

lets also suppose that we are interested in the change to the state vector in the neighborhood of an initial state

$$
\mathbf{x}=\left[\begin{array}{c}
P_{\theta}  \tag{167}\\
\theta \\
P_{\phi} \\
\phi
\end{array}\right]=\left[\begin{array}{c}
P_{\theta} \\
\theta \\
P_{\phi} \\
\phi
\end{array}\right]_{0}+\mathbf{z}
$$

The Hamiltonian equations can then be written

$$
\frac{d \mathbf{z}}{d t}=\left[\begin{array}{c}
\frac{\cos \theta}{m l^{2} \sin ^{3} \theta} P_{\phi}{ }^{2}+m g l \sin \theta  \tag{168}\\
\frac{1}{m l^{2}} P_{\theta} \\
0 \\
\frac{1}{m l^{2} \sin ^{2} \theta} P_{\phi}
\end{array}\right]
$$

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

$$
\begin{equation*}
\frac{d z_{\mu}}{d t}=A_{\mu}\left(x_{v}\right) \approx A_{\mu}\left(\mathbf{x}_{0}\right)+\left.\sum_{\alpha} \frac{\partial A_{\mu}}{\partial x_{\alpha}}\right|_{\mathbf{x}_{0}} z_{\alpha} \tag{169}
\end{equation*}
$$

Our linear approximation is thus

$$
\frac{d \mathbf{z}}{d t} \approx\left[\begin{array}{cccc}
\frac{\cos \theta}{m l^{2} \sin ^{3} \theta} P_{\phi}^{2}+m g l \sin \theta  \tag{170}\\
\frac{1}{m l^{2}} P_{\theta} & 0 \\
\frac{1}{m l^{2} \sin ^{2} \theta} P_{\phi}
\end{array}\right]_{0}+\left[\begin{array}{cccc}
0 & -\frac{P_{\phi}^{2}\left(1+2 \cos ^{2} \theta\right)}{m l^{2} \sin ^{4} \theta}+m g l \cos \theta & \frac{2 \cos \theta}{m l^{2} \sin ^{3} \theta} P_{\phi} & 0 \\
\frac{1}{m l^{2}} & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & \frac{-2 P_{\phi} \cos \theta}{m l^{2} \sin ^{3} \theta} & \frac{1}{m l^{2} \sin ^{2} \theta} & 0
\end{array}\right]_{0} \mathbf{z}
$$

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 $\phi$ leads to a bit of trouble since no explicit $\phi$ 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 $P_{\phi}=$ constant, and to omit this completely from the state vector. Our equations in raw form are now

$$
\begin{align*}
\dot{\theta} & =\frac{1}{m l^{2}} P_{\theta}  \tag{171}\\
\dot{P}_{\theta} & =\frac{\cos \theta}{m l^{2} \sin ^{3} \theta} P_{\phi}^{2}+m g l \sin \theta  \tag{172}\\
\dot{\phi} & =\frac{1}{m l^{2} \sin ^{2} \theta} P_{\phi} \tag{173}
\end{align*}
$$

We can treat the $\phi$ dependence later once we have solved for $\theta$. That equation to later solve is just this last

$$
\begin{equation*}
\dot{\phi}=\frac{1}{m l^{2} \sin ^{2} \theta} P_{\phi} \tag{174}
\end{equation*}
$$

This integrates directly, presuming $\theta=\theta(t)$ is known, and we have

$$
\begin{equation*}
\phi-\phi(0)=\frac{P_{\phi}}{m l^{2}} \int_{0}^{t} \frac{d \tau}{\sin ^{2} \theta(\tau)} \tag{175}
\end{equation*}
$$

Now the state vector and its perturbation can be redefined omitting all but the $\theta$ dependence. Namely

$$
\begin{gather*}
\mathbf{x}=\left[\begin{array}{c}
P_{\theta} \\
\theta
\end{array}\right]  \tag{176}\\
\mathbf{x}=\left[\begin{array}{c}
P_{\theta} \\
\theta
\end{array}\right]=\left[\begin{array}{c}
P_{\theta} \\
\theta
\end{array}\right]_{0}+\mathbf{z} \tag{177}
\end{gather*}
$$

We can now write the remainder of this non-linear system as

$$
\frac{d \mathbf{z}}{d t}=\left[\begin{array}{c}
\frac{\cos \theta}{m l^{2} \sin ^{3} \theta} P_{\phi}^{2}+m g l \sin \theta  \tag{178}\\
\frac{1}{m l^{2}} P_{\theta}
\end{array}\right]
$$

and make the linear approximation around $\mathbf{x}_{0}$ as

$$
\frac{d \mathbf{z}}{d t} \approx\left[\begin{array}{c}
\frac{\cos \theta}{m l^{2} \sin ^{3} \theta} P_{\phi}^{2}+m g l \sin \theta  \tag{179}\\
\frac{1}{m l^{2}} P_{\theta}
\end{array}\right]_{0}+\left[\begin{array}{cc}
0 & -\frac{P_{\phi}^{2}\left(1+2 \cos ^{2} \theta\right)}{m l^{2} \sin ^{4} \theta}+m g l \cos \theta \\
\frac{1}{m l^{2}} & 0
\end{array}\right]_{0} \mathbf{z}
$$

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 $\theta$ dependent motion. That frequency will no longer be $\sqrt{\left|\cos \theta_{0}\right| g / l}$, but also has a $P_{\phi}$ and other more complex trigonometric $\theta$ 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.

### 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 didn't 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 $\theta_{1}$ and length $l_{1}$

$$
\begin{equation*}
z_{1}=\hat{\mathbf{x}} l_{1} e^{i \theta_{1}} \tag{180}
\end{equation*}
$$

If the second mass, dangling from this is at an angle $\theta_{2}$ from the $x$ axis, its position is

$$
\begin{equation*}
z_{2}=z_{1}+\hat{\mathbf{x}} l_{2} e^{i \theta_{2}} \tag{181}
\end{equation*}
$$

We need the velocities, and their magnitudes. For $z_{1}$ this is

$$
\begin{equation*}
\left|\dot{z}_{1}\right|^{2}=l_{1}^{2} \dot{\theta}_{1}^{2} \tag{182}
\end{equation*}
$$

For the second mass

$$
\begin{equation*}
\dot{z}_{2}=\hat{\mathbf{x}} i\left(l_{1} \dot{\theta}_{1} e^{i \theta_{1}}+l_{2} \dot{\theta}_{2} e^{i \theta_{2}}\right) \tag{183}
\end{equation*}
$$

Taking conjugates and multiplying out we have

$$
\begin{equation*}
\left|\dot{z}_{2}\right|^{2}=l_{1}^{2} \dot{\theta}_{1}^{2}+2 l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2} \cos \left(\theta_{1}-\theta_{2}\right)+l_{2}^{2} \dot{\theta}_{2}^{2} \tag{184}
\end{equation*}
$$

That's all that we need for the Kinetic terms in the Lagrangian. Now we need the height for the $m g h$ terms. If we set the reference point at the lowest point for the double pendulum system, the height of the first particle is

$$
\begin{equation*}
h_{1}=l_{2}+l_{1}\left(1-\cos \theta_{1}\right) \tag{185}
\end{equation*}
$$

For the second particle, the distance from the horizontal is

$$
\begin{equation*}
d=l_{1} \cos \theta_{1}+l_{2} \cos \theta_{2} \tag{186}
\end{equation*}
$$

So the total distance from the reference point is

$$
\begin{equation*}
h_{2}=l_{1}\left(1-\cos \theta_{1}\right)+l_{2}\left(1-\cos \theta_{2}\right) \tag{187}
\end{equation*}
$$

We now have the Lagrangian

$$
\begin{align*}
\mathcal{L}^{\prime} & =\frac{1}{2} m_{1} l_{1}^{2} \dot{\theta}_{1}^{2}+\frac{1}{2} m_{2}\left(l_{1}^{2} \dot{\theta}_{1}^{2}+2 l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2} \cos \left(\theta_{1}-\theta_{2}\right)+l_{2}^{2} \dot{\theta}_{2}^{2}\right)  \tag{188}\\
& -m_{1} g\left(l_{2}+l_{1}\left(1-\cos \theta_{1}\right)\right)-m_{2} g\left(l_{1}\left(1-\cos \theta_{1}\right)+l_{2}\left(1-\cos \theta_{2}\right)\right) \tag{189}
\end{align*}
$$

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

$$
\begin{equation*}
\mathcal{L}=\frac{1}{2} M l_{1}^{2} \dot{\theta}_{1}^{2}+\frac{1}{2} m_{2} l_{2}^{2} \dot{\theta}_{2}^{2}+m_{2} l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2} \cos \left(\theta_{1}-\theta_{2}\right)+M l_{1} g \cos \theta_{1}+m_{2} l_{2} g \cos \theta_{2} \tag{190}
\end{equation*}
$$

The conjugate momenta that we need for the Hamiltonian are

$$
\begin{align*}
& P_{\theta_{1}}=M l_{1}^{2} \dot{\theta}_{1}+m_{2} l_{1} l_{2} \dot{\theta}_{2} \cos \left(\theta_{1}-\theta_{2}\right)  \tag{191}\\
& P_{\theta_{2}}=m_{2} l_{2}^{2} \dot{\theta}_{2}+m_{2} l_{1} l_{2} \dot{\theta}_{1} \cos \left(\theta_{1}-\theta_{2}\right) \tag{192}
\end{align*}
$$

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

$$
\left[\begin{array}{c}
P_{\theta_{1}}  \tag{193}\\
P_{\theta_{2}}
\end{array}\right]=\left[\begin{array}{cc}
M l_{1}{ }^{2} & m_{2} l_{1} l_{2} \cos \left(\theta_{1}-\theta_{2}\right) \\
m_{2} l_{1} l_{2} \cos \left(\theta_{1}-\theta_{2}\right) & m_{2} l_{2}{ }^{2}
\end{array}\right]\left[\begin{array}{l}
\dot{\theta}_{1} \\
\dot{\theta}_{2}
\end{array}\right]
$$

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

$$
\boldsymbol{\Theta} \equiv\left[\begin{array}{l}
\theta_{1}  \tag{194}\\
\theta_{2}
\end{array}\right]
$$

Then the relation between the conjugate momenta in vector form

$$
\mathbf{p} \equiv\left[\begin{array}{l}
P_{\theta_{1}}  \tag{195}\\
P_{\theta_{2}}
\end{array}\right]
$$

and the angular velocity vector can be written

$$
\begin{equation*}
\mathbf{p}=\mathcal{I}(\delta) \dot{\Theta} \tag{196}
\end{equation*}
$$

Can we write the Lagrangian in terms of $\dot{\boldsymbol{\Theta}}$ ? The first Kinetic term is easy, just

$$
\frac{1}{2} m_{1} l^{2} \dot{\theta}_{1}^{2}=\frac{1}{2} m_{1} \dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left[\begin{array}{cc}
l_{1}^{2} & 0  \tag{197}\\
0 & 0
\end{array}\right] \dot{\boldsymbol{\Theta}}
$$

For the second mass, going back to 183 , we can write

$$
\dot{z}_{2}=\hat{\mathbf{x}} i\left[\begin{array}{ll}
l_{1} e^{i \theta_{1}} & l_{2} e^{i \theta_{2}} \tag{198}
\end{array}\right] \dot{\boldsymbol{\Theta}}
$$

Writing $\mathbf{r}$ for this $1 x 2$ 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

$$
\begin{equation*}
\left|\dot{z}_{2}\right|^{2}=(\mathbf{r} \dot{\boldsymbol{\Theta}})(\overline{\mathbf{r}} \dot{\boldsymbol{\Theta}})=\dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left(\mathbf{r}^{\mathrm{T}} \overline{\mathbf{r}}\right) \dot{\boldsymbol{\Theta}} \tag{199}
\end{equation*}
$$

The Lagrangian kinetic can all now be grouped into a single quadratic form

$$
\begin{gather*}
Q \equiv m_{1}\left[\begin{array}{c}
l_{1} \\
0
\end{array}\right]\left[\begin{array}{ll}
l_{1} & 0
\end{array}\right]+m_{2}\left[\begin{array}{l}
l_{1} e^{i \theta_{1}} \\
l_{2} e^{i \theta_{2}}
\end{array}\right]\left[\begin{array}{ll}
l_{1} e^{-i \theta_{1}} & l_{2} e^{-i \theta_{2}}
\end{array}\right]  \tag{200}\\
\mathcal{L}=\frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} Q \dot{\boldsymbol{\Theta}}+M l_{1} g \cos \theta_{1}+m_{2} l_{2} g \cos \theta_{2} \tag{201}
\end{gather*}
$$

It is also clear that this generalize easily to multiple connected pendulums, as follows

$$
\begin{align*}
K & =\frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} \sum_{k} m_{k} Q_{k} \dot{\boldsymbol{\Theta}}  \tag{202}\\
Q_{k} & =\left[l_{r} l_{s} e^{i\left(\theta_{r}-\theta_{s}\right)}\right]_{r, s \leq k}  \tag{203}\\
\phi & =-g \sum_{k} l_{k} \cos \theta_{k} \sum_{j=k}^{N} m_{j}  \tag{204}\\
\mathcal{L} & =K-\phi \tag{205}
\end{align*}
$$

In the expression for $Q_{k}$ above, it is implied that the matrix is zero for any indexes $r, s>k$, so it would perhaps be better to write explicitly

$$
\begin{equation*}
Q=\sum_{k} m_{k} Q_{k}=\left[\sum_{j=\max (r, s)}^{N} m_{j} l_{r} l_{s} e^{i\left(\theta_{r}-\theta_{s}\right)}\right]_{r, s} \tag{206}
\end{equation*}
$$

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 201 directly

$$
\begin{aligned}
P_{\theta_{1}} & =\frac{\partial}{\partial \dot{\theta}_{1}} \frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} Q \dot{\boldsymbol{\Theta}} \\
& =\frac{\partial}{\partial \dot{\theta}_{1}} \frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} Q\left[\begin{array}{l}
1 \\
0
\end{array}\right]+\frac{1}{2}\left[\begin{array}{ll}
1 & 0
\end{array}\right] Q \dot{\boldsymbol{\Theta}} \\
& =\left[\begin{array}{ll}
1 & 0
\end{array}\right]\left(\frac{1}{2}\left(Q+Q^{\mathrm{T}}\right)\right) \dot{\boldsymbol{\Theta}}
\end{aligned}
$$

Similarly the $\theta_{2}$ conjugate momentum is

$$
P_{\theta_{2}}=\left[\begin{array}{ll}
0 & 1
\end{array}\right]\left(\frac{1}{2}\left(Q+Q^{\mathrm{T}}\right)\right) \dot{\boldsymbol{\Theta}}
$$

Putting both together, it is straightforward to verify that this recovers 193, which can now be written

$$
\begin{equation*}
\mathbf{p}=\frac{1}{2}\left(Q+Q^{\mathrm{T}}\right) \dot{\boldsymbol{\Theta}}=\mathcal{I} \dot{\boldsymbol{\Theta}} \tag{207}
\end{equation*}
$$

Observing that $\mathcal{I}=\mathcal{I}^{\mathrm{T}}$, and thus $\left(\mathcal{I}^{\mathrm{T}}\right)^{-1}=\mathcal{I}^{-1}$, we now have everything required to express the Hamiltonian in terms of the conjugate momenta

$$
\begin{equation*}
H=\mathbf{p}^{\mathrm{T}}\left(\frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1}\right) \mathbf{p}-M g l_{1} \cos \theta_{1}-m_{2} l_{2} g \cos \theta_{2} \tag{208}
\end{equation*}
$$

This is now in a convenient form to calculate the first set of Hamiltonian equations.

$$
\begin{aligned}
\dot{\theta}_{k} & =\frac{\partial H}{\partial P_{\theta_{k}}} \\
& =\frac{\partial \mathbf{p}^{\mathrm{T}}}{\partial P_{\theta_{k}}} \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \mathbf{p}+\mathbf{p}^{\mathrm{T}} \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \frac{\partial \mathbf{p}^{\mathrm{T}}}{\partial P_{\theta_{k}}} \\
& =\left[\delta_{k j}\right]_{j} \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1} \mathbf{p}+\mathbf{p}^{\mathrm{T}} \frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1}\left[\delta_{i k}\right]_{i} \\
& =\left[\delta_{k j}\right]_{j} \mathcal{I}^{-1} \underbrace{\frac{1}{2}\left(Q+Q^{\mathrm{T}}\right)}_{\mathcal{I}} \mathcal{I}^{-1} \mathbf{p} \\
& =\left[\delta_{k j}\right]_{j} \mathcal{I}^{-1} \mathbf{p}
\end{aligned}
$$

So, when the velocity dependence is a quadratic form as identified in 200 , the first half of the Hamiltonian equations in vector form are just

$$
\dot{\boldsymbol{\Theta}}=\left[\begin{array}{ccc}
\frac{\partial}{\partial P_{\theta_{1}}} & \cdots & \frac{\partial}{\partial P_{\theta_{N}}} \tag{209}
\end{array}\right]^{\mathrm{T}} H=\mathcal{I}^{-1} \mathbf{p}
$$

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

$$
\begin{equation*}
H=\mathbf{p}^{\mathrm{T}}\left(\frac{1}{2} \mathcal{I}^{-1} Q \mathcal{I}^{-1}\right) \mathbf{p}+\phi(\boldsymbol{\Theta}) \tag{210}
\end{equation*}
$$

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

$$
\begin{aligned}
\frac{\partial K}{\partial \theta_{a}} & =\mathbf{p}^{\mathrm{T}}\left(\frac{1}{2} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{a}} Q \mathcal{I}^{-1}\right) \mathbf{p}+\mathbf{p}^{\mathrm{T}}\left(\frac{1}{2} \mathcal{I}^{-1} \frac{\partial Q}{\partial \theta_{a}} \mathcal{I}^{-1}\right) \mathbf{p}+\mathbf{p}^{\mathrm{T}}\left(\frac{1}{2} \mathcal{I}^{-1} Q \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{a}}\right) \mathbf{p} \\
& =\mathbf{p}^{\mathrm{T}} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{a}} \underbrace{\frac{1}{2}\left(Q+Q^{\mathrm{T}}\right)}_{=\mathcal{I}} \mathcal{I}^{-1} \mathbf{p}+\mathbf{p}^{\mathrm{T}}\left(\frac{1}{2} \mathcal{I}^{-1} \frac{\partial Q}{\partial \theta_{a}} \mathcal{I}^{-1}\right) \mathbf{p} \\
& =\mathbf{p}^{\mathrm{T}} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{a}} \mathbf{p}+\mathbf{p}^{\mathrm{T}}\left(\frac{1}{2} \mathcal{I}^{-1} \frac{\partial Q}{\partial \theta_{a}} \mathcal{I}^{-1}\right) \mathbf{p}
\end{aligned}
$$

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

$$
\begin{aligned}
0 & =\frac{\partial I}{\partial \theta_{a}} \\
& =\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{a}} \\
& =\frac{\partial \mathcal{I}}{\partial \theta_{a}} \mathcal{I}^{-1}+\mathcal{I} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{a}}
\end{aligned}
$$

Thus the derivative of the inverse (moment of inertia?) matrix is

$$
\begin{aligned}
\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{a}} & =-\mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{a}} \mathcal{I}^{-1} \\
& =-\mathcal{I}^{-1} \frac{1}{2}\left(\frac{\partial Q}{\partial \theta_{a}}+\frac{\partial Q^{\mathrm{T}}}{\partial \theta_{a}}\right) \mathcal{I}^{-1}
\end{aligned}
$$

This gives us for the Hamiltonian equation

$$
\begin{equation*}
\frac{\partial H}{\partial \theta_{a}}=-\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{a}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}+\frac{\partial \phi}{\partial \theta_{a}} \tag{211}
\end{equation*}
$$

If we introduce a phase space position gradients

$$
\nabla \equiv\left[\begin{array}{lll}
\frac{\partial}{\partial \theta_{1}} & \cdots & \frac{\partial}{\partial \theta_{N}} \tag{212}
\end{array}\right]^{\mathrm{T}}
$$

Then for the second half of the Hamiltonian equations we have the vector form

$$
\begin{equation*}
-\nabla H=\dot{\mathbf{p}}=\left[\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right]_{r}-\nabla \phi \tag{214}
\end{equation*}
$$

The complete set of Hamiltonian equations for 210, in block matrix form, describing all the phase space change of the system is therefore

$$
\frac{d}{d t}\left[\begin{array}{c}
\mathbf{p}  \tag{215}\\
\boldsymbol{\Theta}
\end{array}\right]=\left[\begin{array}{c}
{\left[\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right]_{r}-\nabla \phi} \\
\mathcal{I}^{-1} \mathbf{p}
\end{array}\right]=\left[\begin{array}{c}
{\left[\frac{1}{2} \dot{\boldsymbol{\Theta}}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \dot{\boldsymbol{\Theta}}\right]_{r}-\nabla \phi} \\
\dot{\boldsymbol{\Theta}}
\end{array}\right]
$$

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

$$
\mathbf{z} \equiv\left[\begin{array}{c}
\mathbf{p}  \tag{216}\\
\boldsymbol{\Theta}
\end{array}\right]-\left[\begin{array}{c}
\mathbf{p} \\
\boldsymbol{\Theta}
\end{array}\right]_{t=0}
$$

and the Hamiltonian equations 215 as

$$
\frac{d}{d t}\left[\begin{array}{c}
\mathbf{p}  \tag{217}\\
\mathbf{\Theta}
\end{array}\right]=A(\mathbf{p}, \boldsymbol{\Theta})
$$

Then the linearization, without simplifying or making explicit yet is

$$
\dot{\mathbf{z}} \approx\left[\begin{array}{c}
\left.\left[\begin{array}{c}
\frac{1}{2} \dot{\boldsymbol{\Theta}}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \dot{\boldsymbol{\Theta}} \\
\dot{\boldsymbol{\Theta}}
\end{array}\right]_{r}-\nabla \phi\right]_{t=0}+\left.\left[\begin{array}{llllll}
\frac{\partial A}{\partial P_{\theta_{1}}} & \cdots & \frac{\partial A}{\partial P_{\theta_{N}}} & \frac{\partial A}{\partial \theta_{1}} & \cdots & \frac{\partial A}{\partial \theta_{N}}
\end{array}\right]\right|_{t=0} \mathbf{z} \tag{218}
\end{array}\right.
$$

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's 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 $\Delta$ for the matrix of first order derivatives in our Taylor expansion, as in

$$
\begin{equation*}
f(\mathbf{p}, \boldsymbol{\Theta})=\left.f(\mathbf{p}, \boldsymbol{\Theta})\right|_{0}+\left.\Delta f\right|_{0} \mathbf{z}+\cdots \tag{219}
\end{equation*}
$$

First, lets do the potential gradient.

$$
\Delta(\nabla \phi)=\left[\begin{array}{ll}
0 & {\left[\frac{\partial^{2} \phi}{\partial \theta_{r} \partial \theta_{c}}\right]_{r, c}} \tag{220}
\end{array}\right]
$$

Next in terms of complexity is the first order term of $\dot{\boldsymbol{\Theta}}$, for which we have

$$
\Delta\left(\mathcal{I}^{-1} \mathbf{p}\right)=\left[\left[\mathcal{I}^{-1}\left[\delta_{r c}\right]_{r}\right]_{c} \quad\left[\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}} \mathbf{p}\right]_{c}\right]
$$

The $\delta$ over all rows $r$ and columns $c$ is the identity matrix and we are left with

$$
\begin{equation*}
\Delta\left(\mathcal{I}^{-1} \mathbf{p}\right)=\left[\mathcal{I}^{-1} \quad\left[\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}} \mathbf{p}\right]_{c}\right] \tag{221}
\end{equation*}
$$

Next, consider just the $P_{\theta}$ dependence in the elements of the row vector

$$
\begin{equation*}
\left[\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right]_{r} \tag{222}
\end{equation*}
$$

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

$$
\begin{aligned}
\frac{\partial}{\partial P_{\theta_{c}}}\left(\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right) & =\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1}\left[\delta_{r c}\right]_{r}+\frac{1}{2}\left(\left[\delta_{r c}\right]_{r}\right)^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p} \\
& =\mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial}{\partial \theta_{r}} \frac{1}{2}\left(Q+Q^{T}\right)\right) \mathcal{I}^{-1}\left[\delta_{r c}\right]_{r} \\
& =\mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}} \mathcal{I}^{-1}\left[\delta_{r c}\right]_{r}
\end{aligned}
$$

Since we also have $B^{\prime} B^{-1}+B\left(B^{-1}\right)^{\prime}=0$, for invertible matrixes $B$, this reduces to

$$
\frac{\partial}{\partial P_{\theta_{c}}}\left(\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right)=-\mathbf{p}^{\mathrm{T}} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{r}}\left[\delta_{r c}\right]_{r}
$$

Forming the matrix over all rows $r$, and columns $c$, we get a trailing identity multiplying from the right, and are left with

$$
\begin{equation*}
\left[\frac{\partial}{\partial P_{\theta_{c}}}\left(\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right)\right]_{r, c}=\left[-\mathbf{p}^{\mathrm{T}} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{r}}\right]_{r}=\left[-\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}} \mathbf{p}\right]_{c} \tag{223}
\end{equation*}
$$

Okay, getting closer. The only thing left is to consider the remaining $\theta$ dependence of 222, and now want the theta partials of the scalar matrix elements

$$
\begin{aligned}
\frac{\partial}{\partial \theta_{c}} & \left(\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right) \\
& =\mathbf{p}^{\mathrm{T}}\left(\frac{\partial}{\partial \theta_{c}}\left(\frac{1}{2} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1}\right)\right) \mathbf{p} \\
& =\mathbf{p}^{\mathrm{T}} \frac{1}{2} \mathcal{I}^{-1} \frac{\partial^{2} Q^{\mathrm{T}}}{\partial \theta_{c} \partial \theta_{r}} \mathcal{I}^{-1} \mathbf{p}+\mathbf{p}^{\mathrm{T}} \frac{1}{2}\left(\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1}+\mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}}\right) \mathbf{p} \\
& =\mathbf{p}^{\mathrm{T}} \frac{1}{2} \mathcal{I}^{-1} \frac{\partial^{2} Q^{\mathrm{T}}}{\partial \theta_{c} \partial \theta_{r}} \mathcal{I}^{-1} \mathbf{p}+\mathbf{p}^{\mathrm{T}} \frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}} \frac{\partial \mathcal{I}}{\partial \theta_{r}} \mathcal{I}^{-1} \mathbf{p}
\end{aligned}
$$

There is a slight asymmetry between the first and last terms here that can possibly be eliminated. Using $B^{-1^{\prime}}=-B^{-1} B^{\prime} B^{-1}$, we can factor out the $\mathcal{I}^{-1} \mathbf{p}=\dot{\boldsymbol{\Theta}}$ terms

$$
\frac{\partial}{\partial \theta_{c}}\left(\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right)=\dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left(\frac{1}{2} \frac{\partial^{2} Q^{\mathrm{T}}}{\partial \theta_{c} \partial \theta_{r}}-\frac{\partial \mathcal{I}}{\partial \theta_{c}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}}\right) \dot{\boldsymbol{\Theta}}
$$

Is this any better? Maybe a bit. Since we are forming the matrix over all $r, c$ indexes and can assume mixed partial commutation the transpose can be dropped leaving us with

$$
\begin{equation*}
\left[\frac{\partial}{\partial \theta_{c}}\left(\frac{1}{2} \mathbf{p}^{\mathrm{T}} \mathcal{I}^{-1}\left(\frac{\partial \mathrm{Q}}{\partial \theta_{r}}\right)^{\mathrm{T}} \mathcal{I}^{-1} \mathbf{p}\right)\right]_{r, c}=\left[\dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left(\frac{1}{2} \frac{\partial^{2} Q}{\partial \theta_{c} \partial \theta_{r}}-\frac{\partial \mathcal{I}}{\partial \theta_{c}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}}\right) \dot{\boldsymbol{\Theta}}\right]_{r, c} \tag{224}
\end{equation*}
$$

We can now assemble all these individual derivatives

$$
\dot{\mathbf{z}} \approx\left[\begin{array}{cc}
\left.\left[\frac{1}{2} \dot{\boldsymbol{\Theta}}\left(\frac{\partial \mathrm{Q}}{\partial \theta_{r}}\right)^{\mathrm{T}} \dot{\boldsymbol{\Theta}}\right]_{r}-\nabla \phi\right]_{t=0}+\left[\begin{array}{cc}
-\left[\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}} \mathbf{p}\right]_{c} & {\left[\dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left(\frac{1}{2} \frac{\partial^{2} \mathrm{Q}}{\partial \theta_{c} \partial \theta_{r}}-\frac{\partial \mathcal{I}}{\partial \theta_{c}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}}\right) \dot{\boldsymbol{\Theta}}-\frac{\partial^{2} \phi}{\partial \theta_{r} \partial \theta_{c}}\right]_{r, c}} \\
\mathcal{I}^{-1} & {\left[\frac{\partial\left(\mathcal{I}^{-1}\right)}{\partial \theta_{c}} \mathbf{p}\right]_{c}}
\end{array}\right]_{t=0} \mathbf{z}, ~ \tag{225}
\end{array}\right.
$$

We have both $\partial\left(\mathcal{I}^{-1}\right) / \partial \theta_{k}$ 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

$$
\dot{\mathbf{z}} \approx\left[\begin{array}{cc}
\left.\left[\frac{1}{2} \dot{\boldsymbol{\Theta}}\left(\frac{\partial Q}{\partial \theta_{r}}\right)^{\mathrm{T}} \dot{\boldsymbol{\Theta}}\right]_{r}-\nabla \phi\right]_{t=0}+\left[\begin{array}{cc}
{\left[\mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{c}} \dot{\boldsymbol{\Theta}}\right]_{c}} & {\left[\dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left(\frac{1}{2} \frac{\partial^{2} Q}{\partial \theta_{c} \partial \theta_{r}}-\frac{\partial \mathcal{I}}{\partial \theta_{c}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}}\right) \dot{\boldsymbol{\Theta}}-\frac{\partial^{2} \phi}{\partial \theta_{r} \partial \theta_{c}}\right]_{r, c}} \\
\mathcal{I}^{-1} & -\left[\mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{c}} \dot{\boldsymbol{\Theta}}\right]_{c}
\end{array}\right]_{t=0} \mathbf{z}, ~ \tag{226}
\end{array}\right.
$$

### 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

$$
\begin{align*}
& Q=\mathcal{I}=m l^{2}  \tag{227}\\
& \phi=-m g l \cos \theta \tag{228}
\end{align*}
$$

So all the $\theta$ partials except that of the potential are zero. For the potential we have

$$
\begin{equation*}
-\left.\frac{\partial^{2} \phi}{\partial^{2} \theta}\right|_{0}=-m g l \cos \theta_{0} \tag{229}
\end{equation*}
$$

and for the angular gradient

$$
\begin{equation*}
-\left.\nabla \phi\right|_{0}=\left[-m g l \sin \theta_{0}\right] \tag{230}
\end{equation*}
$$

Putting these all together in this simplest application of 226 we have for the linear approximation of a single point mass pendulum about some point in phase space at time zero:

$$
\dot{\mathbf{z}} \approx\left[\begin{array}{c}
-m g l \sin \theta_{0}  \tag{231}\\
\dot{\theta}_{0}
\end{array}\right]+\left[\begin{array}{cc}
0 & -m g l \cos \theta_{0} \\
\frac{1}{m l^{2}} & 0
\end{array}\right] \mathbf{z}
$$

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'd dropped the $\dot{\theta}_{0}$ term)).

### 5.9.2 Double pendulum explicitly.

For the double pendulum, with $\delta=\theta_{1}-\theta_{2}$, and $M=m_{1}+m_{2}$, we have

$$
Q=\left[\begin{array}{cc}
M l_{1}{ }^{2} & m_{2} l_{2} l_{1} e^{i\left(\theta_{2}-\theta_{1}\right)}  \tag{232}\\
m_{2} l_{1} l_{2} e^{i\left(\theta_{1}-\theta_{2}\right)} & m_{2} l_{2}{ }^{2}
\end{array}\right]=\left[\begin{array}{cc}
M l_{1}{ }^{2} & m_{2} l_{2} l_{1} e^{-i \delta} \\
m_{2} l_{1} l_{2} e^{i \delta} & m_{2} l_{2}{ }^{2}
\end{array}\right]
$$

$$
\begin{aligned}
\frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left(\frac{\partial Q}{\partial \theta_{1}}\right)^{\mathrm{T}} \dot{\boldsymbol{\Theta}} & =\frac{1}{2} m_{2} l_{1} l_{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left[\begin{array}{cc}
0 & -e^{-i \delta} \\
e^{i \delta} & 0
\end{array}\right]^{\mathrm{T}} \dot{\boldsymbol{\Theta}} \\
& =\frac{1}{2} m_{2} l_{1} l_{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}}\left[\begin{array}{c}
e^{i \delta} \dot{\theta}_{2} \\
-e^{-i \delta} \dot{\theta}_{1}
\end{array}\right] \\
& =\frac{1}{2} m_{2} l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2}\left(e^{i \delta}-e^{-i \delta}\right) \\
& =-m_{2} l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2} \sin \delta
\end{aligned}
$$

The $\theta_{2}$ 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

$$
\begin{equation*}
\phi=-M l_{1} g \cos \theta_{1}-m_{2} l_{2} g \cos \theta_{2} \tag{233}
\end{equation*}
$$

So, the gradient is

$$
\nabla \phi=\left[\begin{array}{l}
M l_{1} g \sin \theta_{1}  \tag{234}\\
m_{2} l_{2} g \sin \theta_{2}
\end{array}\right]
$$

Putting things back together we have for the linear approximation of the two pendulum system

$$
\dot{\mathbf{z}}=\left[\begin{array}{c}
m_{2} l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2} \sin \left(\theta_{1}-\theta_{2}\right)\left[\begin{array}{c}
-1 \\
1
\end{array}\right]-g\left[\begin{array}{c}
M l_{1} \sin \theta_{1} \\
m_{2} l_{2} \sin \theta_{2}
\end{array}\right]  \tag{235}\\
\dot{\theta}_{1} \\
\dot{\theta}_{2}
\end{array}\right]_{t=0}+A \mathbf{z}
$$

Where $A$ is still to be determined (from 226).
One of the elements of $A$ are the matrix of potential derivatives. These are

$$
\left[\begin{array}{ll}
\frac{\partial \nabla \phi}{\partial \theta_{1}} & \frac{\partial \nabla \phi}{\partial \theta_{2}}
\end{array}\right]=\left[\begin{array}{cc}
M l_{1} g \cos \theta_{1} & 0  \tag{236}\\
0 & m_{2} l_{2} g \cos \theta_{2}
\end{array}\right]
$$

We also need the inertial matrix and its inverse. These are

$$
\begin{gather*}
\mathcal{I}=\left[\begin{array}{cc}
M l_{1}{ }^{2} & m_{2} l_{2} l_{1} \cos \delta \\
m_{2} l_{1} l_{2} \cos \delta & m_{2} l_{2}{ }^{2}
\end{array}\right]  \tag{237}\\
\mathcal{I}^{-1}=\frac{1}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{cc}
m_{2} l_{2}{ }^{2} & -m_{2} l_{2} l_{1} \cos \delta \\
-m_{2} l_{1} l_{2} \cos \delta & M l_{1}{ }^{2}
\end{array}\right] \tag{238}
\end{gather*}
$$

Since

$$
\frac{\partial Q}{\partial \theta_{1}}=m_{2} l_{1} l_{2} i\left[\begin{array}{cc}
0 & -e^{-i \delta}  \tag{239}\\
e^{i \delta} & 0
\end{array}\right]
$$

We have

$$
\begin{align*}
\frac{\partial}{\partial \theta_{1}} \frac{\partial Q}{\partial \theta_{1}} & =-m_{2} l_{1} l_{2}\left[\begin{array}{cc}
0 & e^{-i \delta} \\
e^{i \delta} & 0
\end{array}\right]  \tag{240}\\
\frac{\partial}{\partial \theta_{2}} \frac{\partial Q}{\partial \theta_{1}} & =m_{2} l_{1} l_{2}\left[\begin{array}{cc}
0 & e^{-i \delta} \\
e^{i \delta} & 0
\end{array}\right]  \tag{241}\\
\frac{\partial}{\partial \theta_{1}} \frac{\partial Q}{\partial \theta_{2}} & =m_{2} l_{1} l_{2}\left[\begin{array}{cc}
0 & e^{-i \delta} \\
e^{i \delta} & 0
\end{array}\right]  \tag{242}\\
\frac{\partial}{\partial \theta_{2}} \frac{\partial Q}{\partial \theta_{2}} & =-m_{2} l_{1} l_{2}\left[\begin{array}{cc}
0 & e^{-i \delta} \\
e^{i \delta} & 0
\end{array}\right] \tag{243}
\end{align*}
$$

and the matrix of derivatives becomes

$$
\frac{1}{2} \dot{\boldsymbol{\Theta}}^{\mathrm{T}} \frac{\partial}{\partial \theta_{c}} \frac{\partial Q}{\partial \theta_{r}} \dot{\boldsymbol{\Theta}}=m_{2} l_{1} l_{2} \dot{\theta_{1}} \dot{\theta_{2}} \cos \left(\theta_{1}-\theta_{2}\right)\left[\begin{array}{cc}
-1 & 1  \tag{244}\\
1 & -1
\end{array}\right]
$$

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

$$
\frac{\partial \mathcal{I}}{\partial \theta_{k}}=-m_{2} l_{1} l_{2}\left(\delta_{k 1}-\delta_{k 2}\right)\left[\begin{array}{cc}
0 & \sin \delta  \tag{245}\\
\sin \delta & 0
\end{array}\right]
$$

Computing the product

$$
\begin{aligned}
\mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{k}} & =\frac{-m_{2} l_{1} l_{2}\left(\delta_{k 1}-\delta_{k 2}\right)}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{cc}
m_{2} l_{2}{ }^{2} & -m_{2} l_{2} l_{1} \cos \delta \\
-m_{2} l_{1} l_{2} \cos \delta & M l_{1}{ }^{2}
\end{array}\right]\left[\begin{array}{cc}
0 & \sin \delta \\
\sin \delta & 0
\end{array}\right] \\
& =\frac{-m_{2} l_{1} l_{2}\left(\delta_{k 1}-\delta_{k 2}\right) \sin \delta}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{cc}
-m_{2} l_{2} l_{1} \cos \delta & m_{2} l_{2}{ }^{2} \\
M l_{1}{ }^{2} & -m_{2} l_{1} l_{2} \cos \delta
\end{array}\right]
\end{aligned}
$$

We want the matrix of $\mathcal{I}^{-1} \partial \mathcal{I} / \partial \theta_{c} \dot{\Theta}$ over columns $c$, and this is

$$
\left[\mathcal{I}^{-1} \partial \mathcal{I} / \partial \theta_{c} \dot{\boldsymbol{\Theta}}\right]_{c}=\frac{m_{2} l_{1} l_{2} \sin \delta}{l_{1}^{2} l_{2}^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{cc}
m_{2} l_{2} l_{1} \cos \delta \dot{\theta}_{1}-m_{2} l_{2}^{2} \dot{\theta}_{2} & -m_{2} l_{2} l_{1} \cos \delta \dot{\theta}_{1}+m_{2} l_{2}^{2} \dot{\theta}_{2}  \tag{246}\\
-M l_{1}{ }^{2} \dot{\theta}_{1}+m_{2} l_{1} l_{2} \cos \delta \dot{\theta}_{2} & M l_{1}^{2} \dot{\theta}_{1}-m_{2} l_{1} l_{2} \cos \delta \dot{\theta}_{2}
\end{array}\right]
$$

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

$$
\begin{aligned}
-\frac{\partial \mathcal{I}}{\partial \theta_{c}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}} & =\frac{-m_{2}{ }^{2} l_{1}{ }^{2} l_{2}{ }^{2}\left(\delta_{c 1}-\delta_{c 2}\right)\left(\delta_{r 1}-\delta_{r 2}\right) \sin ^{2} \delta}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right]\left[\begin{array}{cc}
-m_{2} l_{2} l_{1} \cos \delta & m_{2} l_{2}{ }^{2} \\
M l_{1}{ }^{2} & -m_{2} l_{1} l_{2} \cos \delta
\end{array}\right] \\
& =\frac{-m_{2}{ }^{2} l_{1}{ }^{2} l_{2}{ }^{2}\left(\delta_{c 1}-\delta_{c 2}\right)\left(\delta_{r 1}-\delta_{r 2}\right) \sin ^{2} \delta}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{cc}
M l_{1}{ }^{2} & -m_{2} l_{1} l_{2} \cos \delta \\
-m_{2} l_{2} l_{1} \cos \delta & m_{2} l_{2}{ }^{2}
\end{array}\right]
\end{aligned}
$$

With a sandwich of this between $\dot{\boldsymbol{\Theta}}^{\mathrm{T}}$ and $\dot{\boldsymbol{\Theta}}$ we are almost there

$$
-\dot{\boldsymbol{\Theta}}^{\mathrm{T}} \frac{\partial \mathcal{I}}{\partial \theta_{c}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}} \dot{\boldsymbol{\Theta}}=\frac{-m_{2}^{2} l_{1}^{2} l_{2}^{2}\left(\delta_{c 1}-\delta_{c 2}\right)\left(\delta_{r 1}-\delta_{r 2}\right) \sin ^{2} \delta}{l_{1}{ }^{2} l_{2}^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left(M l_{1}^{2} \dot{\theta}_{1}^{2}-2 m_{2} l_{1} l_{2} \cos \delta \dot{\theta}_{1} \dot{\theta}_{2}++m_{2} l_{2}^{2} \dot{\theta}_{2}^{2}\right)
$$

we have a matrix of these scalars over $r, c$, and that is

$$
\left[-\dot{\boldsymbol{\Theta}}^{\mathrm{T}} \frac{\partial \mathcal{I}}{\partial \theta_{c}} \mathcal{I}^{-1} \frac{\partial \mathcal{I}}{\partial \theta_{r}} \dot{\boldsymbol{\Theta}}\right]_{r c}=\frac{m_{2}{ }^{2} l_{1}{ }^{2} l_{2}^{2} \sin ^{2} \delta}{l_{1}^{2} l_{2}^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left(M l_{1}^{2} \dot{\theta}_{1}^{2}-2 m_{2} l_{1} l_{2} \cos \delta \dot{\theta}_{1} \dot{\theta}_{2}+m_{2} l_{2}{ }^{2} \dot{\theta}_{2}^{2}\right)\left[\begin{array}{cc}
-1 & 1  \tag{247}\\
1 & -1
\end{array}\right]
$$

Putting all the results for the matrix $A$ together is going to make a disgusting mess, so lets summarize in block matrix form

$$
\begin{aligned}
A= & {\left[\begin{array}{cc}
B & C \\
\mathcal{I}^{-1} & -B
\end{array}\right]_{t=0} } \\
B= & \frac{m_{2} l_{1} l_{2} \sin \delta}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{cc}
m_{2} l_{2} l_{1} \cos \delta \dot{\theta}_{1}-m_{2} l_{2}{ }^{2} \dot{\theta}_{2} & -m_{2} l_{2} l_{1} \cos \delta \dot{\theta}_{1}+m_{2} l_{2}{ }^{2} \dot{\theta}_{2} \\
-M l_{1}{ }^{2} \dot{\theta}_{1}+m_{2} l_{1} l_{2} \cos \delta \dot{\theta}_{2} & M l_{1}{ }^{2} \dot{\theta}_{1}-m_{2} l_{1} l_{2} \cos \delta \dot{\theta}_{2}
\end{array}\right] \\
C= & \left.\left(\begin{array}{cc}
m_{2} l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2} \cos \delta+\frac{m_{2}{ }^{2} l_{1}{ }^{2} l_{2}{ }^{2} \sin ^{2} \delta}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left(M l_{1}{ }^{2} \dot{\theta}_{1}^{2}-2 m_{2} l_{1} l_{2} \cos \delta \dot{\theta}_{1} \dot{\theta}_{2}+m_{2} l_{2}{ }^{2} \dot{\theta}_{2}^{2}\right)
\end{array}\right]\right)\left[\begin{array}{cc}
-1 & 1 \\
1 & -1
\end{array}\right] \\
& +\left[\begin{array}{cc}
M l_{1} g \cos \theta_{1} & 0 \\
0 & m_{2} l_{2} g \cos \theta_{2}
\end{array}\right] \\
\mathcal{I}^{-1}= & \frac{1}{l_{1}{ }^{2} l_{2}{ }^{2} m_{2}\left(M-m_{2} \cos ^{2} \delta\right)}\left[\begin{array}{cc}
m_{2} l_{2}{ }^{2} & -m_{2} l_{2} l_{1} \cos \delta \\
-m_{2} l_{1} l_{2} \cos \delta & M l_{1}{ }^{2}
\end{array}\right] \\
b= & {\left[\begin{array}{cc}
m_{2} l_{1} l_{2} \dot{\theta}_{1} \dot{\theta}_{2} \sin \left(\theta_{1}-\theta_{2}\right)\left[\begin{array}{c}
-1 \\
1
\end{array}\right]-g\left[\begin{array}{cc}
M l_{1} \sin \theta_{1} \\
m_{2} l_{2} \sin \theta_{2}
\end{array}\right] \\
\dot{\theta}_{1} \\
\dot{\theta}_{2}
\end{array}\right.}
\end{aligned}
$$

where these are all related by the first order matrix equation

$$
\begin{equation*}
\frac{d \mathbf{z}}{d t}=\left.\mathbf{b}\right|_{t=0}+\left.A\right|_{t=0} \mathbf{z} \tag{248}
\end{equation*}
$$

Wow, even to just write down the equations required to get a linear approximation of the two pendulum system is horrendously messy, and this isn't 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 didn't turn out to be simple even for just the two pendulum system, then one is probably better off going all the way back to 215 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).

### 5.10. Dangling mass connected by string to another.

TODO.

### 5.11. Non-covariant Lorentz force.

In [3], the Lagrangian for a charged particle is given as (12.9) as

$$
\begin{equation*}
\mathcal{L}=-m c^{2} \sqrt{1-\mathbf{u}^{2} / c^{2}}+\frac{e}{c} \mathbf{u} \cdot \mathbf{A}-e \Phi . \tag{249}
\end{equation*}
$$

Let's 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.

### 5.11.1 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

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot{x}_{i}} & =-\frac{1}{2} m c^{2} \gamma\left(-2 / c^{2}\right) \dot{x}_{i}+\frac{e}{c} A_{i} \\
& =m \gamma \dot{x}_{i}+\frac{e}{c} A_{i} .
\end{aligned}
$$

In vector form the canonical momenta are then

$$
\begin{equation*}
\mathbf{P}=\gamma m \mathbf{u}+\frac{e}{c} \mathbf{A} . \tag{250}
\end{equation*}
$$

### 5.11.2 Euler-Lagrange evaluation.

Completing the Euler-Lagrange equation evaluation is the calculation of

$$
\begin{equation*}
\frac{d \mathbf{P}}{d t}=\nabla \mathcal{L} . \tag{251}
\end{equation*}
$$

On the left hand side we have

$$
\begin{equation*}
\frac{d \mathbf{P}}{d t}=\frac{d(\gamma m \mathbf{u})}{d t}+\frac{e}{c} \frac{d \mathbf{A}}{d t}, \tag{252}
\end{equation*}
$$

and on the right, with implied summation over repeated indexes, we have

$$
\begin{equation*}
\nabla \mathcal{L}=\frac{e}{c} \mathbf{e}_{k}\left(\mathbf{u} \cdot \partial_{k} \mathbf{A}\right)-e \nabla \Phi . \tag{253}
\end{equation*}
$$

Putting things together we have

$$
\begin{aligned}
\frac{d(\gamma m \mathbf{u})}{d t} & =-e\left(\nabla \Phi+\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}+\frac{1}{c}\left(\frac{\partial \mathbf{A}}{\partial x_{a}} \frac{\partial x_{a}}{\partial t}-\mathbf{e}_{k}\left(\mathbf{u} \cdot \partial_{k} \mathbf{A}\right)\right)\right) \\
& =-e\left(\nabla \Phi+\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}+\frac{1}{c} \mathbf{e}_{b} u_{a}\left(\frac{\partial A_{b}}{\partial x_{a}}-\frac{\partial A_{a}}{\partial x_{b}}\right)\right) .
\end{aligned}
$$

With

$$
\begin{equation*}
\mathbf{E}=-\nabla \Phi-\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}, \tag{254}
\end{equation*}
$$

the first two terms are recognizable as the electric field. To put some structure in the remainder start by writing

$$
\begin{equation*}
\frac{\partial A_{b}}{\partial x_{a}}-\frac{\partial A_{a}}{\partial x_{b}}=\epsilon^{f a b}(\boldsymbol{\nabla} \times \mathbf{A})_{f} . \tag{255}
\end{equation*}
$$

The remaining term, with $\mathbf{B}=\boldsymbol{\nabla} \times \mathbf{A}$ is now

$$
\begin{aligned}
-\frac{e}{c} \mathbf{e}_{b} u_{a} \varepsilon^{g a b} B_{g} & =\frac{e}{c} \mathbf{e}_{a} u_{b} \epsilon^{a b g} B_{g} \\
& =\frac{e}{c} \mathbf{u} \times \mathbf{B} .
\end{aligned}
$$

We are left with the momentum portion of the Lorentz force law as expected

$$
\begin{equation*}
\frac{d(\gamma m \mathbf{u})}{d t}=e\left(\mathbf{E}+\frac{1}{c} \mathbf{u} \times \mathbf{B}\right) . \tag{256}
\end{equation*}
$$

Observe that with a small velocity Taylor expansion of the Lagrangian we obtain the approximation

$$
\begin{equation*}
-m c^{2} \sqrt{1-\mathbf{u}^{2} / c^{2}} \approx-m c^{2}\left(1-\frac{1}{2} \mathbf{u}^{2} / c^{2}\right)=\frac{1}{2} m \mathbf{u}^{2} \tag{257}
\end{equation*}
$$

If that is our starting place, we can only obtain the non-relativistic approximation of the momentum change by evaluating the Euler-Lagrange equations

$$
\begin{equation*}
\frac{d(m \mathbf{u})}{d t}=e\left(\mathbf{E}+\frac{1}{c} \mathbf{u} \times \mathbf{B}\right) . \tag{258}
\end{equation*}
$$

That was an exercise previously attempting working the Tong Lagrangian problem set [5].

### 5.11.3 Hamiltonian.

Having confirmed the by old fashioned Euler-Lagrange equation evaluation that our Lagrangian provides the desired equations of motion, let's now try it using the Hamiltonian approach. First we need the Hamiltonian, which is nothing more than

$$
\begin{equation*}
H=\mathbf{P} \cdot \mathbf{u}-\mathcal{L} \tag{259}
\end{equation*}
$$

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 $H=H(\mathbf{u}, \mathbf{P}, \mathbf{x})$.

$$
\begin{equation*}
H=\mathbf{P} \cdot \mathbf{u}+m c^{2} \sqrt{1-\mathbf{u}^{2} / c^{2}}-\frac{e}{c} \mathbf{u} \cdot \mathbf{A}+e \Phi . \tag{260}
\end{equation*}
$$

Or

$$
\begin{equation*}
H=\mathbf{u} \cdot\left(\mathbf{P}-\frac{e}{c} \mathbf{A}\right)+m c^{2} \sqrt{1-\mathbf{u}^{2} / c^{2}}+e \Phi . \tag{261}
\end{equation*}
$$

We can rearrange 250 for $\mathbf{u}$

$$
\begin{equation*}
\mathbf{u}=\frac{1}{m \gamma}\left(\mathbf{P}-\frac{e}{c} \mathbf{A}\right), \tag{262}
\end{equation*}
$$

but $\gamma$ also has a u dependence, so this is not complete. Squaring gets us closer

$$
\begin{equation*}
\mathbf{u}^{2}=\frac{1-\mathbf{u}^{2} / c^{2}}{m^{2}}\left(\mathbf{P}-\frac{e}{c} \mathbf{A}\right)^{2}, \tag{263}
\end{equation*}
$$

and a bit of final rearrangement yields

$$
\begin{equation*}
\mathbf{u}^{2}=\frac{(c \mathbf{P}-e \mathbf{A})^{2}}{m^{2} c^{2}+\left(\mathbf{P}-\frac{e}{c} \mathbf{A}\right)^{2}} . \tag{264}
\end{equation*}
$$

Writing $\mathbf{p}=\mathbf{P}-e \mathbf{A} / c$, we can rearrange and find

$$
\begin{equation*}
\sqrt{1-\mathbf{u}^{2} / c^{2}}=\frac{m c}{\sqrt{m^{2} c^{2}+\mathbf{p}^{2}}} \tag{265}
\end{equation*}
$$

Also, taking roots of 264 we must have the directions of $\mathbf{u}$ and $\left(\mathbf{P}-\frac{e}{c} \mathbf{A}\right)$ differ only by a rotation. From 262 we also know that these are colinear, and therefore have

$$
\begin{equation*}
\mathbf{u}=\frac{c \mathbf{P}-e \mathbf{A}}{\sqrt{m^{2} c^{2}+\left(\mathbf{P}-\frac{e}{c} \mathbf{A}\right)^{2}}} . \tag{266}
\end{equation*}
$$

This and 265 can now be substituted into 261, for

$$
\begin{equation*}
H=\frac{c}{m^{2} c^{2}+\mathbf{p}^{2}}\left(\left(\mathbf{P}-\frac{e}{c} \mathbf{A}\right)^{2}+m^{2} c^{2}\right)+e \Phi . \tag{267}
\end{equation*}
$$

Dividing out the common factors we finally have the Hamiltonian in a tidy form

$$
\begin{equation*}
H=\sqrt{(c \mathbf{P}-e \mathbf{A})^{2}+m^{2} c^{4}}+e \Phi \tag{268}
\end{equation*}
$$

### 5.11.4 Hamiltonian equation evaluation.

Let's now go through the exercise of evaluating the Hamiltonian equations. We want the starting point to be just the energy expression 268, 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

$$
\begin{align*}
& \frac{\partial H}{\partial P_{k}}=u_{k}  \tag{269a}\\
& \frac{\partial H}{\partial x_{k}}=-\dot{P}_{k} \tag{269b}
\end{align*}
$$

how far can we go?
For the particle velocity we have no $\Phi$ dependence and get

$$
\begin{equation*}
u_{k}=\frac{c\left(c P_{k}-e A_{k}\right)}{\sqrt{(c \mathbf{P}-e \mathbf{A})^{2}+m^{2} c^{4}}} \tag{270}
\end{equation*}
$$

This is 266 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

$$
\begin{equation*}
\frac{d \mathbf{P}}{d t}=\frac{e\left(c P_{k}-e A_{k}\right) \boldsymbol{\nabla} A_{k}}{\sqrt{(c \mathbf{P}-e \mathbf{A})^{2}+m^{2} c^{4}}}-e \boldsymbol{\nabla} \Phi . \tag{271}
\end{equation*}
$$

Or

$$
\begin{equation*}
\frac{d \mathbf{P}}{d t}=e\left(\frac{u_{k}}{c} \nabla A_{k}-\nabla \Phi\right) \tag{272}
\end{equation*}
$$

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 $d / d t=\partial / \partial t+\mathbf{u} \cdot \nabla$ (ie. chain rule) yielding

$$
\begin{equation*}
\frac{d}{d t}(\mathbf{P}-e \mathbf{A} / c)=e\left(-\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}-\frac{1}{c}(\mathbf{u} \cdot \boldsymbol{\nabla}) \mathbf{A}+\frac{u_{k}}{c} \boldsymbol{\nabla} A_{k}-\nabla \Phi\right) . \tag{273}
\end{equation*}
$$

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

$$
\begin{equation*}
\frac{d}{d t}(\mathbf{P}-e \mathbf{A} / c)=e\left(\mathbf{E}+\frac{\mathbf{u}}{c} \times \mathbf{B}\right) . \tag{274}
\end{equation*}
$$

The only untidy detail left is that $\mathbf{P}-e \mathbf{A} / c$ doesn't look much like $\gamma m \mathbf{u}$, what we recognize as the relativistically corrected momentum. We ought to have that implied somewhere and 270 looks like the place. That turns out to be the case, and some rearrangement gives us this directly

$$
\begin{equation*}
\mathbf{P}-\frac{e}{c} \mathbf{A}=\frac{m \mathbf{u}}{\sqrt{1-\mathbf{u}^{2} / c^{2}}} \tag{275}
\end{equation*}
$$

This completes the exercise, and we've 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.

### 5.12. Covariant force free case.

## TODO.

### 5.13. Covariant Lorentz force.

TODO.

## References

[1] Peeter Joot. Solutions to David Tong's mf1 Lagrangian problems [online]. Available from: http://sites.google.com/site/peeterjoot/geometric-algebra/tong_mf1.pdf.
[2] C. Doran and A.N. Lasenby. Geometric algebra for physicists. Cambridge University Press New York, Cambridge, UK, 1st edition, 2003.
[3] JD Jackson. Classical Electrodynamics Wiley. 2nd edition, 1975.
[4] H. Goldstein. Classical mechanics. Cambridge: Addison-Wesley Press, Inc, 1st edition, 1951.
[5] Dr. David Tong. Classical Mechanics Lagrangian Problem Set 1. [online]. Available from: http://www.damtp.cam.ac.uk/user/tong/dynamics/mf1.pdf.

