Extended Essay

What is the mathematics underlying the Lagrangian Formalism and motivating the derivation of the Euler-Lagrange Equations?

A Geometric Approach to the Lagrange Formalism and Derivation of the Euler-Lagrange Equations through tools in manifold analysis

Word Count: 3996

Contents

1 Introduction

Newton's formulation of dynamical laws are used to derive equations of motion (EOMs) of physical systems, with a particular emphasis on the forces that act within it. However, as a consequence of the pedantic requirement for all the forces to be known, the **Newtonian Formalism** faces limitations in complex contexts within and beyond classical mechanics problems, for instance, situations where physical constraint forces can not be explicitly expressed (Thornton and Marion 228).

Derived from **Hamilton's Principle** and formalised in 1834, the **Lagrangian Formalism** provides a more versatile Physics framework (230). Built upon variational notions of kinetic and potential energy, it can circumvent the inability to explicitly express constraint forces.

This paper aims to explore the deeper mathematics underlying the Lagrangian formalism through manifold analysis. We start by constructing the Spherical Pendulum in Section 2, then overview EOMs in the Newtonian framework in Section 3. The geometric breakdown begins in Section 4 and core Lagrangian concepts are defined in Section 5 and 6. With the support of the geometric intuition, Sections 7 and 8 illustrates how **Hamilton's Stationary Principle Action** and **D'Alembert's Virtual Work Principle** motivates the derivation of the **Euler-Lagrange** (EL) equation. Finally, we show Lagrangian procedures recover Newtonian Dynamics in Section 9, ending with further extensions in Section 10.

2 Physical Set Up

We start by defining an arbitrary physical scenario and concrete example to support the definitions and analysis in the rest of the paper.

2.1 General Set Up and Notation

In general definitions and discussions, we assume a scenario of N particles in M dimensional space with K constraint forces applied, unless a particular example is specified. These are indexed with n, m, k respectively. Furthermore, we use both conventions $\dot{x}$ and $\frac{\mathrm{d}x}{\mathrm{d}t}$ to denote derivatives with respect to (wrt) time, and $\vec{x}$ denotes the set of tuples $(x_1, x_2, \dots)$.

2.2 The Spherical Pendulum

The Spherical Pendulum (SP) consists of a ball with mass m attached to a string with length $l = 2$, with its other end fixed in space at the origin $(0, 0, 0)$. Furthermore, the ball can move in 3-dimensional (3-dim) space.

Gravity is assumed to be the sole force acting on the pendulum, inducing a Tension force on the string (Fig 1). There are two additional assumptions:

- $l = 2$ is fixed, hence the motion of the ball will be restricted along the surface of the sphere of radius 2 centered at $(0, 0, 0)$.
- In this ideal case, we assume no Non-conservative forces exist, for instance, the friction due to the string's motion against the fixed point.

Explicit links to SP throughout the paper will be contained within these blue boxes. This helps establish a concrete translation from the geometric principles and the physical interpretation, thus alleviating our grasp of the underlying mathematics.

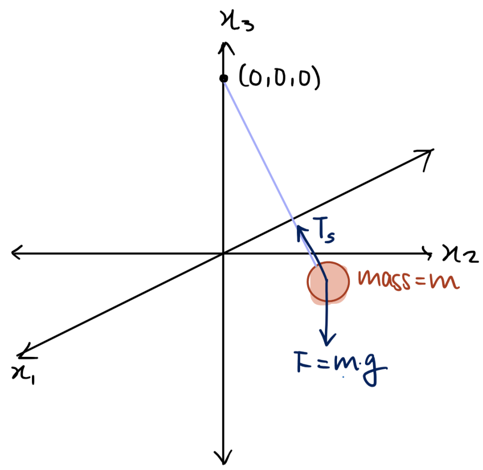

Figure 1: The Spherical Pendulum

3 Newtonian Mechanics

Before exploring the geometry underlying the Lagrangian, we outline a few crucial physical concepts. Kinetic energy, T_n, concerns the relationship between a particle's mass and velocity along all M coordinate axes,

$$T_n = \sum_{m=1}^{M} \frac{1}{2}\mathbf{m}_n \dot{x}_{nm}^2,$$

where n indexes the particle (Tsokos 81). The total kinetic energy T of a system with N particles is the sum of individual kinetic energies,

Definition 3.1: Total Kinetic Energy, T

$$T = \sum_{m,n=1}^{M,N} \frac{1}{2}\mathbf{m}_n \dot{x}_{nm}^2$$

Furthermore, Electric, Gravitational, and Spring potential are examples of Classical potentials (56), and generally represented with,

$$V = V(\vec{x}) = V(x_{11}, \ldots, x_{nm}).$$

. V is defined as the gradient to the Force at a position $\vec{x}$ (Vogtmann et al. 15):

Definition 3.2: Total Potential Energy, V

$$\vec{F}(\vec{x}) = -\nabla V(\vec{x}) = -\frac{\partial V(\vec{x})}{\partial \vec{x}} = -\begin{pmatrix} \frac{\partial V(\vec{x})}{\partial x_{11}} \\ \vdots \\ \frac{\partial V(\vec{x})}{\partial x_{nm}} \end{pmatrix}$$

These two energy concepts are used to define the Lagrangian, as further discussed in Section 5.

3.1 Setting Up Equations of Motion (EOMs)

In this section, we establish the Newtonian framework for setting up EOMs of an arbitrary system. Firstly, we identify all forces $\vec{F}_n$ acting on each particle n and find the net force $\vec{F}$ via,

$$\vec{F}(\vec{x}) = \sum_n \vec{F}_n(\vec{x}_n)$$

Assuming constant mass, Newton's 2nd Law of Motion states that the Force acting on a body is its mass times its acceleration (Thornton and Marion 51), hence,

$$\vec{F}_n(\vec{x}) = \mathbf{m}_n \ddot{\vec{x}}_n \tag{1}$$

Furthermore, substituting Def 3.2, into Eq 1 results in the following general EOMs (Vogtmann et al. 22; Holm et al. 7),

$$-\nabla V(\vec{x}) = m_n \ddot{\vec{x}}_n. \tag{2}$$

This can be equivalently expressed with partial derivatives wrt individual coordinates $x_n m$,

$$-\frac{\partial V(\vec{x})}{\partial x_{nm}} = m_n \ddot{x}_{nm} \tag{3}$$

Systems that can be described by the EOMs in Eq 2 are called **Newtonian Potential Systems** (Vogtmann et al. 53; Holm et al. 7). Towards the end, we show that Lagrangian mechanics recovers these Newtonian results, thus affirming the validity of the mathematical principles underlying the formalism.

4 The Construction of the Manifold

This section covers foundational geometric aspects, including discussions on constraints, coordinates and manifolds.

4.1 Constraints and Generalized Coordinates

The configuration of a system is given by its position coordinates, where there is a range of choice of coordinate systems, such as Cartesian, Cylindrical and Spherical coordinates. Generalized Coordinates (GCs) enables us to deduce a coordinate system that can minimize the information required to describe a system's configuration (Kubiznak 1).

Constraints of a given system are limitations in how it may evolve in time, and is the major deciding factor for picking the set of GCs. Constraints have to be explicitly defined in non-GCs, but are implied within GCs.

According to Thornton and Marion (233), in general, given a K constraints imposed on a system with NM-dim configuration, the minimal degree of freedom (DOF) is,

$$S = NM - K.$$

Subsequently, S is also the number of GCs required to construct the configuration of the system. These GCs are denoted with $q_1, \ldots, q_S$, and conversions between Cartesian coordinates and GCs can be explicitly expressed in terms of one another. For instance, we can construct an equation,

$$x_1 = x_1(q_1, \ldots, q_S),$$

that expresses the value of the x_1 coordinate in terms of $q_1, \ldots, q_S$.

The length constraint mentioned in Section 2 is not implicit in Cartesian coordinates. In order to describe the pendulum's motion, the following equation correspondint to a sphere with radius $l = 2$, whose locus corresponds to the set of possible cartesian states (x_1, x_2, x_3) of SP,

$$x_1^2 + x_2^2 + x_3^2 = 4 \tag{4}$$

must be constructed.

Now consider spherical coordinates ρ, θ, ϕ, corresponding to the radial distance from $(0, 0, 0)$, polar, and azimuthal angle respectively (Fig 2). By setting $\rho = 2$ constant, the SP configuration is only dependent on θ, ϕ, hence the length constraint is implicit in this coordinate system.

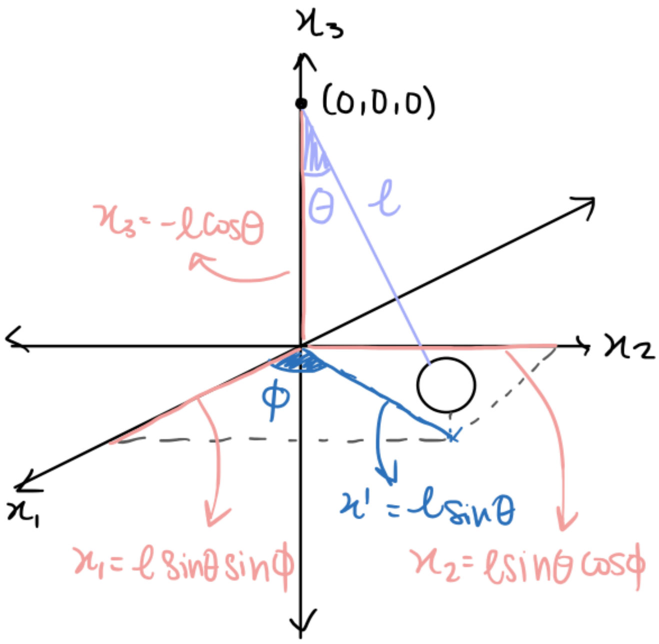

Figure 2: Coordinate components of the position of SP, in terms of θ, ϕ

Therefore, the system's DOF has been reduced from 3 Cartesian coordinates to 2 GCs $q_1, q_2 = \theta, \phi$. Furthermore, as depicted in Fig 2, x_1, x_2, x_3 can be expressed with θ, ϕ,

$$\begin{aligned}
x_1 &= l \sin\theta \cos\phi \\
x_2 &= l \sin\theta \sin\phi \\
x_3 &= -l \cos\theta \\
\dot{x}_1 &= l\dot{\theta} \cos\theta \cos\phi - l\dot{\phi} \sin\theta \sin\phi \\
\dot{x}_2 &= l\dot{\phi} \cos\theta \sin\phi + l\dot{\theta} \sin\theta \cos\phi \\
\dot{x}_3 &= l\dot{\theta} \sin\theta
\end{aligned} \tag{5}$$

Lastly, each k constraint can be defined as a multivariable function C_k, such that (st) any configuration of the system that abides by k evaluates to 0 under C_k (Holm et al. 17; Vogtmann et al. 77),

$$C_k(\vec{x}) = 0.$$

Therefore, a system with K total constraints has functions $C_1, \ldots, C_k$ with the same condition.

We can further define a function F which combines all C_k into a singular function,

Definition 4.1: Constraints Function

$$\begin{aligned} C : \mathbb{R}^{NM} &\to \mathbb{R}^k \\ \vec{x} &\mapsto \begin{pmatrix} C_1(\vec{x}) \\ \vdots \\ C_k(\vec{x}) \end{pmatrix} \end{aligned} \tag{6}$$

As SP only contains a single constraint, it can be constructed based on the explicit expression in Eq 4, where:

$$C(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 4 = 0. \tag{7}$$

Formulating constraints as functions is a crucial precursor to formulating the configuration space as a manifold, as discussed in the following Section.

4.2 Manifolds

The rules of calculus are typically defined for linear spaces such as the reals $\mathbb{R}$ (Ou and Chen). However, when working with configurations often entails curved spaces. Therefore, additional structures and properties must be imposed on such spaces in order to apply calculus to it (Voitsekhovskii).

When located at a point A on the earth's surface, we perceive earth as flat. However, we are only seeing a very small portion of the earth surrounding A, ie. a **local** neighbourhood of A. Instead, when we zoom out and view the earth in its entirety, ie. the **global** viewpoint, we observe the earth is round. Hence the earth is globally "curved", but locally "flat" (Fig 3).

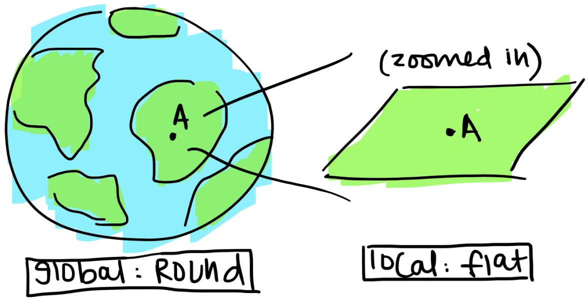

Figure 3: Local and Global properties of the Earth

This is the core idea behind manifolds, which are roughly defined as a topological space that locally resembles Euclidean Space (Rowland). Manifolds consists of a set of points X, and a collection of open subsets $\{O_\alpha\}$, called a **topology**, satisfying a few additional properties.

$\{O_\alpha\}$ is constructed st, for all $a \in X$, there exists at least one subset $O_a \in \{O_\alpha\}$ st. $a \in O_a$ (Kubiznak 5). In other words, $\{O_\alpha\}$ "covers" X. Together, X and $\{O_\alpha\}$ forms a topological space, which enables the construction of continuous functions over X.

Furthermore, for all α, there exists a **homeomorphism**—a continuous map with a one-to-one relation (Holm et al. 45),

$$\Phi_\alpha : O_\alpha \to U_\alpha \tag{8}$$

where $U\alpha \subseteq \mathbb{R}^n$ is an open subset of $\mathbb{R}^n$, for some $n \in \mathbb{N}$. This implies a pairing between each element of O_α and only one element of U_α, which contains no discontinuities or jumps. Φ_α is called a **coordinate map**, and U_α is the **local coordinates or coordinate chart** on O_α (Kubiznak 9; Ou and Chen). Together, the set $(O_\alpha, U_\alpha)_\alpha$ forms an **atlas** on X (Holm et al. 74; Vogtmann et al. 78).

These coordinate charts ultimately characterize the dimensionality of the manifold, where X is an n-dim manifold when its charts are subsets of $\mathbb{R}^n$ (Kubiznak 8).

Recall the length constraint function $C(\vec{x})$ in Eq 7. Consider a possible cartesian configuration of SP, $\vec{x} \in \mathbb{R}^3$, which, by definition, satisfies $C(\vec{x}) = 0$. By collecting all such $\vec{x}$ into a subset $\mathbb{Q} \in \mathbb{R}^3$, every single element of $\mathbb{Q}$ satisfies the constraints of the system.

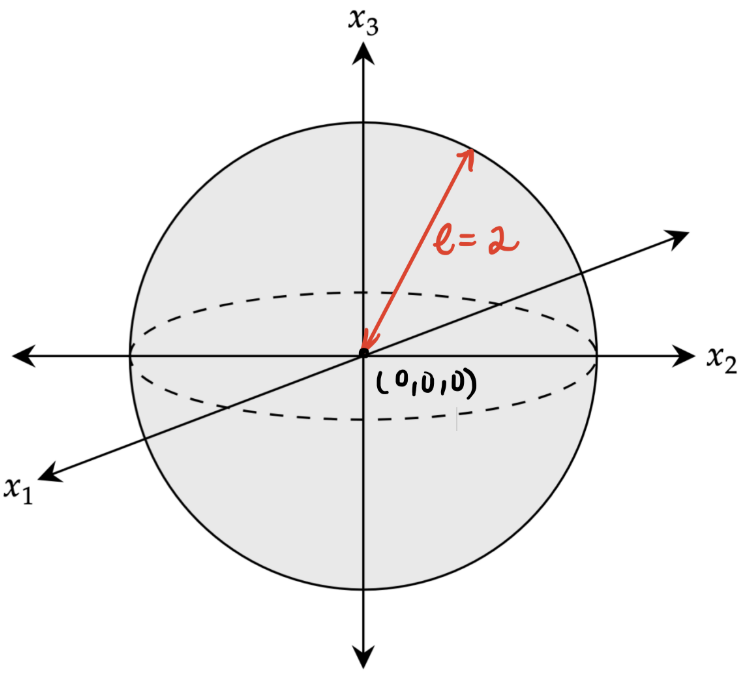

Figure 4: The Configuration Space of SP as a 2-dim submanifold Q

As depicted in Fig 4, Q is the locus of Eq 4. Since any point $(x_1, x_2, x_3) \in Q$ will by definition satisfy the length constraint, Q is the set of points that underlie our Lagrangian manifold.

A topology can be imposed on Q by considering a collection of its open subsets, for instance, the six hemispheres of Q relative to the three axes in Fig 5.

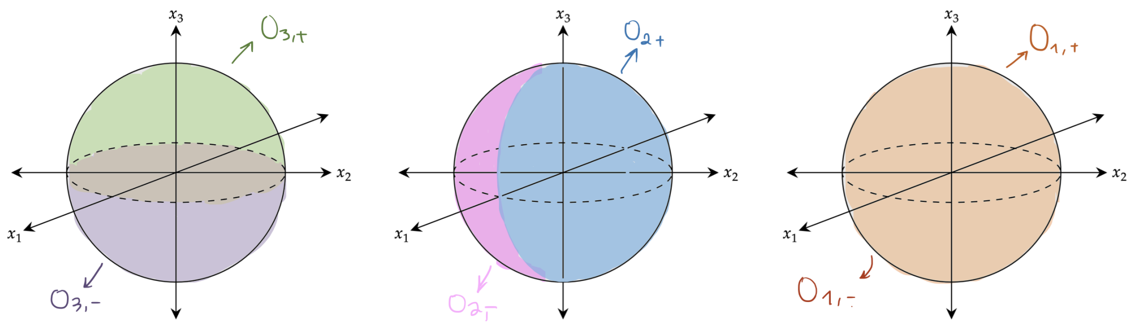

Figure 5: The hemispheres associated with each element of $O_{i,\pm_{i\in\{1,2,3\}}}$

The general construction is,

$$O_{i,\pm} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid \pm x_i > 0 \right\}$$

Furthermore, each open subset can be projected into an open subset of $\mathbb{R}^2$ via the coordinate chart (Grossmann),

$$h_{i,\pm} : \mathbb{R}^3 \to \mathbb{R}^2$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ x_{i-1} \\ x_{i+1} \\ \vdots \end{pmatrix} \tag{9}$$

which effectively "flattens" each hemisphere into a circular disk. This mapping has a one-to-one relation and preserves the closeness of points, hence preserves continuity. For example, Fig 6 depicts the coordinate map $h_{2,+}$, mapping a line section in $O_{2,+}$ onto $\mathbb{R}^2$, where its continuity is preserved under $h_{2,+}$.

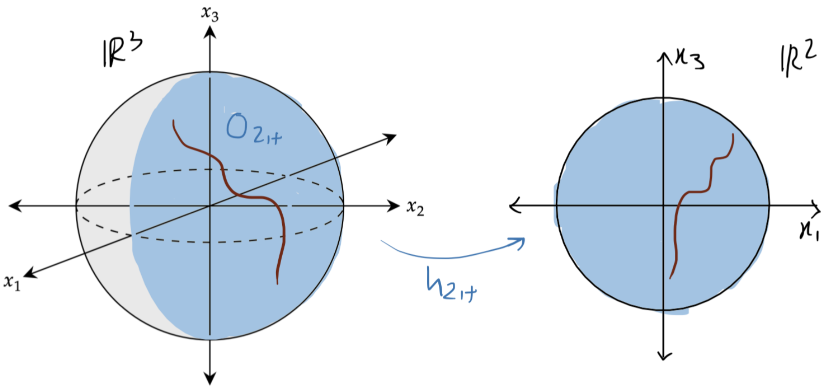

Figure 6: Projection of $O_{2,+}$ onto $\mathbb{R}^2$ via $h_{2,+}$

Together, $(O_{i,\pm}, h_{i,\pm})_{i\in\{1,2,3\}}$ is the atlas on Q and Q is a 2-dim manifold embedded in 3-dim space.

The coordinate charts ensure the non-empty intersections $O_{\alpha\beta} = O_\alpha \cap O_\beta$ is mapped to the subsets $U_\alpha \subseteq \mathbb{R}^n$ and $U_\beta \subseteq \mathbb{R}^n$, under Φ_α and Φ_β respectively (Kubiznak 9). As Φ maps are homeomorphisms, it is possible to construct a continuous inverse map

$$\Phi^{-1} : U \to O$$

which reverses the pairing in the opposite direction. Therefore, we construct the transition map,

$$\tau_{\alpha\beta} : \Phi_\alpha(O_{\alpha\beta}) \xrightarrow{\Phi_\alpha^{-1}} O_{\alpha\beta} \xrightarrow{\Phi_\beta} \Phi_\beta(O_{\alpha\beta})$$

that converts from the local coordinates of O_α to that of O_β for all points in $O_{\alpha\beta}$. In particular, $\tau_{\alpha\beta}$ is a also homeomorphism (Holm et al. 47; Keng).

By making Φ infinitely differentiable, $\tau_{\alpha\beta}$ is also infinitely differentiable. This induces a global differentiable structure on the manifold X, hence forming a **differentiable manifold** (Ou and Chen; Voitsekhovskii).

We can construct the inverse map for $h_{3,+}$ as,

$$h_{3,+}^{-1} : \mathbb{R}^3 \to \mathbb{R}^2$$
$$\begin{pmatrix} x_1 \\ x_2 \\ \sqrt{4-(x_1^2+x_2^2)} \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Since the intersection $O_{32} = O_{3,+} \cap O_{2,+}$ is not empty, the following transition map $\tau_{32} = h_{3,+}^{-1} \circ h_{2,+}$

$$\tau_{32} : h_{3,+}(O_{32}) \xrightarrow{h_{3,+}^{-1}} O_{32} \xrightarrow{h_{2,+}} h_{2,+}(O_{32})$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ \sqrt{4-(x_1^2+x_2^2)} \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \sqrt{4-(x_1^2+x_2^2)} \end{pmatrix}$$

Thus, the continuity of a line section contained in O_{32}, highlighted red in Fig 7, is preserved under τ_{32}.

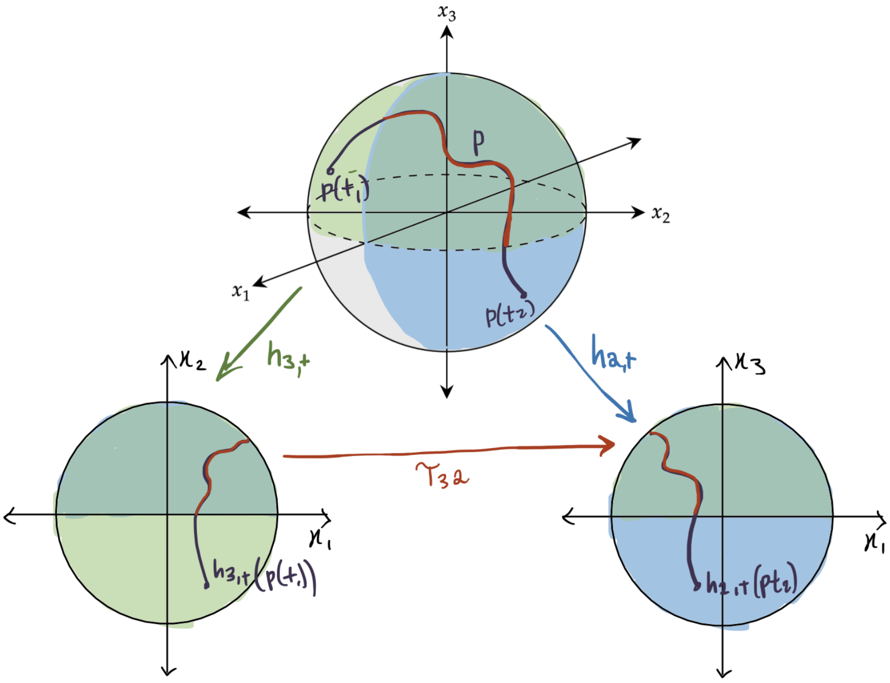

Figure 7: The mapping of a line section in O_{32} under τ_{32}

Differentiable transition maps implies that the computations in one coordinate chart is equally valid in other charts. Hence,it is on these differentiable manifolds that we can apply calculus and analysis to curved surfaces, enabling us to analyze the dynamics of a physical system.

4.3 Function Space and Paths

This section establishes what it means to have paths on a differentiable manifold. Consider the map,

$$P : [t_1, t_2] \subseteq \mathbb{R}^+ \to X \subseteq \mathbb{R}^{NM}$$
$$t \mapsto (r_1(t), \dots, r_{nm}(t)).$$

P parameterizes a path along X starting from the point $P(t_1)$ at the start-time $t = t_1$ and ending at $P(t_2)$ at $t = t_2$ (Holm et al. 15). We assume P is differentiable, hence for all $t \in [t_1, t_2]$, $P(t)$ and $\dot{P}(t)$ is defined.

The set of all such paths $P \subseteq X$ across all possible t_1, t_2, will roughly form a **function space** $X^{\mathbb{R}^+}$, which consists of all functions between $\mathbb{R}^+$ and X. Thus, $X^{\mathbb{R}^+}$ is the collection of paths through time that are consistent with constraints.

Fig 8 depicts some paths contained in $Q^{\mathbb{R}^+}$ for SP.

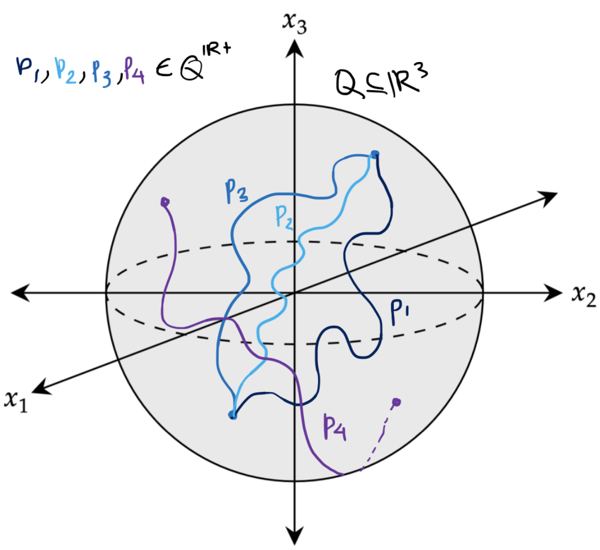

Figure 8: Examples of the elements of the Function Space $Q^{\mathbb{R}^+}$

Paths underlies the EL, and will be used to formulate various notions Section 5 and beyond.

4.4 Tangent Spaces of X

With paths defined, consider the tangential components of X, starting with tangent vectors,

Definition 4.2: Tangent Vectors

Let X be a differentiable manifold and $P : \mathbb{R} \to X$ a smooth path on X. A tangent vector of X with the base-point $P(0) = x$ is the directional derivative $P'(0)$ (Holm et al. 51; Ou and Chen).

By collecting all tangent vectors at a point $x \in X$, we obtain,

Definition 4.3: Tangent Spaces

The tangent space of X at x is composed of tangent vectors at x of all possible smooth maps $P \subseteq X$ (Holm et al. 51; Kubiznak 11). It is denoted T_xX, where

$$TxX = \{(x, x')|x' = P'(0) \text{ for some path } P(t) \subseteq X \text{ and } P(0) = x\} \tag{10}$$

Finally, by collecting the tangent spaces across all $x \in X$, we obtain,

Definition 4.4: Tangent Bundle

The tangent bundle of X is,

$$\begin{aligned} TX &= \bigcup_{x \in X} T_x X \\ &= \{(x, x')|x \in X \wedge P(t) \subseteq X \text{ st } x = P(0) \wedge x' = P'(0)\} \end{aligned} \tag{11}$$

(Vogtmann et al. 81; Kubiznak 11)

In SP, the tangent vectors of a sample of paths through $x \in Q$ and the subsequent tangent space are depicted in Fig 9.

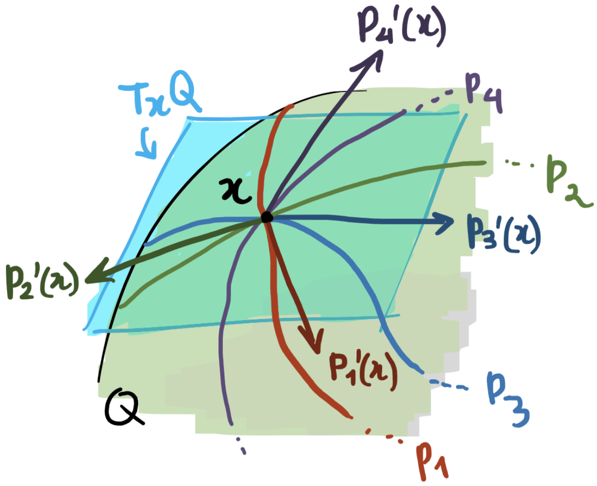

Figure 9: The tangent vectors of paths P_1, P_2, P_3, P_4 and at T_xQ $x \in Q$

We first use the tangent bundle to define the Lagrangian, then explore the physical interpretations of tangent vectors and spaces in Section 7.

5 The Lagrangian

In Classical mechanics, the Lagrangian is defined as total kinetic energy T, less total potential energy V (Vogtmann et al. 53). It is a property that characterizes the configuration of a system.

From the definitions of T and V, in Defs 3.1 and 3.2, we conclude that the Lagrangian inputs the positions and velocities of a system at a point in time, thus

$$\begin{aligned} L(\vec{x}, \dot{\vec{x}}) &= T(\dot{\vec{x}}) - V(\vec{x}) \\ &= \sum_{m,n=1}^{M,N} \frac{1}{2}\mathbf{m}_n {x_{nm}}^2 - V(\vec{x}) \end{aligned}$$

Note that the set of all inputs $\vec{x}, \dot{\vec{x}}$ is precisely the tangent bundle of X in Def 4.4 (Kubiznak 19). Therefore, the Lagrangian is typically defined as a map,

Definition 5.1: Lagrangian for Newtonian Potential Systems

$$\begin{aligned} L :&TX \to \mathbb{R} \\ (\vec{x}, \dot{\vec{x}}) &\mapsto \sum_{m,n=1}^{M,N} \frac{1}{2}\mathbf{m}_n {x_{nm}}^2 - V(\vec{x}) \end{aligned}$$

(Holm et al. 18)

Furthermore, the application of the Lagrangian along a path $P \subseteq X$ underlies EL.

Consider an arbitrary path $P \subseteq Q$ of SP, depicted in Fig 10. By setting the points along P as the x-axis, and their associated Lagrangians on the y-axis, we obtain a graphical representation of the Lagrangian along P.

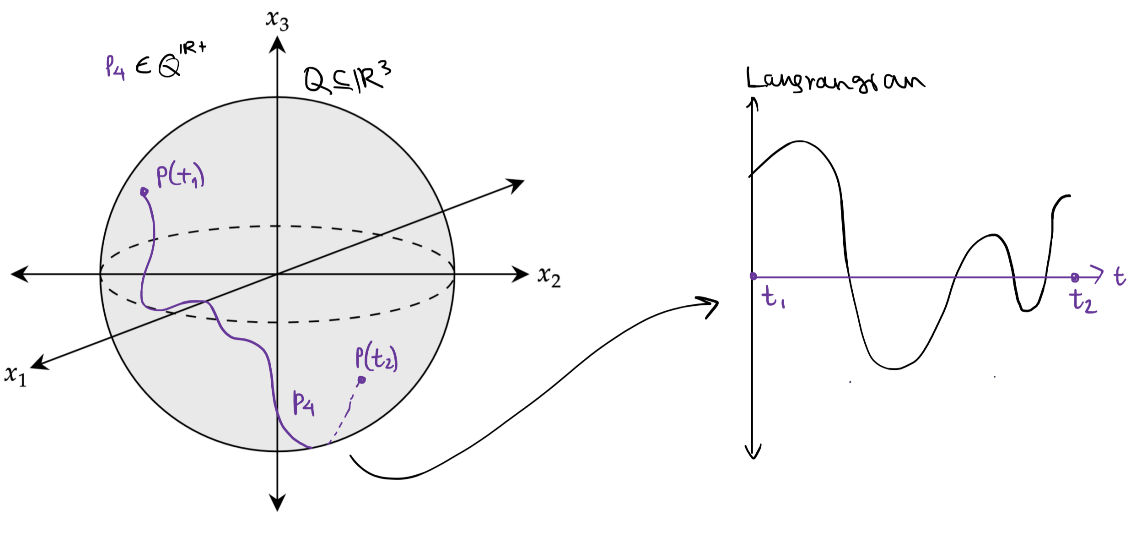

Figure 10: The Lagrangian of $P \subseteq Q$

By defining the Lagrangian with the tangent bundle of the configuration manifold X, applying it application along a path, we have established the geometric picture that underlies the Lagrangian formalism.

6 The Euler-Lagrange Equation

The Euler-Lagrange Equation (EL) is a tool in variational calculus that provides the necessary condition for the extremals of the general functional (Vogtmann et al. 58),

$$I(x) = \int_a^b f(x(t), \dot{x}(t), t)dt$$

The function f is of the same composition as the Lagrangian function L in Def 5.1, hence EL can be applied on L. $I(x)$ is later established as the action in Section 7. For fixed point boundaries, EL is given as (Kubiznak 2; Vogtmann et al. 58),

Definition 6.1: Euler-Lagrange Equations for fixed boundary

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}_z}\right) = \frac{\partial L}{\partial y_z}$$

where z indexes the coordinates involved.

Ultimately, by reducing to a set of second-order EOMs, EL provides a method to find the optimal path $P \subseteq Q$ that a system will take (Kubiznak 2). Furthermore, EL is a coordinate-independent formulation, hence enables transformations between coordinates y_z, particularly from cartesian coordinates, x_i, to GCs, q_j, while its results remain invariant (Holm et al. 13).

In Sections 7 and 8, we derive EL through the Stationary Actions and Virtual Work principle, which relies on the geometric principles introduced. This affirms the validity of the underlying mathematical framework established.

7 Hamilton's Principle of Stationary Action

Formulated by Hamilton in 1834, the Stationary Action Principle is a variational principle in which EL can be derived from (Thornton and Marion 230).

Firstly, the Action is defined as a functional with the function space $X^{\mathbb{R}^+}$ as its domain—it inputs the path of a system $P \subseteq X$ with start and end points t_1, t_2 respectively—and outputs the integral of the Lagrangian along P (Vogtmann et al. 60; Kubiznak 1),

Definition 7.1: Action Functional

$$A : Q^{\mathbb{R}^+} \to \mathbb{R}$$
$$p \mapsto \int_{t_1}^{t_2} L(p(t), \dot{p}(t))dt$$

For instance, for a path $P \subseteq Q$, the shaded area in the Lagrangian graph from Fig 10 represents the value of the action, as shown in Fig 11.

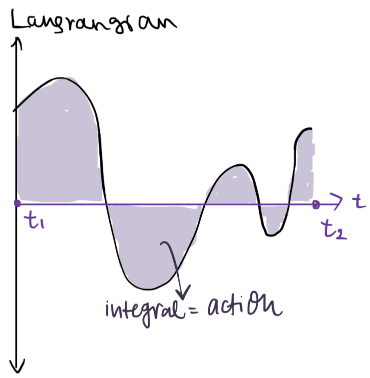

Figure 11: A simplistic breakdown of the Action

By slightly varying the path P, it will induce a change in A. This variation is visually depicted in Fig 13. This leads to,

Principle 7.1: Stationary Action Principle

$$\delta A(P) = \int_{t_1}^{t_2} \delta L(P(t), \dot{P}(t)) = 0$$

(Holm et al. 16)

The differential δ represents an instantaneous change in a value, without a change in time. Therefore, $\delta P(t)$ and $\delta \dot{P}(t)$ represent infinitely small quantities added to the path's position and velocity at each time t, but without a change in t itself (Tchekhovskoy).

7.1 The Role of Tangent Spaces

Recall tangent vectors and spaces in Section 4.4. Of crucial importance is that such displacements $\delta P(t)$ ensures the configuration is still consistent with constraints, ie. $P(t) + \delta P(t) \in Q$. Thus, $\delta P(t)$ must occur along the tangent space of each point $T_{P(t)}X$. Such instantaneous shifts that occur tangentially to the manifold under constant time are called **virtual displacements** (Tchekhovskoy).

Furthermore, the GCs q_j also has a sensible geometric interpretation on the configuration manifold X, as any changes ∂q_j, will be consistent with constraints by definition, and hence remain constricted to X. Consequently, we can construct the tangent space T_xX as a vectorspace with partial derivatives wrt GCs, $\frac{\partial}{\partial q_j}$, as its basis vectors (Kubiznak 11).

In SP, T_xQ as a vectorspace has two basis vectors, $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial \phi}$ as depicted in Fig 12.

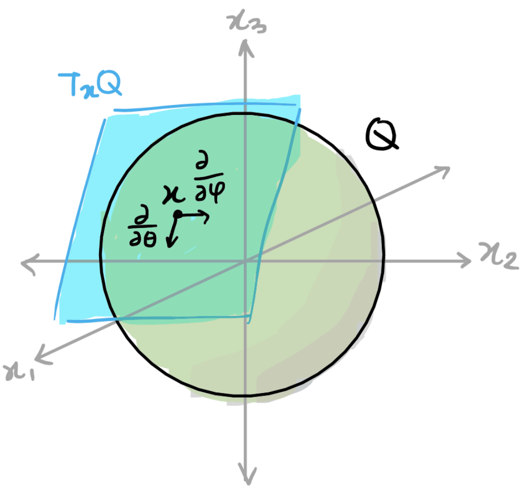

Figure 12: Tangent Space at $x \in Q$, T_xQ, as a 2-dim vectorspace with coordinate basis $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial \phi}$

With this construction, by choosing a certain value of θ and ϕ to virtually displace a point $\vec{x}$ by, then computing the new value $\vec{x}'$ via,

$$\vec{x}' = \vec{x} + \frac{\partial(\vec{x})}{\partial \theta}\theta + \frac{\partial(\vec{x})}{\partial \phi}\phi$$

we can obtain the 3-dim vectors in cartesian coordinates that represent a virtual displacement consistent with constraints.

Therefore, rather than constructing virtual displacements with cartesian coordinates then ensuring it is within X, instead we can first determine the changes in θ, ϕ, then convert back to cartesian coordinates, which automatically satisfies the constraints of the system. Each $\delta P(t)$ along P can be similarly constructed.

7.2 Variations in the Paths of X

When treated as a function of t, $\delta P(t)$ is continuous and has a continuous first derivative $\delta\dot{P}(t)$ which vanishes at fixed points $t = t_1, t_2$. Adding all δP along all P results in a slight variation of the entire path P, with fixed endpoints $P(t_1), P(t_2)$. This new path is denoted $P + \delta P$.

An example path variation is shown for SP in Fig 13, note how $P + \delta P$ shares two fixed points with P, and is contained within Q.

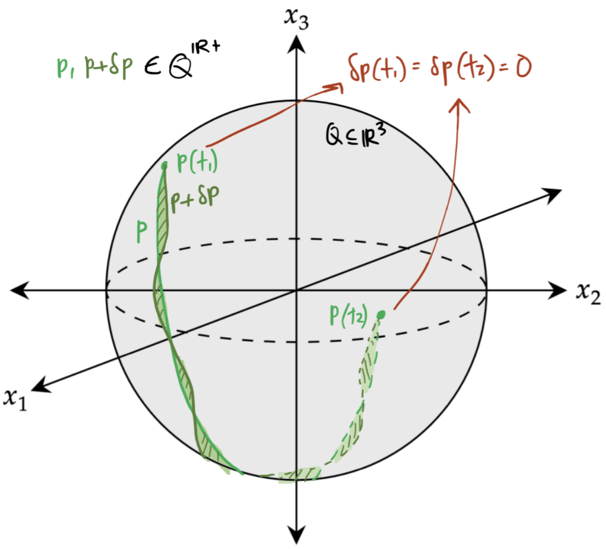

Figure 13: Infinitesimal variations of the path P forming a new path $P + \delta P$

With Hamilton's Principle in 7.1, we are looking for a path, st applying these slight, infinitesimal variations does not change the action, hence $\delta A = 0$. Such a path is the solution to EL.

7.3 Derivation of EL

From Fig 13, consider an arbitrary path P and its variation $P + \delta P$, corresponding to the Lagrangians $L(p, \dot{p})$ and $L(p + \delta p, \dot{p} + \delta\dot{p})$ respectively. Note that $p, \dot{p}$ represents a vector of some arbitrary coordinates p_j, $\dot{p}_j$.

Firstly, we want to obtain a difference between these two Lagrangians to investigate the Action Differential, δA (Cline ch 9.2). By applying the Taylor Expansion for two variables (Seeburger),

$$f(x, y) \approx f(x + a, y + b) + \frac{\partial f}{\partial x}(-a) + \frac{\partial f}{\partial y}(-b)$$

and setting $x = p + \delta p$, $y = \dot{p} + \delta\dot{p}$, we can find derive an expression for $L(p + \delta p, \dot{p} + \delta\dot{p})$,

$$L(p + \delta p, \dot{p} + \delta\dot{p}) \approx L(p, \dot{p}) + \frac{\partial L(p, \dot{p})}{\partial p}\delta p + \frac{\partial L(p, \dot{p})}{\partial \dot{p}}\delta\dot{p}.$$

This gives us a $L(p,\dot{p})$ term in the expansion of $L(p+\delta p, \dot{p}+\delta\dot{p})$ that can be eliminated. For simplicity, $L(p,\dot{p})$ is written as L henceforth.

$$\begin{aligned}\delta L &= L(p+\delta p, \dot{p}+\delta\dot{p}) - L(p,\dot{p}) \\ &= \frac{\partial L}{\partial p}\delta p + \frac{\partial L}{\partial \dot{p}}\delta\dot{p}\end{aligned}$$

As $p, \dot{p}$ are vectors, its partial derivatives can be decomposed into individual coordinates $p_j, \dot{p}_j$, hence

$$\begin{aligned}\delta L &= \left(\frac{\partial L}{\partial p_1}\delta p_1 + \cdots + \frac{\partial L}{\partial p_J}\delta p_J\right) + \left(\frac{\partial L}{\partial \dot{p}_1}\delta\dot{p}_1 + \cdots + \frac{\partial L}{\partial \dot{p}_J}\delta\dot{p}_J\right) \\ &= \left(\frac{\partial L}{\partial p_1}\delta p_1 + \frac{\partial L}{\partial \dot{p}_1}\delta\dot{p}_1\right) + \cdots + \left(\frac{\partial L}{\partial p_J}\delta p_J + \frac{\partial L}{\partial p_J}\delta p_J\right) \\ &= \sum_j \left(\frac{\partial L}{\partial p_j}\delta p_j + \frac{\partial L}{\partial \dot{p}_j}\delta\dot{p}_j\right)\end{aligned}$$

Recall from Principle 7.1 that

$$\delta A = \int_{t_1}^{t_2} \delta L dt$$

Therefore,

$$\delta A = \int_{t_1}^{t_2} \sum_j \left(\frac{\partial L}{\partial p_j}\delta p_j + \frac{\partial L}{\partial \dot{p}_j}\delta\dot{p}_j\right) dt. \tag{12}$$

Utilizing the product rule,

$$u\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}uv}{\mathrm{d}t} - \frac{\mathrm{d}u}{\mathrm{d}t}v \tag{13}$$

and noting the identity (Holm et al. 16),

$$\delta\dot{p}_j = \delta\frac{\mathrm{d}p_j}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\delta p_j\right)$$

setting $u = \frac{\partial L}{\partial \dot{p}_j}$, $v = \delta p_j$, thus $\frac{\mathrm{d}u}{\mathrm{d}t} = \dot{p}_j$, we obtain,

$$\begin{aligned}\frac{\partial L}{\partial \dot{p}_j}\delta\dot{p}_j &= \frac{\partial L}{\partial \dot{p}_j}\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta p_j\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{p}_j}\delta p_j\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{p}_j}\right)\delta p_j\end{aligned} \tag{14}$$

Plugging Eq 14 into Eq 12, then factoring out δp_j,

$$\begin{aligned}\delta A &= \int_{t_1}^{t_2} \sum_j \left(\frac{\partial L}{\partial p_j}\delta p_j + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{p}_j}\delta p_j\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{p}_j}\right)\delta p_j\right) dt \\ &= \int_{t_1}^{t_2} \sum_j \left(\frac{\partial L}{\partial p_j} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{p}_j}\right)\right)\delta p_j dt + \int_{t_1}^{t_2} \sum_j \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{p}_j}\delta p_j\right)\right) dt\end{aligned} \tag{15}$$

Since start and end points $p(t_1), p(t_2)$ are fixed in the variation of the path, $\delta p(t_1) = \delta p(t_2) = 0$, the integral in the second term of Eq 15 ultimately disappears,

$$\begin{aligned}
\int_{t_1}^{t_2} \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{p}_j} \delta p_j \right) \right) dt &= \frac{\partial L(p_j(t), \dot{p}_j(t))}{\partial \dot{p}_j} \delta p_j(t) \Big|_{t_1}^{t_2} \\
&= \frac{\partial L(p_j(t_2), \dot{p}_j(t_2))}{\partial \dot{p}_j} \delta p_j(t_2) - \frac{\partial L(p_j(t_1), \dot{p}_j(t_1))}{\partial \dot{p}_j} \delta p_j(t_1) \\
&= \frac{\partial L(p_j(t_2), \dot{p}_j(t_2))}{\partial \dot{p}_j} 0 - \frac{\partial L(p_j(t_1), \dot{p}_j(t_1))}{\partial \dot{p}_j} 0 \\
&= 0
\end{aligned}$$

Thus Eq 15 reduces to Eq 16, which equals 0 by the Stationary Action Principle,

$$\delta A = \int_{t_1}^{t_2} \sum_j \left(\frac{\partial L}{\partial p_j} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{p}_j} \right) \right) \delta p_j dt = 0 \tag{16}$$

Despite it being infinitely small, each δp_j remains a non-zero quantity (Cline ch 9.2). Thus, the following must be true for the integral in Eq 16 to equal 0,

$$\frac{\partial L}{\partial p_j} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{p}_j} \right) = 0 \Leftrightarrow \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{p}_j} \right) = \frac{\partial L}{\partial p_j}$$

which implies EL,

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{p}_j} \right) = \frac{\partial L}{\partial p_j}.$$

Therefore, we successfully established the mathematics underlying the Stationary Action Principle and from it, derived EL.

8 D'Alembert's Principle of Virtual Work

The Virtual Work Principle utilizes a different viewpoint of the same geometric intuition to derive EL. Firstly, virtual work is the product of Component of Force in the direction of a virtual displacement along X, δx, as depicted in Fig 14,

Definition 8.1: Virtual Work

$$\delta W_i = F_i \cdot \delta x_i$$

for some $i = 1, \ldots, NM$. (Cline ch 6.3; Holm et al. 17)

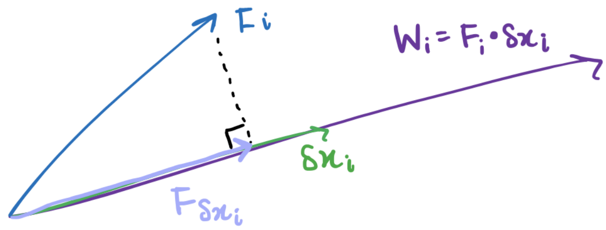

Figure 14: Geometric picture of Work as the dot product of distance and the component of force in the direction of displacement

The Total forces F_i on a system is defined as the inertial force, given by (Holm et al. 17),

$$F_i = m_i \ddot{x}_i.$$

Furthermore, $m_i \ddot{x}_i$, Constraint C_i and Applied A_i forces is related via,

Definition 8.2: Constraint, Applied, and Total Forces

$$m_i \ddot{x}_i = C_i + A_i$$

The Virtual Work Principle is an assumption that this constraint force C_i should not do any work on the system in equilibrium, where the net force is 0 (Cline ch 6.3). Therefore,

Principle 8.1: Virtual Work Principle

For any displacement δx_i of X consistent with constraints,

$$\delta W = \sum_i C_i \cdot \delta x_i = 0$$

The geometric reasoning underlying this principle lies in the normal spaces of the configuration manifold X.

8.1 The Role of Normal Spaces of X

Normal vectors $x \in X$ are orthogonal to the tangent space at x. Moreover, the normal space at x is the set of all normal vectors with base-point x. Thus,

Definition 8.3: Normal Space

Let X be a manifold. The normal space at $x \in X$ is,

$$N_x X = \{N_x | \; \forall V_x \in T_x X, \; N_x \cdot V_x = 0\}$$

Each element of $N_x X$ is a normal vector at x (Holm et al. 51).

Normal spaces of X are physically interpreted as the changes in the configuration that are inconsistent with constraints (Dourmashkin ch 8.3).

In SP, the normal space $N_x Q$ at each point x would only be composed of 2 normal vectors that are directed away or towards the origin.

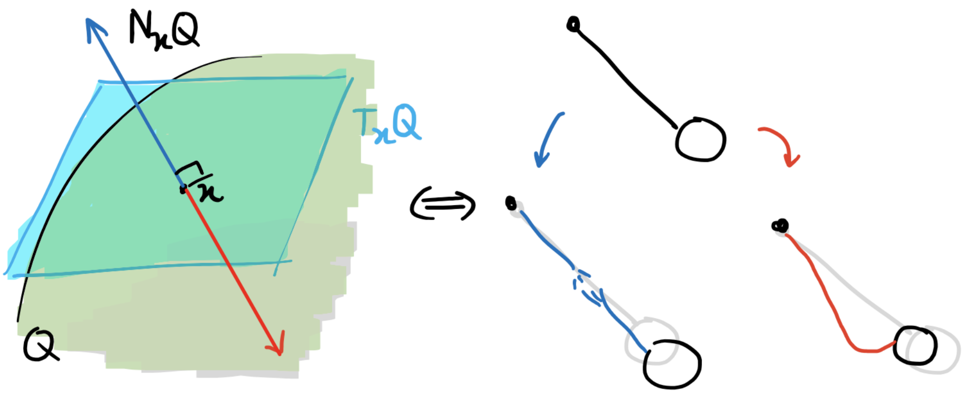

Figure 15: Normal space of $x \in Q$ and its physical interpretation wrt the length constraint

Configuration changes along the normal space implies that the pendulum can extend contract or extend beyond the length of the string, as depicted in Fig 15. However, these changes are inconsistent with the established length constraint.

Therefore, normal vectors correspond to the constraint forces action on a system, as the configuration would only evolve perpendicular to X if the constraint force did work on it. This leads to the geometric construction in Fig 16, where the virtual displacement δx_i occurring tangentially to X is orthogonal to the constraint forces C_i,

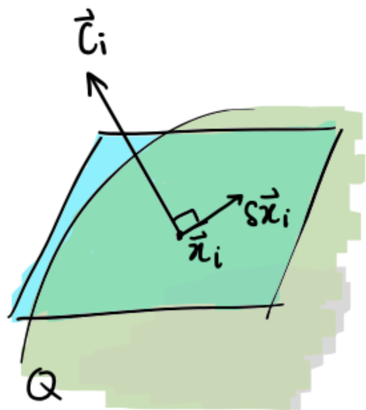

Figure 16: Geometric interpretation of constraint forces and virtual displacement

Since the angle between these two vectors is $\frac{\pi}{2}$, the component of C_i in the direction of δx_i is 0, hence virtual work is 0.

Although we can not explicitly express C_i, it does not negate the ability to apply Principle 8.1. By definition of q_j, changes in q_j will occur tangentially, hence are always orthogonal to the normal space. Thus, the following always holds,

$$\delta W = \sum_i C_i \cdot \delta q_i = 0.$$

Therefore, using GCs in place of arbitrary coordinates x_i implies the Virtual Work Principle is invariably satisfied.

8.2 Derivation of EL

For each i coordinate, using the breakdown of total force in Def 8.2, the work done on a virtual displacement δx_i of the system is,

$$\begin{aligned} \left(\sum_i F_i \cdot \delta x_i =\right) \sum_i (A_i + C_i) \cdot \delta x_i &= \sum_i m_i \ddot{x}_i \cdot \delta x_i \\ \sum_i A_i \cdot \delta x_i + \sum_i C_i \cdot \delta x_i &= \sum_i m_i \ddot{x}_i \cdot \delta x_i \end{aligned} \tag{17}$$

According to Tchekhovskoy, to convert the coordinates x_i to the set of generalized coordinates q_j, we use the following principle,

$$\delta x_i = \frac{\partial x_i}{\partial q_j}\delta q_j$$

thus, substituting into Eq 17,

$$\begin{aligned}\sum_{ij} A_i \cdot \frac{\partial x_i}{\partial q_j}\delta q_j + \sum_{ij} C_i \cdot \frac{\partial x_i}{\partial q_j}\delta q_j &= \sum_{ij} m_i \dot{x}_i \cdot \frac{\partial x_i}{\partial q_j}\delta q_j \\ \sum_{ij} A_i \cdot \frac{\partial x_i}{\partial q_j}\delta q_j + \left(\sum_{ij} C_i \cdot \delta q_j\right)\frac{\partial x_i}{\partial q_j} &= \sum_{ij} m \dot{x}_i \cdot \frac{\partial x_i}{\partial q_j}\delta q_j \end{aligned} \tag{18}$$

By the principle of virtual work (Principle 8.1), the second term, evaluates to 0, thus Eq 18 reduces to,

$$\sum_{ij} A_i \cdot \frac{\partial x_i}{\partial q_j}\delta q_j = \sum_{ij} m_i \ddot{x}_i \cdot \frac{\partial x_i}{\partial q_j}\delta q_j. \tag{19}$$

As q_j are independent of each other, the sum equality in Eq 19 holds for every q_j (Tchekhovskoy). Therefore, we can consider the virtual displacement along δq_j as a whole, eliminating δq_j from both sides.

Consider the LHS, from Def 3.2, we have for a single variable x_i,

$$F(x_i) = -\Delta V(x_i) = -\frac{\partial V(x_i)}{\partial x_i}$$

which implies,

$$F(x_i)\partial x_i = -\partial V(x_i).$$

Therefore, taking F as A,

$$LHS = \sum_i \frac{A_i \partial x_i}{\partial q_j} = -\frac{\partial V}{\partial q_j} \tag{20}$$

Using the product rule in Eq 13, with $u = \frac{m_i \partial x_i}{\partial q_j}$, $dv = \ddot{x}_i$ thus $v = \dot{x}_i$, the RHS becomes,

$$\begin{aligned} RHS &= \sum_i \frac{m_i \ddot{x}_i \partial x_i}{\partial q_j} \\ &= \sum_i \left(\frac{d}{dt}\left(\frac{m_i \dot{x}_i \partial x_i}{\partial q_j}\right) - \dot{x}_i \frac{d}{dt}\left(\frac{m_i \partial x_i}{\partial q_j}\right)\right). \end{aligned} \tag{21}$$

Dividing the numerator and denominator of the first term of Eq 21 with dt and switching the order of differentiation in the second term,

$$\begin{aligned} RHS &= \sum_i \left(\frac{d}{dt}\left(\frac{m_i \dot{x}_i \partial x_i(\frac{1}{dt})}{\partial q_j(\frac{1}{dt})}\right) - m_i \dot{x}_i \frac{\partial}{\partial q_j}\left(\frac{\mathrm{d}\partial x_i}{\mathrm{d}t}\right)\right) \\ &= \sum_i \left(\frac{d}{dt}\left(\frac{m_i \dot{x}_i \partial \dot{x}_i}{\partial \dot{q}_j}\right) - m_i \dot{x}_i \frac{\partial \dot{x}_i}{\partial q_j}\right). \end{aligned} \tag{22}$$

From Def 3.1, we have,

$$T = \sum_i \frac{m_i \dot{x}_i{}^2}{2}$$

Taking the partial derivative of T wrt to $\dot{q}_j$ and q_j respectively,

$$\frac{\partial T}{\partial \dot{q}_j} = \sum_i \frac{2}{2} m_i \dot{x}_i \frac{\partial \dot{x}_i}{\partial \dot{q}_i} \qquad = \sum_i \frac{m_i \dot{x}_i \partial \dot{x}_i}{\partial \dot{q}_i}$$
$$\frac{\partial T}{\partial q_j} = \sum_i \frac{2}{2} m_i \dot{x}_i \frac{\partial \dot{x}_i}{\partial q_j} \qquad = \sum_i m_i \dot{x}_i \frac{\partial \dot{x}_i}{\partial q_j}$$

Therefore, we can substitute T terms into Eq 22

$$RHS = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} \tag{23}$$

Combining LHS = RHS from Eqs 20 and 23 respectively,

$$\begin{aligned} -\frac{\partial V}{\partial q_j} &= \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} \\ \frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j} &= \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) \\ \frac{\partial (T-V)}{\partial q_j} &= \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) \end{aligned}$$

Since V is independent of $\dot{q}_i$, the $\frac{\partial V}{\partial \dot{q}_j}$ term evaluates to 0, hence can be subtracted from the LHS while leaving its value unchanged,

$$\begin{aligned} \frac{\partial (T-V)}{\partial q_j} &= \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial V}{\partial \dot{q}_j}\right) \\ &= \frac{d}{dt}\left(\frac{\partial (T-V)}{\partial \dot{q}_j}\right) \end{aligned} \tag{24}$$

Finally, by replacing $T - V$ with L in Eq 24, we obtain EL,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_j}$$

Consequentially, we successfully established the mathematics underlying the Virtual Work Principle and from it, derived EL.

9 Recovery of Newtonian Dynamics

This section demonstrates that Lagrangian Mechanics is equivalent to Newtonian Mechanics, and applies the former framework on SP.

9.1 General Newtonian Equations of Motion

Starting with the Lagrangian $L = T - V$ from Def 5.1, we replace T and V with their respective definitions in Defs 3.1 and 3.2 (Schneider),

$$L(x_1, \dots, x_M, \dot{x}_1, \dots, \dot{x}_{NM}) = \frac{1}{2}m(\dot{x}_1^2 + \dots + \dot{x}_{NM}^2) - V(x_1, \dots, x_{NM})$$

The partial derivatives of L wrt cartesian coordinates are,

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= -\frac{\partial V(x_1, \dots, x_{NM})}{\partial x_i} \\ \frac{\partial L}{\partial \dot{x}_i} &= \frac{\mathrm{d}}{\mathrm{d}\dot{x}_i}\left(\frac{1}{2}m\dot{x}_i^2\right) = m_i\dot{x}_i \end{aligned} \tag{25}$$

Where $i = 1, \dots, NM$. Further differentiating the latter term wrt time,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_i}\right) = \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}_i) = m\ddot{x}_i \tag{26}$$

Finally, plugging Eqs 25 and 26 into EL (Def 6.1),

$$-\frac{\partial V(x_1, \dots, x_M)}{\partial x_i} = m\ddot{x}_i \tag{27}$$

If we consider the set of Eq 27 for all i, this is exactly the EOMs derived in Eq 3, thus illustrating that the Lagrangian Formalism can recover Newtonian Dynamics.

9.2 Equations of Motion of the Spherical Pendulum

This section breaks down the derivation of SP's EOMs based on the process above. Since SP deals with gravitational potential, it follows that,

$$\begin{aligned} T(x_1, x_2, x_3) &= \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) \\ V(x_1, x_2, x_3) &= mgx_3 \\ L(x_1, x_2, x_3, \dot{x}_1, \dot{x}_2, \dot{x}_3) &= \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) - mgx_3 \end{aligned}$$

Writing x_1, x_2, x_3 in terms of θ and ϕ according to the coordinate change equations for position and velocity in Eqs 5, the Lagrangian becomes,

$$L = \frac{1}{2}ml^2(\dot{\theta}^2 + \dot{\phi}^2\sin^2(\theta)) + mgl\cos(\theta) \tag{28}$$

The full computation of Eq 28 can be found in Appendix A. With the Lagrangian in terms of θ and ϕ, we can find the EOMs of SP:

For $i = \theta$,

$$
\begin{aligned}
\frac{\partial L}{\partial \dot{q}_i} &= \frac{\mathrm{d}}{\mathrm{d}\dot{\theta}} \left(\frac{1}{2} m l^2 \dot{\theta}^2 \right) \\
&= m l^2 \dot{\theta} \\
\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) &= m l^2 \ddot{\theta} \\
\frac{\partial L}{\partial q_i} &= \frac{\mathrm{d}}{\mathrm{d}\theta} \left(\frac{1}{2} m l^2 \dot{\phi}^2 \sin^2(\theta) + m g l \cos(\theta) \right) \\
&= m l^2 \dot{\phi}^2 \sin(\theta) \cos(\theta) - m g l \sin(\theta)
\end{aligned}
$$

substituting into the Lagrangian,

$$
\begin{aligned}
0 &= m l^2 \ddot{\theta} - (m l^2 \dot{\phi}^2 \sin(\theta) \cos(\theta) - m g l \sin(\theta)) \\
&= m l^2 \left(\ddot{\theta} - \dot{\phi}^2 \sin(\theta) \cos(\theta) + \frac{g}{l} \sin(\theta) \right)
\end{aligned}
$$

As $m l^2$ is non-zero, this results in the EOM,

$$
\boxed{0 = \ddot{\theta} - \dot{\phi}^2 \sin(\theta) \cos(\theta) + \frac{g}{l} \sin(\theta) \quad (\text{for } i = \theta)}
$$

For $i = \phi$,

$$
\begin{aligned}
\frac{\partial L}{\partial \dot{q}_i} &= \frac{\mathrm{d}}{\mathrm{d}\dot{\phi}} \left(\frac{1}{2} m l^2 \dot{\phi}^2 \sin^2(\theta) \right) \\
&= m l^2 \dot{\phi} \sin^2(\theta) \\
\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) &= m l^2 \frac{\mathrm{d}}{\mathrm{d}t} \left(\dot{\phi} \sin^2(\theta) \right) \\
&= m l^2 \left(\ddot{\phi} \sin^2(\theta) + 2 \dot{\phi} \sin(\theta) \cos(\theta) \right)
\end{aligned}
$$

Since the $\frac{\partial L}{\partial q_i}$ evaluates to 0, and $m l^2$ is non-zero, the resulting EL is,

$$
\begin{aligned}
0 &= \ddot{\phi} \sin^2(\theta) + 2 \dot{\phi} \sin(\theta) \cos(\theta) \\
\ddot{\phi} \sin^2(\theta) &= -\dot{\phi} 2 \sin(\theta) \cos(\theta) \\
\ddot{\phi} &= -2 \dot{\phi} \cot(\theta)
\end{aligned}
$$

Thus, the EOM is,

$$
\boxed{0 = \ddot{\phi} + 2 \dot{\phi} \cot(\theta) \quad (\text{for } i = \phi)}
$$

While deriving the EOMs of systems like SP through pure Newtonian principles can be extremely tedious, this illustrates that the Lagrangian Formalism enables for a relatively convenient process of arriving at the same EOMs (Hirvonen).

10 Conclusion

In this paper, I introduced how the application of constraints and GCs reduces the DOFs of a system. Then, I illustrated that the configurations consistent with constraints is formulated as a manifold, the system's physical time evolution is modelled as a path on the manifold, and the Lagrangian and Action are functions along such paths.

Tangent spaces of this manifold are interpreted as configurations changes that are consistent with the constraints on the system and underlies the Stationary Action Principle. Infinitesimal variations along a given path will aggregate to construct another path which is also contained within the manifold, otherwise it would violate the constraints. Therefore, the infinitesimal variations must occur tangentially, along the tangent spaces at each point.

Normal spaces of this manifold are interpreted as changes inconsistent with the constraints, where constraint forces on the system applies a force in a direction normal to the manifold at every point. Thus constraint forces are formulated as normal vectors. When combined with tangent vectors, this forms the foundation of the Virtual Work Principle, which states the work done by constraint forces (normal vectors) on virtual displacements consistent with the constraints of the system (tangent vectors) will be 0, as these vectors are orthogonal.

Then, I showed the derivations of EL based on these two principles and their associated mathematical concepts. Finally, I illustrated the equivalence between the Lagrangian and Newtonian Formalism, and deriving the EOMs of SP, thus affirming the validity of the underlying mathematics.

10.1 Evaluation

Many textbook and resources on the subject lack sufficient concrete explanation of the mathematical formulations of the Lagrangian framework, as they tend to be concise, abstract, and assume a substantial background knowledge. With guidance from the research question, the information was primarily synthesized from four sources: Vogtmann et al., Holm et al., Thornton and Marion and Kubiznak, which originates from reputable publishers and organizations, such as Perimeter Institute, Oxford University Press and Springer. The former three textbooks have 2468, 19802 and 391 citations respectively, illustrating their reputation and credibility within academic circles (Google Scholar). Ultimately, this paper is distinctive among existing works, as it illustrates the mathematics underlying the Lagrangian formalism through both abstract and concrete lenses in a sequential, cohesive order, where mathematical definitions are broken down into simpler explanations and references to a concrete example were made throughout to fully consolidate with the material.

10.2 Extensions

Various aspects within elementary Lagrangian Mechanics extends beyond this paper, such as **nonholonomic** constraints that do not reduce a system's DOF (Vogtmann et al. 249) and lead to beautiful geometric insights of **Lagrange multipliers** (Hirvonen). Further tools in manifold analysis also provide deeper insights, for instance, using the **Jacobi Operator** to the constraint function in Eq 6 to define curvature of the manifold (Holm et al. 18).

Furthermore, the **Hamiltonian formalism** is an extension of the Lagrangian formalism (Vogtmann et al. 266), both of which are incredibly powerful and widely applied across Physics. For instance, Lagrangians for fields find applications in **Classical Field Theories** and **Quantum Mechanics**. The Hamiltonian also contains various mathematical intricacies, leading to symplectic geometries and applications in Quantum Computing among other fields.

11 Works Cited

Cline, Douglas. "Variational Principles in Classical Mechanics". Jan. 2021. phys.libretexts.org/Bookshelves/Classical_Mechanics/Variational_Principles_in_Classical_Mechanics_(Cline).

Dourmashkin, Peter. "Classical Mechanics". July 2022. phys.libretexts.org/Bookshelves/Classical_Mechanics/Classical_Mechanics_(Dourmashkin).

Google Scholar. Feb. 2023.

Grossmann, Julian P. "Manifolds - Part 13 - Examples of Smooth Manifolds", May 2022, Youtube Video, www.youtube.com/watch?v=wMt5iaaQxjU.

Hirvonen, Ville. "Lagrangian Mechanics For Dummies: An Intuitive Introduction", profoundphysics.com/lagrangian-mechanics-for-beginners/.

Holm, Darryl, et al. *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Jan. 2009.

Keng, Brian. "Manifolds: A Gentle Introduction", Apr. 2018, bjlkeng.github.io/posts/manifolds.

Kubiznak, David. "Theoretical Mechanics". Study Text from the Theoretical Mechanics course offered at the 2017/2018 Perimeter Scholars International program, 2017.

Ou, Ye-Lin, and Bang-Yen Chen. *Biharmonic Submanifolds and Biharmonic Maps in Riemannian Geometry*. World Scientific Publishing, Apr. 2020, www.worldscientific.com/doi/pdf/10.1142/9789811212383_0001.

Rowland, Todd. "Manifold", From MathWorld–A Wolfram Web Resource, created by Eric W. Weisstein, mathworld.wolfram.com/Manifold.html.

Schneider, Elliot. "Explaining the Principle of Least Action: Physics Mini Lesson", Oct. 2017, Youtube Video, www.youtube.com/watch?v=sUk9y23FPHk.

Seeburger, Paul. "Multivariable Calculus". Nov. 2020. math.libretexts.org/Bookshelves/Calculus/Supplemental_Modules_(Calculus)/Multivariable_Calculus.

Tchekhovskoy, Sasha. "Lecture 3, Virtual Work and Lagrange Equations, Physics-411, Classical Mechanics", Dec. 2020, Youtube Video, www.youtube.com/watch?v=1BfDqrLAHbc.

Thornton, S.T., and J.B. Marion. *Classical Dynamics of Particles and Systems*. Brooks/Cole, 2004. *Google Books*, books.google.com.kh/books?id=HOqLQgAACAAJ.

Tsokos, K.A. *Physics for the IB Diploma*. 6th ed., Cambridge UP, July 2014. *Google Books*, books.google.com.kh/books?id=HOqLQgAACAAJ.

Vogtmann, K., et al. *Mathematical Methods of Classical Mechanics*. Springer New York, 1997. *Google Books*, books.google.com.kh/books?id=Pd8-s6rOt%5C_cC. Graduate Texts in Mathematics.

Voitsekhovskii, M.I. "Differentiable manifold". Dec. 2021. encyclopediaofmath.org/index.php?title=Differentiable_manifold.

Appendix A Computations

$$
\begin{aligned}
\dot{x}_1^2 =& l^2\dot{\theta}^2\cos^2(\theta)\cos^2(\phi) \\
&- 2l^2\dot{\theta}\dot{\phi}\cos^2(\theta)\cos^2(\phi)\sin^2(\theta)\sin^2(\phi) \\
&+ l^2\dot{\phi}^2\sin^2(\theta)\sin^2(\phi) \\
\dot{x}_2^2 =& l^2\dot{\theta}^2\cos^2(\theta)\sin^2(\phi) \\
&+ 2l^2\dot{\theta}\dot{\phi}\cos^2(\theta)\cos^2(\phi)\sin^2(\theta)\sin^2(\phi) \\
&+ l^2\dot{\phi}^2\sin^2(\theta)\cos^2(\phi) \\
\dot{x}_3^2 =& l^2\dot{\theta}^2\sin^2(\theta) \\
\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2 =& l^2\dot{\theta}^2\cos^2(\theta)(\sin^2(\phi) + \cos^2(\phi)) \\
&+ l^2\dot{\phi}^2\sin^2(\theta)(\sin^2(\phi) + \cos^2(\phi)) \\
&+ l^2\dot{\theta}^2\sin^2(\theta) \\
&= l^2(\dot{\theta}^2(\cos^2(\theta) + \sin^2(\theta)) + \dot{\phi}^2\sin^2(\theta)) \\
&= l^2(\dot{\theta}^2 + \dot{\phi}^2\sin^2(\theta))
\end{aligned}
\tag{29}
$$