Reduction by Symmetry in Lagrangian Mechanics *†

Dustin Ragan

April 16, 2002 This version: May 1, 2003

Contents

*Draft for Evaluation by Thesis Committee
†`allthesis.tex` here

A departmental honors thesis submitted to the

Department of Mathematics at Trinity University

in partial fulfillment of the requirements for Graduation with departmental honors

Jeff Lawson, Thesis Advisor	Julio Hasfura, Mathematics Chair

Moya Ball, VPAA

0 Introduction

Physics professors often joke about how they commit egregious crimes against mathematics in their derivations. Naturally, physics students who moonlight as math majors (or vice versa, depending on your persuasion), sometimes wonder about what happens when mathematicians, in all of their rigor, consider the same questions as a physicist does.

Reduction is exploiting the symmetry of a problem to reduce the size of what we are considering. We can often remove variables from the physical system, letting us solve the differential equations more easily. It is quite often taught in classical mechanics, and I do not go beyond that content here. What will be new to a physics student who reads this is the portrayal of the mathematical structure behind the reductions. To such a reader I hope to introduce, with a minimum of pain, this geometrical interpretation of the physicists theorems. There are several running examples throughout. These are intended to provide a clear illustration of what the theorems say in a concrete manner.

For the more mathematically minded reader, this is intended to serve as a casual introduction to the terser content in actual texts in the area of mathematical physics. Here and now I concede that I have occassionally taken liberties with rigor in the name of clarity, but I hope that I have given sufficient warning about them that they will not lead astray; rather I hope that these serve to illustrate the intentions of the definitions that I cheated upon so that when you encounter the full definition, in its mighty fierceness, you will be more capable of grasping its true meaning.

1 The Examples

This section presents the various example systems we will be working with throughout the text. One of them is an almost completely trivial example. The others are chosen for the variety of symmetries that they offer.

1.1 Example 1: The Falling Rock

Example 1 consists of a rock falling solely due to gravity. We assume that the distance fallen is much less than the radius of the Earth, and so the acceleration due to gravity is effectively constant. We assume it falls straight down, with no veering from a straight path. Thus, it is a one dimensional problem. Let x be the distance from the ground of the rock. Then, if m is the mass of the rock, we have a potential $U(x) = mgx$.

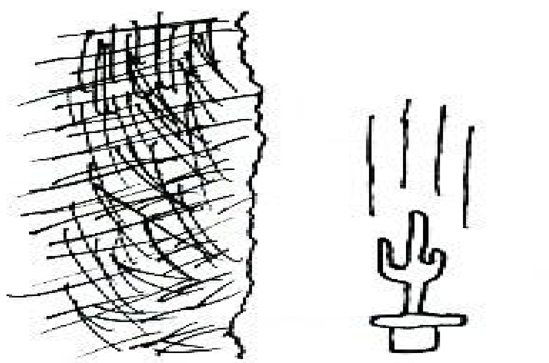

Figure 1.1: Dr. Holder finally tires of the cactus...

1.2 Example 2: Simple Harmonic Oscillator

Bar none (or bar h), one of the most common physical systems is the simple harmonic oscillator. The first example is a spring, with a restoring force proportional to (and opposite in direction) its stretch. This behavior is in fact what happens to most systems when they are near a minimum of their potential. This potential takes the form of $U(x) = \frac{1}{2}kx^2$ near the minimum.

1.3 Example 3: Elroy's Beanie

Elroy's beanie actually refers to both Elroy and his beanie. We assume that Elroy is floating in the middle of space, free from gravitional effects (and friction, and anything else that would complicate the problem). This problem is not interesting because of its potential, but because of its configuration space–how the system's motion is restricted. Elroy and his beanie move independently, each spinning around. Because of this, it is really a system of two rigid bodies spinning around the same axis.

Figure 1.2: Elroy practices his ballet in space, spinning independently of his propeller

1.4 Example 4: Spherical Pendulums

There are actually two problems we will consider involving spherical pendulums. One is a pendulum by itself, being supported and under the influence of gravity. The other example is one spherical pendulum, with another one attached to the bottom of it (see picture). In either example, the pendulums are free to rotate in any direction (like a socket joint).

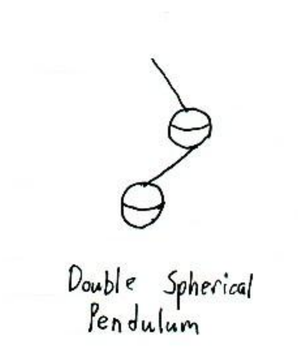

Figure 1.3: A double-spherical pendulum

2 Mathematical Details...

This chapter provides an introduction to many of the miscellaneous mathematical terms used. We will also show how they correspond with the four aforementioned examples.

2.1 Specific Sets

Most readers will readily recognize $\mathbb{R}^n$ as the set of n-tuples of real numbers. Less widely recognized is the set S^1, which is a circle (of any radius) and S^2, which is a sphere. The set $SO(\mathbb{R}, 3)$ is the set of 3×3 real matrices that have determinant $+1$, all of the columns are orthogonal, and all of the the columns have length 1. The set $GL(\mathbb{R}, n)$ is the set of $n \times n$ invertible real matrices.

2.2 Configuration and Phase Spaces

We are ultimately concerned with describing physical systems. We normally describe the system in terms of its *configuration space* and *phase space*. The former is the set of possible physical states that the system can be in; the latter is the configuration space and its time derivatives. This is not a specific coordinate system, just the set of points. See [2]

Example
Most everyday (unconstrained) physical objects can be described as having configuration spaces of $\mathbb{R}^3$ and phase spaces of $\mathbb{R}^3 \times \mathbb{R}^3$.

Example The Falling Rock
The falling rock has a configuration space of $\mathbb{R}$ and a phase space of $\mathbb{R} \times \mathbb{R}$. Its physical position is just its height, its complete description requires specifying both its height and momentum.

Example Simple Harmonic Oscillator
The SHO has the same configuration space as the falling rock. The only difference is the potential.

Example Elroy's Beanie
Since both the beanie and Elroy are free to move in a circle, and their movements are independent, the configuration space is $S^1 \times S^1$. Because angular momentum is not itself measured on a circle, the phase space is $S^1 \times \mathbb{R} \times S^1 \times \mathbb{R}$.

Example Spherical Pendulum
The pendulum has a fixed radius, but is otherwise free to move by rotation. Thus its configuration space is S^2.

2.3 Tangent Spaces

Introductory physics classes always teach that vectors require a magnitude and a direction *and* a point of application. The point of application is then brushed under the rug for the rest of a student's career. However, it is still important. The set of vectors at a point to a manifold is called the *tangent space* at that point. The set of all tangent spaces to a space is called the *tangent bundle* to the space. (See [2].) Tangent spaces can be defined on any manifold. For now think of manifolds as Euclidean space; the definition will be broadened later. More formally,

Definition Let Q be a manifold. For any point $x \in Q$, the tangent space at x, T_xQ, is the collection of all vectors applied at x. The tangent bundle of Q, TQ, is $\bigcup_{x \in Q} T_xQ$.

Example The tangent space of $\mathbb{R}^n$ at any point is $\mathbb{R}^n$. The tangent bundle is $\mathbb{R}^n \times \mathbb{R}^n$.

Since at any point in $\mathbb{R}^n$ we can have a vector going in any direction of any length, $T_x\mathbb{R}^n = \mathbb{R}^n$. Since we have one tangent space for every point of $\mathbb{R}^n$, $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$.

Example The tangent space at any point of S^1 is $\mathbb{R}$. The tangent bundle is $S^1 \times \mathbb{R}$.

At any point on the circle, the tangent vector is a straight line of arbitrary magnitude.

Notice that anytime we calculate the velocity of a point traveling on a path, since the velocity is the temporal derivative of the path, we get a vector tangent to the path (the derivative may in fact be defined in terms of the tangent vector). Tangent spaces are important for this reason—the phase space is the tangent space adjoined to the configuration space.

2.4 Dual Spaces

Definition The dual space of a vector space V, V^*, is the set of linear transformations of the elements of V into $\mathbb{R}$.

Example $\mathbb{R}^{n*} = \mathbb{R}^{1 \times n}$

In $\mathbb{R}^n$ the dual space is obviously the set of $1 \times n$ matrices.

While the tangent space corresponds to temporal derivatives, the dual space corresponds to the space of spatial derivatives. In other words, this is the set of derivatives of curves defined on the space, while the tangent space is derivatives of curves [1] moving in time along the surface. One may think of this as the space of possible gradients of functions on the manifold.

While often not introduced as such, duality is a common notion in mechanics. When considering a scalar quantity that is the dot product of two vectors, for example, we can often think of one of the two vectors as being in the dual space of the other. For example, work is the dot product of force and position; we may think of force as being contained within the dual space of the configuration space. Notice that if the force is derivable from a potential, then the force is the derivative of a function of the space.

2.5 Groups and Group Actions

Groups are one of the basic extensions of a set. They are not just a set but a set with an operation defined on it. Examples are legion; the real numbers (and rationals) under addition (and multiplication, if you remove 0), the integers (under addition), matrix spaces, and most anything else we commonly perform operations with. For the sake of completeness,

Definition A group is a set G and an operation $+$ such that

1. $(a+b)+c = a+(b+c)$
2. There exists $0 \in G \ni a+0 = 0+a = a$.
3. For every $a \in G, \exists -a \ni -a+a = a+-a = 0$.

There is a barely finite number of resources on groups and other algebraic topics available, so we will not go into more detail here, except for a few definitions and specific groups that relate directly to our topic. (Any modern algebra book will include this definition.)

There are many modern algebra texts available, however the definition is also present in [2].

Example The *general linear group on* $\mathbb{R}^n$, written $GL(\mathbb{R}, n)$ is the set of invertible $n \times n$ matrices.

Example The *special orthogonal group on* $\mathbb{R}^n$, written $SO(\mathbb{R}, n)$ is the set of $n \times n$ matrices with determinant 1 such that $A^T = A^{-1}$. This group consists of matrices that simply rotate $\mathbb{R}^n$ and do not distort lengths or change from a right to a left handed coordinate system.

One concept that we deal with is that of a group action. Technically, a group action is a set of functions that respects the group operations, that is:

Definition A group action on a set $\mathcal{X}$ is a group of functions $F : \mathcal{X} \to \mathcal{X}$ and a group $(G, \cdot)$ such that

[1] curve here refers to the geometrical idea; more precisely an embedding of a real interval into the manifold

1. f $g, h \in G$ and $f_g, f_h \in F$, then $f_{g+h} = f_g \circ f_h$, for each $g, h \in G$..
2. f_e is the identity function

This definition, taken from [2], obfuscates its content more than some. Several commonly encountered examples will both illuminate the definition and be of interest in their own right later.

Example Consider the group $(\mathbb{R}^n, +)$. Then the *left translation map* $L_a : \mathbb{R}^n \to \mathbb{R}^n$, for each $a \in \mathbb{R}^n$, defined by $L_a(x) = a + x$, is a group action.

Proof Let $a, b \in \mathbb{R}^n$. Then

$$\begin{aligned}\big(L_a \circ L_b\big)(x) &= L_a(b + x)\\ &= a + (b + x)\\ &= (a + b) + x\\ &= L_{a+b}(x).\end{aligned}$$

Consider, for example, the case where $n = 1$. Then we want the composition of L_3 and L_5, which increment a real number by 3 and 5, to increase the number by 8 (which it does). We get an analagous case for all other values of n.

In this example, the distinction between the left and right translation maps (the latter being define analogously) is unimportant, but if we have a non-abelian (an abelian group as a commutative operation) group they are distinct. Also notice that the domain and range of the function need not be the group that indexes the function, as in the next example.

Example Consider the group $\big((SO(\mathbb{R}, 3), \cdot\big)$. Then the *left translation map* L_a : $\mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, defined by $L_A(x) = Ax$ is a group action.
Proof Almost identical to above.

The above group consists of all coordinate changes applied after a linear transformation.

Example The *conjugate map* of $GL(\mathbb{R}, n)$ on $\mathbb{R}^n$ (or $SO(\mathbb{R}, n)$ on $\mathbb{R}^n$), $F_g : \mathbb{R}^n \to \mathbb{R}^n$, defined by $x \mapsto gxg^{-1}$ is a group action.

Proof Let $g, h \in GL(\mathbb{R}, n)$ (or $SO(\mathbb{R}, n)$. Then

$$
\begin{aligned}
\big(F_g \circ F_h\big)(x) &= F_g(hxh^{-1}) \\
&= ghxh^{-1}g^{-1} \\
&= (gh)x(gh)^{-1} \\
&= F_{gh}(x).
\end{aligned}
$$

Notice that we can also define the conjugate map on a matrix space as well, as long as the matrix multiplication makes sense.

3 Manifolds and Lie Groups

A purpose of mathematics is often to bridge the gap between the intuitive real world and a rigorous understanding based upon fundamental principles. For example, calculus arose from the idea that many functions, when examined under a sufficiently strong microscope, are linear. The formal concept of differentiation is a method of quantifying exactly what this means. Likewise, the dreaded ϵ-δ proofs from analysis define what "sufficiently close" means.

3.1 Manifolds

Manifolds are another case of this. Many surfaces, when examined "sufficiently close", resemble some Euclidean n-space. Casually, a manifold is a set of points to which we may attach some local coordinate system at each point. The formal definition is considerably more challenging than the previous naïve description, but its complexity is derived from the requirements of smoothing rough details of a vague statement (much like the complexity of taking the derivative using ϵ and δ comes from a need to avoid pathologies).

Definition 3.1 *A **manifold** is a set of points M such that there exists an open cover $\mathcal{U}$ of M such that for every $U \in \mathcal{U}$ there is a homeomorphism $\phi_U : U \to \mathbb{R}^n$ with the requirement that for any two non-disjoint subsets U and U' the maps $\phi_U \circ \phi_{U'}^{-1}$ and $\phi_{U'} \circ \phi_U^{-1}$ are infinitely differentiable.*

A reader seeking a more formal and complete definition of manifolds is advised to see one of the many texts available on manifold theory. See, for example, [4].For our purposes this simplified definition is sufficient and conveys the essential nature of manifolds: that they are locally isomorphic to Euclidean space and that components near each other are compatible.

Example $\mathbb{R}^n$
Proof
As one may expect, this is trivial. Just use the identity map on the entire set.

Example S^1
Proof
This requires more subtlety. Fix two distinct points θ, θ' in S^1. Then $\{S^1 - \{\theta\}, S^1 - \{\theta'\}\}$ is an open cover of S^1. Let $\phi_\theta : S^1 - \{\theta\} \to \mathbb{R}$ be such that if $x \in S^1 - \{\theta\}$, then $\phi_\theta(x)$ is the angular rotation (in radians) from θ to x. Define $\phi_{\theta'}$ similarly. These two functions are sufficient to show that S^1 is a manifold. Notice that we could not use one function; if we used just one function there would be a point of non-differentiability in the inverse (actually, the inverse would not be defined everywhere because of the fact that a circle wraps back around on itself).

3.2 Lie Groups

Often a mathematical object is nothing more than an amalgamation of two or more previous ones, like a ring is simply a double group. Thus, a ring can be used to recognize the structures of both $(\mathbb{R},+)$ and $(\mathbb{R}^{\times},\times)$ as one object. Also, $\mathbb{R}^n$ has both a manifold and an additive group structure. In this context, we call $\mathbb{R}^n$ a *Lie group.* (Once again, more explanation can be found in [2].)

Definition 3.2 *A **Lie group** is a manifold that has a C^∞ group operation.*

Lie groups are numerous. As expected, $\mathbb{R}^n$ is a Lie group (under the normal vector addition operation). The general linear group $GL(\mathbb{R},n)$ is a Lie group under matrix multiplication, as is any other vector space. We can also write S^1 as a Lie group, but it requires a minor slight of hand: if $\theta_1,\theta_2 \in S^1$, we can define an operation by writing the elements of the group at $e^{i\theta_1}$ and $e^{i\theta_2}$ and defining the group operation as $(e^{i\theta_1},e^{i\theta_2}) \mapsto e^{i(\theta_1+\theta_2)}$. This avoids the problem of circular addition (which must effectively be addition modulo 2π).

Example $\left(\mathbb{R}^n,+\right)$ and S^2 are Lie groups.

Addition is clearly C^∞. For S^2, consider a 2-dimensional sphere embedded in $\mathbb{R}^3$.

The operation on S^2 corresponds to matrix multiplication[2], which is also C^∞. This same argument holds for S^1.

3.3 Lie Algebras

Unlike manifolds and Lie groups, Lie algebras are not easily related to a familiar structure. However, they are intimately related to Lie groups in addition to be interesting structures in their own right. The salient feature is the *Lie bracket*, which is an anticommutative bilinear operation on some vector space. (See [2].)

Definition 3.3 *A **Lie algebra** is a vector space $\mathfrak{g}$ with a bilinear antisymmetric operation $[\cdot,\cdot]$ called a Lie bracket that satisfies the Jacobi identity, that is*

$$\big[[\xi,\eta],\zeta\big] + \big[[\eta,\zeta],\xi\big] + \big[[\zeta,\xi],\eta\big] = 0.$$

for every $\xi,\eta,\zeta \in \mathfrak{g}$.

As a break from our pattern, $\mathbb{R}^n$ is not, in general, a Lie algebra (except under the trivial bracket where every pair of elements maps to 0). However, $\mathbb{R}^3$ is, under the cross product. Also $n \times n$ matrices are a Lie algebra, under the commutator bracket, $[A,B] = AB - BA$.

[2]Any orthonormal 3×3 matrix represents a rotation of a vector in $\mathbb{R}^3$

There is, in fact, a way to derive a Lie algebra from a Lie group (as the similarity of the names would hint at).

Let G be a Lie group. Let $\Gamma = \{\gamma(t) : \gamma(t)$ is a one parameter curve in G with $\gamma(0) = e\}$ (where e here is the identity element of the group G) and define

$$\mathfrak{g} = \{\xi : \xi = \frac{d}{dt}\gamma(t)|_{t=0}, \gamma(t) \in \Gamma\}.$$

Now let

$$[\xi, \eta] = \frac{d}{dt}\frac{d}{ds} g(t)h(s)g(t)^{-1}\Big|_{s,t=0}$$

where $g(0) = h(0) = e, g'(0) = \xi, h'(0) = \eta$. Then $\mathfrak{g}$ is the Lie algebra of G with Lie bracket $[\cdot, \cdot]$.

The above construction takes the derivatives of all curves at 0. As stated earlier, derivatives and tangent vectors are much the same thing. In fact, we have that $\mathfrak{g} = T_eG$.

Example Let $G = S^1$. Then $\mathfrak{g} = \mathbb{R}$.
Proof We can represent any curve in S^1 by $e^{i\theta(t)}$, where $\theta(t)$ is some curve in $\mathbb{R}$. Then

$$\begin{aligned}\mathfrak{g} &= \left\{\eta : \eta = \frac{d}{dt}e^{i\theta(t)}\Big|_{t=0}\right\} \\ &= \left\{\eta : \eta = i\dot{\theta}(t)e^{i\theta(t)}\Big|_{t=0}\right\} \\ &= \left\{\eta : \eta = i\dot{\theta}(0)\right\} \\ &= i\mathbb{R}.\end{aligned}$$

We may identify this set with $\mathbb{R}$. Now, if $g(t) = e^{i\theta(t)}$ generates the element ξ and $h(s) = e^{i\phi(s)}$ produces the element η,

$$\begin{aligned}[\xi, \eta] &= \frac{d}{dt}\frac{d}{ds} g(t)h(s)g(t)^{-1}\Big|_{s,t=0} \\ &= \frac{d}{dt}\frac{d}{ds} e^{i\theta(t)}e^{i\phi(s)}e^{-i\theta(t)}\Big|_{s,t=0} \\ &= 0.\end{aligned}$$

So the Lie bracket is just the trivial one.

4 Lagrangian and Hamiltonian Mechanics

Introductory physics students are often led to believe that Newton's Laws are the end-all of mechanics. Eventually they are introduced to Lagrangian and Hamiltonian mechanics, which are subtler and in some sense more aesthetically pleasing. They rely upon principles such as conservation of energy, a more comfortable (philosophically, at least) axiom than the somewhat arbitrary creation of "force" found in Newton's Laws. For a more complete exposition than can be found here, we direct the reader to [1].

4.1 Generalized Coordinates

Introductory physics courses give much attention to choosing coordinate systems. This exclusively means choosing the origin and orientation of three axes representing an object's position in space, except on the rare occasions when the student must make an (obvious) choice between spherical and Cartesian coordinates.

However, nature is not restricted to units of distance. Consider a simple situation: a rock being dropped off a cliff. Suppose it is dropped straight down so we have a one-dimensional problem. We will consider the system fully described if a set of coordinates can always tell us what the distance from the ground and velocity of the rock are. What are some sets of coordinate we can use?

We can make the obvious choice, position and velocity. We can make a slight variation and choose position and momentum. These would all be considered valid *generalized coordinates*. Notice that for any choice we make, we ultimately need some "positional" quantity and a change.

An advantage of Lagrangian or Hamiltonian mechanics is that they easily work with generalized coordinates, while Newton's laws are often awkward when dealing with coordinates more complex than rectangular positions, velocities, and momenta. We can literally just substitute any generalized positions, position derivatives, and momenta in the Lagrangian and Hamiltonian equations.

We need to use care when we compare momentum and the derivative of a position. They are not always proportional; but they are connected. We will elaborate more when we discuss Hamiltonian mechanics.

4.2 Lagrangian Mechanics

Call the generalized positions $q_i, i = 1, 2, \ldots n$. If the kinetic energy of a system is K and the potential energy is U (remembering that these are implicitly functions of our coordinates, which are in turn implicitly functions of time), we can define the *Lagrangian* as $L = K - U$. This is not the only possible Lagrangian; but for our purposes we will not need a more sophisticated definition.

There is ultimately a variational principle at work here, which we will elaborate on later. From it we can derive the *Euler-Lagrange equations*:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, i = 1, 2, \ldots n.$$

These equations are a means of obtaining the equations of motion easily from the Lagrangian. As the following examples show, this is because the kinetic energy is related to the square of the velocity, thus giving us a differential equation in $\ddot{q}_i$.

Example The Falling Rock

Let the generalized position be the height above sea level x. Then the kinetic energy is $K = \frac{1}{2}m\dot{x}^2$. The potential is $U = mgx$. Thus, the Lagrangian is $L = \frac{1}{2}m\dot{x}^2 - mgx$. Then,

$$\begin{aligned}
0 &= \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} \\
&= \frac{d}{dt}\frac{\partial}{\partial \dot{x}}\big(\frac{1}{2}m\dot{x}^2 - mgx\big) - \frac{\partial}{\partial x}\big(\frac{1}{2}m\dot{x}^2 - mgx\big) \\
&= \frac{d}{dt}\big(m\dot{x}\big) - mg \\
&= m\ddot{x} - mg \\
&\Rightarrow \ddot{x} = g.
\end{aligned}$$

This is exactly what we would get from Newton's laws.

Example The Simple Harmonic Oscillator (Undamped)

Call the stretch of the spring x. Then $K = \frac{1}{2}m\dot{x}^2$ and $U = \frac{1}{2}kx^2$, where k is the spring constant. Then $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$. So,

$$\begin{aligned}
0 &= \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} \\
&= \frac{d}{dt}\frac{\partial}{\partial \dot{x}}\big(\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2\big) - \frac{\partial}{\partial x}\big(\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2\big) \\
&= \frac{d}{dt}\big(m\dot{x}\big) + kx \\
&= m\ddot{x} + kx \\
&\Rightarrow \ddot{x} = -\frac{k}{m}x.
\end{aligned}$$

This is again what we would get from Newton's Laws.

Example Elroy's Beanie

Choosing an arbitrary but fixed coordinate system, let θ_b be the angle of the Beanie's rotation and let θ_p be Elroy's (the Person's) angle. Call the corresponding moments of intertia I_b and I_p. Then, absent a potential, the Lagrangian is just the kinetic energy, so $L = K = \frac{1}{2}I_b\dot{\theta}_b^2 + \frac{1}{2}I_p\dot{\theta}_p^2$. We perform the calculation solely for the beanie, Elroy's equations are similar:

$$\begin{aligned}
0 &= \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}_b} - \frac{\partial L}{\partial \theta_b} \\
&= \frac{d}{dt}\frac{\partial}{\partial \dot{\theta}_b}\left(\frac{1}{2}I_b\dot{\theta}_b^2 + \frac{1}{2}I_p\dot{\theta}_p^2\right) - \frac{\partial}{\partial \theta_b}\left(\frac{1}{2}I_b\dot{\theta}_b^2 + \frac{1}{2}I_p\dot{\theta}_p^2\right) \\
&= I_b\ddot{\theta}_b \qquad\qquad \Rightarrow \ddot{\theta}_b = 0
\end{aligned}$$

We get constant angular velocity, as we expect from the absence of any forces. Notice that $\frac{\partial L}{\partial \theta_b} = 0$. This is the source of the constant velocity. Because that term was 0, the other term must also have a derivative of 0.

Example Single Spherical Pendulum

We treat the single pendulum in terms of spherical coordinates. Specifically, θ is the angle from the positive xz-half plane, and ϕ is the angle with the vertical. This leaves r as the (fixed) radius. Kinetic energy is $K = \frac{1}{2}mr^2\left(\dot{\theta}^2 + \dot{\phi}^2\right)$. The potential is $U = -mgr\cos\phi$. The Lagrangian is $\frac{1}{2}mr^2\left(\dot{\theta}^2 + \dot{\phi}^2\right) + mgr\cos\phi$.

Now, invoking the Euler-Lagrange equations,

$$\begin{aligned}
0 &= \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} \\
&= \frac{d}{dt}\frac{\partial}{\partial \dot{\theta}}\left(\frac{1}{2}mr^2\left(\dot{\theta}^2 + \dot{\phi}^2\right) + mgr\cos\phi\right) - \frac{\partial}{\partial \theta}\left(mr^2\left(\dot{\theta}^2 + \dot{\phi}^2\right) + mgr\cos\phi\right) \\
&= mr^2\ddot{\theta} \\
&\Rightarrow \ddot{\theta} = 0
\end{aligned}$$

and

$$\begin{aligned}
0 &= \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} \\
&= \frac{d}{dt}\frac{\partial}{\partial \dot{\phi}}\left(\frac{1}{2}mr^2\left(\dot{\theta}^2 + \dot{\phi}^2\right) + mgr\cos\phi\right) - \frac{\partial}{\partial \phi}\left(mr^2\left(\dot{\theta}^2 + \dot{\phi}^2\right) + mgr\cos\phi\right) \\
&= mr^2\ddot{\phi} - mgr\dot{\phi}\sin\phi \\
&\Rightarrow r\ddot{\phi} - g\dot{\phi}\sin\phi = 0.
\end{aligned}$$

Again notice that because the potential is independent of one coordinate, we simply get that the velocity is constant in that coordinate. This observation is crucial to Routhian reduction. However, ϕ is still a complicated differential equation–we cannot simplify it by symmetry.

4.3 Derivation of the Euler-Lagrange Equations from Hamilton's Principle

Hamilton's Principle states that, if we write the action as $I = \int_{t_1}^{t_2} L dt$ then the motion of the system is such that I is an extreme value. This is normally written in terms of a variation,

$$\delta \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt = 0.$$

The above equation looks like a derivative. Since we are discussing extrema, the "suspected derivative" is equal to 0, it smells like a derivative. But we are not taking a derivative with respect to anything in particular (unless we want to get into function spaces), so it cannot be a true derivative. Then what is it?

We have a variation, which is analagous to a derivative. But with a derivative, we know what we are varying and we are examining how a function changes with respect to some known parameter. We are looking at a known function and we do not know how the parameters behave, except that they are all functions of time. We are looking at how changes of those parameters affect our Lagrangian.

This looks like the chain rule from calculus. We have a function L, which is expressed in terms of two sets of functions, $q_i, \dot{q}_i$, which are both parameterized by time (t). We are effectively looking for the functions $q_i, \dot{q}_i$ that produce an extreme value of I.

Theorem 4.1 *Hamilton's Principle implies the Euler-Lagrange equations*
Proof

Assume Hamilton's Principle. Then let $q_i, \dot{q}_i$ be the correct motion of the system. Let $\eta(t)$ be a $\mathcal{C}^2$ function such that $\eta(t_1) = \eta(t_2) = 0$, for every i. Then define the curve $q_i(t, \alpha) = q_i(t) + \alpha\eta(t)$. (Notice that we assume that α is constant with respect to time.)

$$\begin{aligned}
\delta I &= \frac{\partial I}{\partial \alpha} d\alpha \\
&= \int_{t_1}^{t_2} \frac{\partial L(q_i(t,\alpha), \dot{q}_i(t,\alpha), t)}{\partial \alpha} d\alpha dt \\
&= \int_{t_1}^{t_2} \sum_i \left(\frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} d\alpha + \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \alpha} d\alpha \right) dt.
\end{aligned}$$

By integration by parts, using $\frac{\partial}{\partial \alpha} \dot{q}_i = \frac{d}{dt} \frac{\partial q_i}{\partial \alpha}$,

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \frac{\partial q_i}{\partial \alpha} d\alpha dt = \frac{\partial L}{\partial \dot{q}_i} \frac{\partial q_i}{\partial \alpha} d\alpha \bigg|_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \frac{\partial q_i}{\partial \alpha} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) d\alpha dt.$$

Notice that the first term vanishes because the curves pass through the endpoints.

Then, using the above fact and continuing where we left off,

$$\delta I = \int_{t_1}^{t_2} \sum_i \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i dt.$$

Now, since each q_i is independent, the variations are as well. There is a result from the calculus of variations (which follows from the fact that the second derivative of η is continuous; the details of the result are ignored in mechanics texts and we continue the tradition here) that says that the integral can vanish only if each term of the sum does (because the variations are independent). This gives us the Euler-Lagrange equations,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, 2, \ldots n.$$

4.4 Hamiltonian Mechanics

Earlier we stated that, while connected, the time-derivatives of position ($\dot{q}_i$) and the momenta conjugate to the positions (p_i) are distinct. Specifically,

$$p_i = \frac{\partial L}{\partial \dot{q}_i}.$$

This may serve to define momentum.

Changing the variables of Lagrangian mechanics to position-momentum gives us Hamiltonian mechanics. We will almost always be able to say that,

$$H = K + U.$$

Formally, though, we must derive the Hamiltonian from the Lagrangian by the Legendre transformation,

$$H(q, p, t) = \sum_i \dot{q}_i p_i - L(q, \dot{q}, t).$$

Using the relationship between p_i and $\dot{q}_i$ we may eliminate $\dot{q}_i$ from the Hamiltonian, leaving us with just a function of q, p, and t.

If we take a differential and perform the appropriate substitutions (see [1] for more details). From this, analagous to the Euler-Lagrange equations, we get *Hamilton's Equations*:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}$$
$$\dot{p}_i = -\frac{\partial H}{\partial q_i}$$
$$\frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t}.$$

Hamiltonian mechanics is similar to Lagrangian mechanics; they are in fact, equivalent systems. Some problems are easier to approach from one perspective or another. In particular, constant momentum systems are easier to work with in Hamiltonian mechanics. It is actually possible to blend the two systems in one problem; this is known as Routhian reduction and will be discussed later.

Example The Falling Rock

First, using our prior work,

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{x}} \\ &= \frac{\partial}{\partial \dot{x}}\left(\frac{1}{2}m\dot{x}^2 - mgx\right) \\ &= m\dot{x}. \end{aligned}$$

This just the classical momentum. The Hamiltonian is then $H = \frac{p^2}{2m} + mgx$. Thus,

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

These are again what we would get using Newton's equations.

Example Simple Harmonic Oscillator

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{x}} \\ &= \frac{\partial}{\partial \dot{x}}\left(\frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2\right) \\ &= m\dot{x} \end{aligned}$$

This is again the classical momentum. The Hamiltonian is $H = \frac{1}{2m}p^2 + \frac{1}{2}kx^2$. Then

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

This is exactly what we get from Newtonian mechanics.

Example Elroy's Beanie

Starting with the Lagrangian we had previously, (and again only explicitly calculating the beanie's motion):

$$\begin{aligned}
p_b &= \frac{\partial L}{\partial \dot{\theta}_b} \\
&= \frac{\partial}{\partial \dot{\theta}_b}\big(\frac{1}{2}I_b\dot{\theta}_b^2 + \frac{1}{2}I_p\dot{\theta}_p^2\big) \\
&= I_b\dot{\theta}_b.
\end{aligned}$$

and

$$p_p = I_p\dot{\theta}_p.$$

Thus

$$\begin{aligned}
H &= p_b\dot{\theta}_b + p_p\dot{\theta}_p - L \\
&= p_b\dot{\theta}_b + p_p\dot{\theta}_p - \frac{1}{2}I_b\dot{\theta}_b^2 - \frac{1}{2}I_p\dot{\theta}_p^2 \\
&= \frac{p_b^2}{I_b} + \frac{p_p^2}{I_p} - \frac{p_b^2}{2I_b} - \frac{p_p^2}{2I_p} \\
&= \frac{p_b^2}{2I_b} + \frac{p_p^2}{2I_p}.
\end{aligned}$$

$$\begin{aligned}
\dot{\theta}_b &= \frac{\partial H}{\partial p_b} \\
&= \frac{p_b}{I_b} \\
\dot{p}_b &= -\frac{\partial H}{\partial \theta_b} \\
&= 0.
\end{aligned}$$

and $\dot{\theta}_p = \frac{p_p}{I_p}$ and $\dot{p}_p = 0$.

Example Single Spherical Pendulum

Recall that $L = \frac{1}{2}mr^2\left(\dot{\theta}^2 + \dot{\phi}^2\right) + mgr\cos\phi$. Then

$$\begin{aligned}
p_\theta &= \frac{\partial L}{\partial \dot{\theta}} \\
&= mr^2\dot{\theta} \\
p_\phi &= \frac{\partial L}{\partial \dot{\phi}} \\
&= mr^2\dot{\phi}
\end{aligned}$$

So,

$$\begin{aligned}
H &= p_\theta\dot{\theta} + p_\phi\dot{\phi} - L \\
&= \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2} - mgr\cos\phi.
\end{aligned}$$

By Hamilton's equations,

$$
\begin{aligned}
\dot{\theta} &= \frac{\partial H}{\partial p_\theta} \\
&= \frac{p_\theta}{mr^2} \\
\dot{p}_\theta &= -\frac{\partial H}{\partial \theta} \\
&= 0 \\
\dot{\phi} &= \frac{\partial H}{\partial p_\phi} \\
&= \frac{p_\phi}{mr^2} \\
\dot{p}_\phi &= -\frac{\partial H}{\partial \phi} \\
&= -mgr\sin\phi.
\end{aligned}
$$

5 A Few Last Spaces...

Physics is an attempt to describe one large, complicated, and (supposedly) logical system. We see the results and attempt to derive the axioms that produce them. Mathematics is an attempt to describe any conceivable logical system. It begins with axioms and finds their implications. This reversal of roles appears when we shift from the physics "dialect" of mathematics to the more formal mathematical "dialect." In the texts on mathematical physics we usually ignore the origin of the Lagrangian and Hamiltonian. They become another class of functions. They are in a sense more defined by the laws they obey and how they map between spaces than by the motion they represent. In this chapter we attempt to recast Lagrangian mechanics in a more abstract setting, and provide additional mathematical framework.

5.1 Lagrangian Redux

Recall that the tangent bundle, TQ to a manifold Q, is the space of tangent vectors at each point of Q. The tangent vectors are essentially the first derivatives of curves in Q. That is, we may consider $\dot{q}$ to be an element of T_qQ. In other words, a Lagrangian, which is defined in terms of coordinates and their derivatives, is a map from TQ into $\mathbb{R}$. Up until this point, we have not paid much attention to the space on which we define the Lagrangian. Now we will pay more attention to the geometry of those spaces. But first we must proceed through a few more mathematical definitions.

5.2 Aside: Quotient Spaces

The integers $\mathbb{Z}$ and the set $3\mathbb{Z} = \{3z : z \in \mathbb{Z}\}$ are both groups under addition. Now, consider the function $\phi : \mathbb{Z} \to \{0, 1, 2\}$ such that ϕ maps an integer to its remainder when divided by 3.

In abstract math classes we can define a quotient group. Without going into the technical details, the computation would produce $\frac{\mathbb{Z}}{3\mathbb{Z}} = \big\{\{\dots, -6, -3, 0, 3, 6, \dots\}, \{\dots, -5, -2, 1, 4, 7, \dots\}, \{\dots, -4, -1, 2, 5, 8, \dots\}\big\}$.

Obviously, we have a map induced from ϕ on $\frac{\mathbb{Z}}{3\mathbb{Z}}$, say $\hat{\phi}$, because the sets of $\frac{\mathbb{Z}}{3\mathbb{Z}}$ partition $\mathbb{Z}$ into sets that all map to the same point. This is an example of reduction. We have contracted our domain into a smaller set which still contains all of the information needed for our function by removing the cyclicity. This is exactly what we are going to do when we reduce Lagrangians.

5.3 Hamiltonian Vector Fields

We recall from a class in differential equations or vector calculus that a vector field is a mapping from a set of vectors to another set of vectors. These were mostly used for visualizing the dynamics of a given set of differential equations. This connection to rates of change appears again in the form of the **Hamiltonian Vector Field**. Given a Hamiltonian H, the Hamiltonian vector field, X_H, is defined by

$$\begin{aligned} X_H(q^i, p_i) &= \left(\frac{\partial H}{\partial p_i}, -\frac{\partial H}{\partial q^i} \right) \\ &= \left(\dot{q}^i, -\dot{p}_i \right). \end{aligned}$$

5.4 Infinitesimal Generators

Let $\mathfrak{g}$ be a Lie algebra on a manifold M. Then the infinitesimal generator of $\mathfrak{g}$ is the vector field ξ_M is given by

$$\xi_M(z) = \frac{d}{dt}\Big(exp(t\xi) \cdot z \Big)\Big|_{t=0}.$$

Notice that these two vector fields both map into $\mathfrak{g}^*$. The Hamiltonian vector field maps into a row vector of elements of T_0G, which is $\mathfrak{g}^*$. (These last two definitions can be found in [2].)

5.5 Momentum Maps

We have one final definition to make before we may proceed to a meatier area. Suppose that we have a set of Hamiltonian functions on TG, $f(\xi)$, parameterized by the elements of $\mathfrak{g}$ such that

$$X_{f(\xi)} = \xi_G, \forall \xi \in \mathfrak{g}.$$

Then the map $\mathbf{J} : TG \to \mathfrak{g}^*$ implicitly defined by

$$\big\langle \mathbf{J}(z), \xi \big\rangle = f(\xi)(z), \forall \xi \in \mathfrak{g}, z \in TG$$

is called the momentum map.

This is defined, and further explained in [5] and [2].

5.6 What Pray Tell is $\mathfrak{g}^*$?

Now that we have dealt with several consecutive entities, we are left with the question of what $\mathfrak{g}^*$ *is.* We may, if desired, view it as a formal construction, but when possible we should try and discover a physical quantity, which even if $\mathfrak{g}^*$ is not quite isomorphic to, it is at least metaphoric to.

As a first suggestion we feel obliged to comment that we just defined momentum maps, which map into $\mathfrak{g}^*$. This suggests that $\mathfrak{g}^*$ is somehow connected to momentum, or a generalization thereof.

More subtly, recall that by definition $\mathfrak{g}^*$ is a set of linear maps from $\mathfrak{g}$ into $\mathbb{R}$. That is, whenever we multiply an element of $\mathfrak{g}^*$ and an element of $\mathfrak{g}$ we get a scalar.

Now, there is only one type of scalar quantity that is associated with the dynamics of a system, the energy terms. Given that we have already associated $\mathfrak{g}$ (via T_0G) with velocity, and that the product of momentum and velocity has units of energy, the association with momentum seems quite justified.

6 Reductions by Symmetry

6.1 The Reduced Euler-Lagrange Equations

Now that we have many complicated mathematical constructions, we have the question of how to perform a reduction of a symmetry. The first method we will describe is effectively a direct computation on a projected space that is equivalent to our original space, at least as far as the dynamics of the system are concerned. It is taken from a paper by Marsden and Scheurle [2]:

Euler-Poincaré: Let G be a Lie group and let $L : TG \to \mathbb{R}$ be a left invariant Lagrangian (that is the left translation action of the Lie group does not change the Lagrangian). Let $l : \mathfrak{g} \to \mathbb{R}$ be its restriction to the identity. For a curve $g(t) \in G$, let $\xi(t) = g(t)^{-1} \cdot \dot{g}(t)$. Then the following are equivalent:

1. $g(t)$ satisfies the Euler-Lagrange equations for L on G
2. The variational principle

$$\delta \int_a^b L\big(g(t), \dot{g}(t)\big) dt = 0$$

holds, for variations with fixed endpoints

3. The Euler-Poincaré equations hold:

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}_\xi^* \frac{\delta l}{\delta \xi}$$

4. The variational principle

$$\delta \int l\big(\xi(t)\big) dt = 0$$

holds on $\mathfrak{g}$, using variations of the form

$$\delta\xi = \dot{\eta} + [\xi, \eta]$$

where η vanishes at the endpoints.

In some sense, this theorem tells us exactly what we would expect. It says that we can effectively disregard the redundant variables. However, it has a certain weakness, in that it requires that the Lagrangian be invariant under the entire group (which must also be the configuration space).

6.2 Sketch of a Proof of the Euler-Poincaré Theorem

That (1) and (2) are equivalent is one of the most fundamental results of Lagrangian mechanics.

The two variational principles are the same because by construction l takes on the values of L when ξ takes on values corresponding with the g it is derived from.

The variation restriction in (4) effectively constrains us to variations that could be derived from variations on g and that make sense in $\frac{TG}{G}$. This calculation is shown in [2]. (It is recommended that one also reads the proof of the special case; the general proof often does not receive full elucidation because of already existing exposition in the proof of the special case.)

The equivalence of (3) and (4) is really little more than the previous calculation regarding the equivalence of Hamilton's principle and the Euler-Lagrange equations performed in a different space.

Example Elroy's Beanie

In this version of the example, the two angles, θ_b , the angle of the beanie, and θ_p, the angle of Elroy, are completely disjoint. This non-coupling simplifies the problem. Thus the Lagrangian is simply $L = \frac{1}{2} I_b \dot{\theta}_b^2 + \frac{1}{2} I_e \dot{\theta}_p^2$. Clearly the equations of motion are simply that $\dot{\theta}_b$ and $\dot{\theta}_p$ are constant, which we will confirm.

This means that $G = S^1 \times S^1$ and that $g(t) = \begin{pmatrix} e^{i\theta_b(t)} & 0 \\ 0 & e^{i\theta_p(t)} \end{pmatrix}$. Define $\mathbb{I} = \begin{bmatrix} I_b & 0 \\ 0 & I_p \end{bmatrix}$. (Using this representation, we get that $L = \mathrm{tr}\frac{1}{2}\Big(g(\dot{t})^{-1}\mathbb{I}g(\dot{t})\Big)$.) Then we get that

$$\begin{aligned} \xi(g) &= g^{-1}(t)\dot{g}(t) \\ &= \begin{pmatrix} e^{-i\theta_b(t)} & 0 \\ 0 & e^{-i\theta_p(t)} \end{pmatrix} \begin{pmatrix} i\dot{\theta}_b(t)e^{i\theta_b(t)} & 0 \\ 0 & i\dot{\theta}_p(t)e^{i\theta_p(t)} \end{pmatrix} \\ &= \begin{pmatrix} i\dot{\theta}_b(t) & 0 \\ 0 & i\dot{\theta}_p(t) \end{pmatrix}. \end{aligned}$$

We are used to representing states as vectors, but in this case the matrix format allows us to easily do the multiplication (if you change a column vector to a diagonal matrix, componentwise multiplication is easily represented). We are using preexisting notation to best allow us to work in the $S^1 \times S^1$ space, in which addition is best represented using complex multiplication.

Let $l = \big\langle \mathbb{I}\xi(t), \xi(t) \big\rangle$. Notice that this is essentially the same kinetic energy except that we have stripped it of all positional information (notice that no θ's appear in the formula for ξ).

Since we have that g satisfies the equations of motion, we must have that l also does. Taken as a variational principle, this says that

$$\delta \int l\big(\xi(t)\big)dt = 0.$$

So,

$$
\begin{aligned}
0 &= \delta l \\
&= \delta\big(\frac{1}{2}I_b\dot{\theta}_b^2 + \frac{1}{2}I_p\dot{\theta}_p^2\big) \\
&= I_b\dot{\theta}_b\delta\theta_b + I_p\dot{\theta}_p\delta\theta_p.
\end{aligned}
$$

Now, since $\delta\theta_b$ and $\delta\theta_p$ could be anything, and since they are independent, they can never sum to 0. The only way that $\delta l = 0$, then, is if $\dot{theta}_b = \dot{\theta}_e = 0$.

Example Elroy's Beanie (modified)

Now consider what happens when we put a spring on Elroy's beanie so that there is a restoring force whenever Elroy and his beanie are out of alignment. This means that we get a potential (assuming it is an ideal spring) which is $U = \frac{1}{2}k(\theta_b - \theta_p)^2$. If we define $\phi_1 = \frac{1}{\sqrt{2}}\big(\theta_b + \theta_p\big)$ and $\phi_2 = \frac{1}{\sqrt{2}}\big(\theta_b - \theta_p\big)$, we can rewrite the Lagrangian as $\frac{1}{2}I_1'\dot{\phi}_1^2 + \frac{1}{2}I_2'\dot{\phi}_2^2 - \frac{1}{2}k'\phi_2$.

We are stuck (at least as far as this theorem is concerned). The theorem fails because it requires that the group act invariantly on the entire manifold. Even though we change coordinates so that one of the two angles is cyclic, the theorem still fails. What we want is an analogue to the Routhian procedure, where we can at least eliminate ϕ_1.

Aside: In the interests of full disclosure, I concede that the above potential is technically best described as being a function on $\mathbb{R} \times \mathbb{R}$, instead of $S^1 \times S^1$ because this better describes the distinction between a difference of say 78° and 438°. These two difference angles are the same on an S^1 space, but are (correctly) distinguished on $\mathbb{R}$. Assume that we never rotate Elroy too far with respect to his beanie if this bothers you.

6.3 Routhian Reduction

There is also an analogue of the Routhian procedure.

Routhian Reduction Theorem [3]: Let Q be a manifold and let a Lie group G act freely (it has no fixed points) on it. Let J be its momentum map on TQ. Assume that $\mu \in \mathfrak{g}^*$ is a regular value (a point where the derivative is surjective [2]) of J. Consider a simple mechanical system given by a Lagrangian $L : TQ \to \mathbb{R}$ which is G-invariant. Define the Routhian $R^\mu : TQ \to \mathbb{R}$ to be

$$
R^\mu(q, \dot{q}) = \big\langle A(q, \dot{q}), \mu \big\rangle - L(q, \dot{q})
$$

where A is the mechanical connection. (This Routhian is the negative of the one in the cited paper; this is done to make the Routhian match the Hamiltonian).

Suppse that $q(t)$ satisfies the Euler-Lagrange equations for L and lies on the level set $J\big(q(t), \dot{q}(t)\big) = \mu$. Then the induced curve on Q/G_μ satisfies the reduced Lagrangian variational principle dropped to $T(Q/G_\mu)$.

We have not yet discussed the mechanical connection, or in fact any type of connection. Recognize that the quotient group is a type of projection. Then we can reconstruct the original group as $TQ = \frac{TQ}{G} \oplus \mathfrak{g}$. (Imagine that the quotient group removes a copy of G, which is restored by $\mathfrak{g}$). Thus the mechanical connection effectively adds in a copy of what was projected out by the quotient.

The quotient removes an entire coordinate from TQ. This is the same situation as in the classical Routhian, where we take the difference of the Lagrangian and a momentum multiplied by the removed coordinate's velocity.

Example Elroy's Beanie (Modified)

Using the alternate coordinate system described earlier, we have one cyclic variable. This corresponds with the fact that we have an $S^1 \times S^1$ configuration space and a S^1 symmetry, corresponding with the fact that the overall orientation of Elroy and his beanie does not matter.

6.4 Pendulum Madness

We leave you with a question to consider. Foucault's pendulum is a staple of science museums everywhere. It consists of a single spherical pendulum, like the one we have been considering throughout. However, now we include the rotation of the Earth in our calculations. This causes the pendulum to precess. There is clearly a symmetry here. But what form does it take, since S^1 does not break apart S^2 in an obvious manner?

7 Acknowledgements

Thanks go to my family, who still are supportive and remember my existence, in spite of the fact that I was particularly bad about keeping them aware of the state of my life during this project. Thanks to my girlfriend Anne, who provided an uncountably infinite number of "Don't worry, you'll finish this" and finite, but large, number of cups of tea.

I would like to thank Dr. Lawson, for being my thesis advisor, and to apologize for the many stressful moments he incurred during my work with him.

Thanks also to everyone else who assisted me in various ways, both academically and personally, in bringing me to this point where I had the ability to write this.

References

[1] H. Goldstein. *Classical Mechanics*. Addison-Wesley, 1980.

[2] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*. Springer-Verlag, 1999.

[3] J. E. Marsden and J. Scheurle. The reduced euler-lagrange equations. *Fields Institute Communications*, 1, 1993.

[4] J. R. Munkres. *Analysis on Manifolds*. Perseus Publishing, 1991.

[5] S. F. Singer. *Symmetry in Mechanics*. Birkhäuser, 2001.