Coordinate-free derivation of the Euler–Lagrange equations and identification of global solutions via local behavior*

Elsa K. Hansen†

2005/02/21

Abstract

Results concerning C^2-minimizing curves on manifolds are presented. A coordinate-free derivation of the Euler–Lagrange equation is presented. Using a variational approach, two vector fields are defined along the minimizing curve; the tangent to the curve $\dot{\gamma}$, and the infinitesimal variation $\delta\sigma$. The derivation presented involves complete lifts of arbitrary extensions of these vector fields and it is shown that the derivation is independent of the particular choice of extensions. Special care is also taken to ensure that the derivation does not require any additional differentiability constraints, other than γ being of class C^2.

Minimizing curves are also characterized in terms of their local behaviour. It is shown that if a curve is minimizing then any sub-arc of the curve is also minimizing. An important corollary of this result is that a curve, γ, on a manifold will be minimizing only if any collection of admissible charts which cover γ have minimizing local representations.

Contents

*Report for project in fulfilment of requirements for MSc

†Graduate student, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada

Email: `elsa@mast.queensu.ca`

Research supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

1. Introduction

It is well known that a necessary condition for a C^2-curve to solve the calculus of variations minimization problem, is that it satisfies the Euler–Lagrange equations. Since the nature of the minimization problem is coordinate-independent, it follows that the solution should also be coordinate-independent. This is in fact the case and is shown directly by observing how the Euler–Lagrange equations change when one changes coordinates [Bullo and Lewis 2004]. The coordinate-independent nature of the Euler–Lagrange equations recommends the existence of a coordinate-free expression. The standard approach to finding such an expression is to postulate its form and then prove that it agrees with the standard coordinate expression [Crampin and Pirani 1986]. There have been a few attempts at the intellectually more satisfying approach of finding a coordinate-free derivation that parallels the variational method used in the standard coordinate derivation.

Two approaches towards a coordinate-free derivation which we found in the literature are given in [Nester 1988] and [Gamboa Saraví and Solomin 2003]. The most striking difference between these two approaches is that [Gamboa Saraví and Solomin 2003] postulates the existence of a covariant derivative and then proceeds to write an intrinsic expression for the Euler–Lagrange equations in terms of that covariant derivative. In contrast, [Nester 1988] makes no use of a covariant derivative and derives the Euler–Lagrange equations using only vector fields defined on TM. Since the second method promises to be more general, this is the method that is the topic of this report.

More specifically, we make an in-depth study of the derivation given by Nester in [Nester 1988]. In particular, we show that the vector fields used in the derivation can be expressed as the complete lifts of the two vector fields on M which arise naturally from the variational formulation. These two vector fields are the tangent vector field of the curve in question and the vector field representing the infinitesimal variation. Since these vector fields on M are not necessarily defined on open subsets, it is necessary to extend them in order to define the complete lifts. Since the choice of extension is arbitrary, this immediately brings into question whether or not the derivation given by Nester is independent of vector field extension. We first show that it is possible to extend these vector fields, and then we show that the derivation is in fact independent of vector field extension.

Using complete lifts introduces a further wrinkle. The derivation given in [Nester 1988] takes the Lie bracket of the two vector fields defined on TM. This effectively tightens the differentiability constraints on the curve. We show that with a slight modification of the derivation we can remove these additional differentiability constraints. In [Nester 1988] the matter of smoothness is blithely sidestepped.

In the final section of this report we investigate the possibility of identifying a minimizing curve on a manifold M by its local behaviour. The approach here is to require that given any set of charts covering a minimizing curve on M, the curves coordinate representations solve corresponding minimization problems. This approach is only valid if a minimizing curve on a certain interval must also be a minimizing curve on any subinterval. We prove that this is in fact the case.

This report is organized as follows. In Chapter 2 we review some of the main concepts and definitions which will be used. In Chapter 3 we begin by defining the concept of a variation and giving the coordinate-free statement of the minimization problem. Next we show that, provided the vector fields defining the variation are defined in an open neighborhood, then the problem statement can be rephrased using complete lifts. We conclude Chapter 3 by showing that we can always extend the tangent vector of the curve and the infinitesimal variation to a neighbourhood of the curve. In Chapter 4 we present the main results of our report. We first give the coordinate-free derivation of the Euler–Lagrange equations and then show that if a curve is a minimizer on a certain interval, I, then it is also a minimizing curve on any subinterval of I.

2. Preliminaries

In this chapter we review some of the major definitions and concepts that will play prominent roles in the sequel. We begin with a basic review of topology with the aim of defining a partition of unity. Next we define the Lie derivative in terms of flows of vector fields. We then describe two different ways of lifting a vector field from M to TM. Following this, we define the exponential map and describe those properties of the exponential map which will be used in our development. Lastly we give a basic review of some algebraic terminology.

2.1. Topological Definitions and Partitions of Unity.

In this section only definitions and results directly used in this work are presented. A detailed discussion of partitions of unity can be found in [Abraham, Marsden, and Ratiu 1988]. Let $(S, \mathcal{O})$ be a topological space.

2.1 Definition: The subset $\mathcal{B} \subset \mathcal{O}$ is a ***basis for*** $\boldsymbol{\mathcal{O}}$ if every element $\mathcal{G} \in \mathcal{O}$ can be written as an arbitrary union of elements belonging to $\mathcal{B}$. •

2.2 Definition: $(S, \mathcal{O})$ is ***second countable*** if it has a countable basis. •

2.3 Definition: $(S, \mathcal{O})$ is ***Hausdorff*** if, for any two elements $x_1, x_2 \in S$, there exist open sets $\mathcal{G}_1, \mathcal{G}_2 \in \mathcal{O}$ satisfying $x_1 \in \mathcal{G}_1$, $x_2 \in \mathcal{G}_2$ and $\mathcal{G}_1 \cap \mathcal{G}_2 = \varnothing$. •

2.4 Definition: A ***cover*** of $(S, \mathcal{O})$ is a subset $\{\mathcal{G}_i\}_{i \in I} \subset \mathcal{O}$ satisfying $S = \cup_{i \in I} \mathcal{G}_i$, where I is an arbitrary index set. •

2.5 Definition: $\{\mathcal{G}_i\}_{i \in I} \subset \mathcal{O}$ is a ***locally finite cover*** if it is a cover and if, for each $x \in S$, there exists a neighbourhood U of x for which the set $\{\, i \mid U \cap \mathcal{G}_i \neq \varnothing\}$ is finite. •

2.6 Definition: A ***refinement*** of a cover $\{\mathcal{G}_i\}_{i \in I}$ is a cover $\{\mathcal{H}_j\}_{j \in J}$ if for each $j \in J$, there exists an $i \in I$ such that $\mathcal{H}_j \subset \mathcal{G}_i$. •

2.7 Definition: $(S, \mathcal{O})$ is ***paracompact*** if it is Hausdorff and if every cover has a locally finite refinement. •

2.8 Definition: A ***C^r-partition of unity*** on a manifold M is a collection of pairs $\{(U_i, g_i)\}_{i\in I}$ satisfying

(i) $\{U_i\}$ is a locally finite covering of M,

(ii) $g_i \in C^r(M, \mathbb{R})$ and satisfies $g_i(x) \geq 0 \ \forall\ x \in U_i$ and $g_i(x) = 0 \ \forall\ x \in M \backslash U_i$,

(iii) $\sum_i g_i(x) = 1$ for each $x \in M$.

If $\mathcal{A} = \{(V_j, \phi_j)\}_{j\in A}$ is an atlas for M then a partition of unity $\{(U_i, g_i)\}_{i\in I}$ is ***subordinate to $\mathcal{A}$*** if for each i there exists a j such that $U_i \subset V_j$. If all atlases for a manifold M have a subordinate partition of unity then ***M admits partitions of unity***. •

The proof of the following theorem can be found in [Abraham, Marsden, and Ratiu 1988].

2.9 Theorem: *Every second countable, Hausdorff n-dimensional manifold admits a C^∞-partition of unity.*

2.2. The Flow Interpretation of the Lie Derivative.

In the sequel, a useful characterization of the Lie derivative will be its flow interpretation. Before giving the flow interpretation we must provide a few definitions.

2.10 Definition: Let $f : M \to N$ be a C^1-map between manifolds M and N and let w be a k-form on N. The ***pull-back of w by f*** is defined as $f^*w(x)(v_1, \ldots, v_k) = w(f(x))(T_xf(v_1), \ldots, T_xf(v_k))$, where $v_1, \ldots, v_k \in T_xM$. •

That is, the pull-back of a differential form on N defines a new differential form on M by pulling back the operation of w through f. In the case where w is a function, the pull-back takes the form $f^*w = w \circ f$.

2.11 Definition: Let $f : M \to N$ be a C^1-diffeomorphism between manifolds M and N and let X be a vector field on M. The ***push-forward of X by f*** is defined as $f_*X(y) = T_xf \circ X \circ f^{-1}(y)$. •

That is, the push-forward of a vector field on M defines a new vector field on N by pushing forward the operation of X through f. An important difference between Definitions 2.10 and 2.11 is that, in order for f_*X to exist, f must be a diffeomorphism so that every element $y \in N$ can be mapped to an element $x \in M$ through the inverse of f.

It is also possible to define the push-forward of a k-form by a diffeomorphism:

$$f_*w = w(f^{-1}(x))(T_xf^{-1}(v_1), \ldots, T_xf^{-1}(v_k)).$$

Similarly, one can define the pull-back of a vector field:

$$f^*X = Tf^{-1} \circ X \circ f.$$

The push-forward of a vector field can be used to transport a vector from one point on a manifold to another. The transport construction is as follows. For t small enough, let $\Phi_t^X(x)$ denote the one-parameter group of local diffeomorphisms which assigns to $x \in M$ the point which corresponds to starting at x and flowing along X for time t. For a fixed t and any $v_x \in T_xM$, the Lie transport of v_x with respect to the flow of X at time t is $T\Phi_t^X(v_x)$.

With the Lie transport defined, the Lie derivative can now be interpreted as a measure of how similar a vector field is to its Lie transport. The details are specified in the next definition.

2.12 Definition: Let Y be a vector field defined along the integral curve γ of X. The ***Lie derivative of Y with respect to X*** is given by,

$$\mathcal{L}_X Y(\gamma(s)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} T\Phi^X_{-t} Y(\gamma(t+s)). \qquad \bullet$$

In Definition 2.12 the Lie derivative is used as a measure of how $Y(\gamma(t+s))$ varies from the Lie transport of $Y(\gamma(s))$ along γ. Thus $\mathcal{L}_X Y(\gamma(s)) = 0$ means that $Y(\gamma(t+s))$ *is* the Lie transport of $Y(\gamma(s))$ along γ.

The Lie derivative of a C^1-function can also be given a flow interpretation.

2.13 Definition: The ***Lie derivative of a C^1-function f with respect to X*** is given by

$$\mathcal{L}_X f(\gamma(s)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi^{X*}_t f(\gamma(s)). \qquad \bullet$$

2.14 Definition: The ***Lie bracket*** of two vector fields $X, Z \in \Gamma^r(TM)$, $r \geq 1$, produces a new vector field $[X, Z]$ which satisfies, for any $g \in C^2(M, \mathbb{R})$,

$$\mathcal{L}_{[X,Z]} g = \mathcal{L}_X \mathcal{L}_Z g - \mathcal{L}_Z \mathcal{L}_X g. \qquad \bullet$$

2.15 Theorem: *Let X, $Z \in \Gamma^r(M)$, Y, $W \in \Gamma^r(N)$, $r \geq 1$, and $f \in C^k(M, N)$, $2 \leq k \leq r$. We say that X and Y are **f-related**, denoted by $X \sim_f Y$, if $Y \circ f = Tf \circ X$. If $X \sim_f Y$ and $Z \sim_f W$, then $[X, Z] \sim_f [Y, W]$.*

Proof: This proof follows from one presented in [Kolář, Michor, and Slovák 1993]. Let $h \in C^2(N, \mathbb{R})$ so that $h \circ f$ defines a C^2-function on M. Using the chain rule,

$$(\mathcal{L}_Z(h \circ f))(x) = \mathrm{d}(h \circ f)(Z(x)) = \mathrm{d}h(Tf(Z(x))).$$

Now, since $Z \sim_f W$,

$$\mathrm{d}h(Tf(Z(x))) = \mathrm{d}h(W \circ f(x)) = (\mathcal{L}_W h)(f(x)).$$

Therefore, we have the identity,

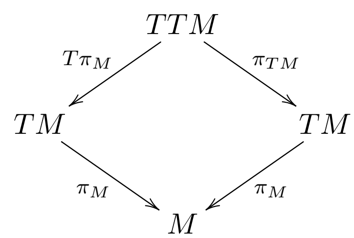
$$\mathcal{L}_Z(h \circ f)(x) = \mathcal{L}_W h(f(x)). \tag{2.1}$$

Substituting (2.1) into the definition of the Lie bracket,

$$\begin{aligned}
\mathcal{L}_{[X,Z]}(h \circ f)(x) &= \mathcal{L}_X \mathcal{L}_Z (h \circ f)(x) - \mathcal{L}_Z \mathcal{L}_X (h \circ f)(x) \\
&= \mathcal{L}_X \mathcal{L}_W h(f(x)) - \mathcal{L}_Z \mathcal{L}_Y h(f(x)) \\
&= \mathcal{L}_Y \mathcal{L}_W h(f(x)) - \mathcal{L}_W \mathcal{L}_Y h(f(x)) \\
&= \mathcal{L}_{[Y,W]} h(f(x)),
\end{aligned}$$

where the second equality comes from recognizing that the Lie derivative of a function $h : N \to \mathbb{R}$ is a new function $\mathcal{L}_W h : N \to \mathbb{R}$ and then substituting (2.1). ■

2.3. Lifting Vector Fields to the Tangent Bundle. In the sequel we use two different constructions to lift vector fields defined on M to vector fields defined on TM. To introduce a discussion of these constructions, we first review a few facts about TTM. In coordinates, an element in TTM has the form $((u, v), (e_1, e_2))$. There are two commonly used projections for TTM. The first is the tangent bundle projection $\pi_{TM} : TTM \to TM$, which in coordinates looks like $\pi_{TM}((u, v), (e_1, e_2)) = (u, v)$. The second is the projection $T\pi_M : TTM \to TM$, which in coordinates has the form $T\pi_M((u, v), (e_1, e_2)) = (u, e_1)$. The following commutative diagram shows how these two projections commute with the standard projection $\pi_M : TM \to M$:



The ***vertical subbundle*** of TM is intrinsically defined as $\ker(T\pi_M)$. The first construction we present is the complete lift.

2.16 Definition: The ***complete lift*** of a vector field $X \in \Gamma^r(TM)$, $r \geq 1$, is denoted by $X^T \in \Gamma^{r-1}(TTM)$, and is defined as

$$X^T(v_x) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} T_x\Phi_t^X(v_x). \qquad \bullet$$

Given a local chart (U, ϕ) on M, the coordinate representation of the principle part of X^T is given by $X_\phi^T(x, v) = (X_\phi(x), \frac{\partial X_\phi(x)}{\partial x} \cdot v)$, where X_ϕ^T and X_ϕ are the coordinate representations of the principle parts of X^T and X respectively. Comparing Definition 2.16 with that of the Lie transport, we see that the complete lift can also be defined in terms of the Lie transport. That is, the complete lift of a vector field evaluated at some point v_x in the tangent bundle is the tangent vector of the curve defined by the Lie transport of v_x along the flow of X.

The second lifting construction is the vertical lift.

2.17 Definition: The ***vertical lift*** of a tangent vector $w_x \in T_xM$ is defined as

$$\mathrm{vlft}_{v_x}(w_x) = \frac{d}{dt}\bigg|_{t=0} (v_x + tw_x) \in T_{v_x}TM. \qquad \bullet$$

If X is a vector field then define the vector field $\mathrm{vlft}(X) \in \Gamma(TTM)$ by $\mathrm{vlft}(X)(v_x) = \mathrm{vlft}_{v_x}(X(x))$.

Given a local coordinate chart (U, ϕ) on M, the coordinate representation of the principle part of $\mathrm{vlft}_{v_x}(X(x))$ is given by $(\mathrm{vlft}_{v_x}(X(x)))_\phi = (0, X_\phi(x))$. Note that $\mathrm{vlft}_{v_x}(X(x))$ belongs to the vertical subbundle of TM.

A significant difference between the complete lift and the vertical lift of X is that the vertical lift is a pointwise construction whereas the complete lift depends on how X is defined in a neighbourhood of the point of evaluation. This distinction is evident for example, because X^T depends on the Lie transport of v_x. It is also evident by looking at the coordinate expressions of the two different lifts; the coordinate expression for X^T contains derivatives of X.

Two important constructions which use the vertical lift are the almost tangent structure $J : TTM \to TTM$ which is defined as

$$J(w_{v_x}) := \mathrm{vlft}_{v_x}(T\pi_M(w_{v_x})) \text{ for all } w_{v_x} \in TTM$$

and the Liouville vector field $V : TM \to TTM$, which is defined as

$$V(v_x) := \mathrm{vlft}_{v_x}(v_x).$$

2.4. The Exponential Map. Before presenting properties of the exponential map, we give a few definitions.

2.18 Definition: For $r \geq 1$, a ***C^r-affine connection*** on M, denoted by ∇, assigns to any two vector fields X, $Z \in \Gamma^r(TM)$ a new vector field $\nabla_X Z \in \Gamma^{r-1}(TM)$. For any $f \in C^r(M, \mathbb{R})$ let $fX(x) = f(x)X(x)$, then the assignment satisfies the following properties:

(i) the map $(X, Z) \mapsto \nabla_X Z$ is $\mathbb{R}$-bilinear;

(ii) $\nabla_{fX} Z = f\nabla_X Z$;

(iii) $\nabla_X fZ = f\nabla_X Z + (\mathscr{L}_X f)Z$. •

2.19 Definition: A ***geodesic*** of a C^r-connection is a differentiable curve $\gamma : I \to M$ that satisfies $\nabla_{\dot\gamma(t)}\dot\gamma(t) = 0$ for all $t \in I$. •

2.20 Definition: A ***Riemannian manifold***, $(M, \mathbb{G})$ is a manifold M with a Riemannian metric $\mathbb{G}$. •

The Riemannian metric of a Riemannian manifold uniquely determines an affine connection. This affine connection is called the Levi-Civita connection.

Let M be a manifold with connection ∇. For $x \in M$, the exponential map $\exp : T_xM \to M$ is a local diffeomorphism from an open neighbourhood of the zero vector of T_xM to an open neighbourhood of x. Let $\beta_{v_x}(t)$ be the geodesic, associated with the connection ∇, that satisfies the initial condition $\beta_{v_x}(0) = v_x$. Then the exponential map is defined by $\exp(v_x) = \beta_{v_x}(1)$. In order to show that the exponential map is a local diffeomorphism, we will follow the approach in [Crampin and Pirani 1986] and show that, on an appropriate set, $T\exp$ is the identity and then apply the Inverse Function Theorem. First, for completeness we state the Inverse Function Theorem. The following version of the Inverse Function Theorem is from [Bullo and Lewis 2004].

2.21 Theorem: (Inverse Function Theorem) *Let $f \in C^1(M, N)$. If Tf is an isomorphism at x_0, then there exists a neighbourhood U of x_0 for which $f|U : U \to f(U)$ is a diffeomorphism.*

Let Id denote the following identification of $v_x \in T_xM$ with $w_{0_x} \in T_{0_x}TM$,

$$w_{0_x} = \mathrm{vlft}_{0_x}(v_x).$$

To show that $T_{0_x}\exp : T_{0_x}(T_xM) \to T_xM$ is equal to Id, first note that, for $s > 0$ sufficiently small, geodesics have the homogenity property that $\beta_{sv_x}(1) = \beta_{v_x}(s)$.

Applying the chain rule,

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp((tv)_x) = T_{0_x}\exp\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} tv\right) = T_{0_x}\exp\circ\,\mathrm{vlft}_{0_x}(v_x).$$

But by definition we know that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp((tv)_x) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \beta_{tv}(1) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \beta_v(t) = v_x.$$

Therefore we have the desired result,

$$T_{0_x}\exp(\mathrm{vlft}_{0_x}(v)) = v_x. \tag{2.2}$$

Finally, the identity is certainly non-singular so, by the Inverse Function Theorem, exp is a local diffeomorphism of a neighbourhood of $0_x \in T_xM$ to a neighbourhood of $x \in M$.

The exponential map can be used to define a local diffeomorphism from an open neighbourhood of the zero section of an embedded submanifold to an open neighbourhood of the submanifold. This construction is defined in the following theorem and will be used later to extend vector fields from embeddings to open neighbourhoods.

2.22 Theorem: *Let M be a Riemannian manifold and N an embedded submanifold of M. Let $N^{\perp}$ denote the normal bundleto TN in TM. Then $\exp|\,N^{\perp}$ defines a local diffeomorphism from an open neighbourhood $V \subset N^{\perp}$ to an open neighbourhood $U \subset M$, where U satisfies $N \subset U$. U is called the* ***tubular neighbourhood*** *of N.*

2.5. Algebraic Terminology. This section states some basic definitions and identities which will be used in later sections.

2.23 Definition: The ***permutation group***, denoted by S_n, is the set of bijections σ of the set of n elements $\{1, \ldots, n\}$. The operation of σ can be diagrammatically explained by

$$\begin{pmatrix} 1 & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) \end{pmatrix}.$$

•

2.24 Definition: $\sigma \in S_n$ is a ***transposition*** if

$$\begin{pmatrix} 1 & \ldots & i & \ldots & j & \ldots & n \\ \sigma(1) & \ldots & j & \ldots & i & \ldots & \sigma(n) \end{pmatrix}.$$

That is, if σ transposes exactly two elements of $\{1, \ldots, n\}$ and leaves all other elements fixed. Each $\sigma \in S_n$ can be written as a composition of transpositions. A permutation σ is called ***even*** if it is a composition of an even number of transpositions and ***odd*** otherwise. It can be shown that this distinction is well-defined. The ***sign*** of σ, is 1 if σ is odd and 0 if it is even. •

Let M be a manifold and let Λ^k denote the set of differential k-forms on M. Then $\Lambda = \oplus_{k=0}^{\dim(M)}\Lambda^k$ is the ***graded exterior algebra*** of differential forms.

2.25 Definition: The ***exterior product***, denoted by $\wedge$, is a map $\wedge : \Lambda^r(M) \times \Lambda^s(M) \to \Lambda(M)^{r+s}$ defined by

$$(\alpha\wedge\beta)(X_1,\dots,X_{r+s}) = \frac{(r+s)!}{r!s!}\sum_{\sigma\in S_{r+s}}(-1)^{\mathrm{sign}(\sigma)}\alpha(X_{\sigma(1)},\dots,X_{\sigma(r)})\beta(X_{\sigma(r+1)},\dots,X_{\sigma(r+s)}).$$

•

The exterior product has the following properties:

1. $\alpha \wedge (\beta + \gamma) = \alpha \wedge \beta + \alpha \wedge \gamma$;
2. $\alpha \wedge \beta = (-1)^{\deg(\alpha)\deg(\beta)}\beta \wedge \alpha$;
3. $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$.

2.26 Definition: Let r be an even integer. A ***derivation of degree r*** on $\Lambda(M)$ is a map $D : \Lambda(M) \to \Lambda(M)$ such that $D(\Lambda^k(M)) \subset \Lambda^{k+r}(M)$, and which satisfies the following properties:

(i) $D(a\alpha + b\beta) = aD(\alpha) + bD(\beta)$;
(ii) $D(\alpha \wedge \beta) = D(\alpha) \wedge \beta + \alpha \wedge D(\beta)$,

for $a, b \in \mathbb{R}$ and $\alpha, \beta \in \Lambda(M)$. •

2.27 Definition: Let r be an odd integer. An ***antiderivation of degree r*** on $\Lambda(M)$ is a map $D : \Lambda(M) \to \Lambda(M)$ such that $D(\Lambda^k(M)) \subset \Lambda^{k+r}(M)$, and which satisfies the following properties:

(i) $D(a\alpha + b\beta) = aD(\alpha) + bD(\beta)$;
(ii) $D(\alpha \wedge \beta) = D(\alpha) \wedge \beta + (-1)^{\deg(\alpha)}\alpha \wedge D(\beta)$. •

The set of all derivations and antiderivations $\mathcal{D}$, is a graded Lie algebra with commutator,

$$[D_1, D_2] = D_1 \circ D_2 - (-1)^{\deg(D_1)\deg(D_2)}D_2 \circ D_1,$$

where, $D_1, D_2 \in \mathcal{D}$.

Two basic antiderivations are the interior product and the exterior derivative.

2.28 Definition: The ***interior product with respect to a vector V***, $i_V : \Lambda^k(M) \to \Lambda^{k-1}(M)$ is an antiderivation of degree -1 which is defined by the following two properties:

(i) $i_V f = 0 \ \forall f \in \Lambda^0(M)$;
(ii) $i_V \alpha = \alpha(V) \ \forall \alpha \in \Lambda^1(M)$. •

2.29 Definition: The ***interior product with respect to a linear endomorphism J***, $i_J : \Lambda^k(M) \to \Lambda^k(M)$ is a derivation of degree 0 which is defined by the following two properties:

(i) $i_J f = 0 \ \forall f \in \Lambda^0(M)$;
(ii) $\langle i_J \alpha, V\rangle = \alpha(JV) \ \forall\ \alpha \in \Lambda^1(M)$. •

2.30 Definition: The ***exterior derivative*** $d : \Lambda^k(M) \to \Lambda^{k+1}(M)$ is the unique antiderivation of degree 1 which is defined by the following two properties:

(i) $dd\alpha = 0 \ \forall \ \alpha \in \Lambda(M)$;

(ii) $df = \frac{\partial f}{\partial x^i} dx^i \ \forall f \in \Lambda^0(M)$. •

The bracket of i_J and d forms a new antiderivation, the vertical derivative:

$$d_J := [i_J, d] = i_J d - d i_J,$$

which has the property, $d_J^2 = 0$.

The different posible brackets of the above defined antiderivations are

1. $[i_Z, d] = i_Z d + d i_Z = \mathcal{L}_Z$,
2. $[d, d_J] = d d_J + d_J d = 0$,
3. $[d_J, i_V] = d_J i_V + i_V d_J = i_J$,
4. $[i_X, i_J] = i_X i_J - i_J i_X = i_{JX}$.

2.6. Notation. We will use some non-standard notation. Let $\sigma : I \times J \to M$ be of class C^1. Then $\dot{\sigma}$ will denote a vector field in TM defined by,

$$\dot{\sigma}(s,t) = \frac{\mathrm{d}}{\mathrm{d}t}\sigma(s,t) \in T_{\sigma(s,t)}M.$$

Furthermore, if we define $\sigma_t(s)$ and $\sigma_s(t)$ to be curves parameterized by $s \in I$ and $t \in J$, respectively, then $\dot{\sigma}(s,t) = \dot{\sigma}_s(t) = \dot{\sigma}_t(s)$. That is, " $\dot{}$ " will always be used to denote a tangent vector with respect to the variable t.

3. Problem Statement Using Tangent Bundle Geometry

In preperation for giving a coordinate free derivation of the Euler–Lagrange equations, we must first reformulate the classical version of the calculus of variations minimization problem in terms of vector fields on TM. This reformulation is the goal of the present chapter. We therefore begin by giving a careful description of the minimization problem, followed by its reformulation in terms of vector fields on TM. Since our reformulation will require extending vector fields defined along integral curves to neighbourhoods in M, we also show that such extensions exist.

3.1. Classical Minimization Problem Statement. In this section we give a coordinate free statement of the minimization problem. This section parallels the discussion found in [Bullo and Lewis 2004].

Let M be a C^r-manifold, $r > 1$. Given a C^2-function $L : TM \to \mathbb{R}$ and a certain class of curves $\mathcal{D}$, the goal is to find a curve $\gamma_0 \in \mathcal{D}$ such that the action defined as,

$$A_L[\gamma] = \int_a^b L(\dot{\gamma}(t))\mathrm{d}t$$

is minimized with respect to all curves $\gamma \in \mathcal{D}$. If the class of curves is chosen to be $\mathcal{D}(C^2, x_a, x_b) := \{\gamma \mid \gamma \in C^2([a,b], M), \gamma(a) = x_a, \gamma(b) = x_b\}$ then the following useful definitions can be made.

3.1 Definition: Let $I \subset \mathbb{R}$ be an interval satisfying $0 \in \mathrm{int}(I)$. A ***C^2-variation*** of a C^2-curve γ is a C^2-map $\sigma : I \times [a, b] \to M$ which satisfies the following properties:

(i) $\sigma(s, a) = \gamma(a)$, $s \in I$;

(ii) $\sigma(s, b) = \gamma(b)$, $s \in I$;

(iii) $\sigma(0, t) = \gamma(t)$, $t \in [a, b]$.

Every C^2-variation σ has a corresponding C^1-***infinitesimal variation*** defined by

$$\delta\sigma(t) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \sigma(s, t). \qquad \bullet$$

From the definition of a C^2-variation we see that $\delta\sigma(a) = \delta\sigma(b) = 0$.

Let $\sigma_s : [a, b] \to M$ denote the curve satisfying $\sigma_s(t) = \sigma(s, t)$ for fixed s. If $\gamma \in \mathcal{D}(C^2, x_a, x_b)$, then all C^2-variations of γ give rise to curves σ_s, $s \in I$, which also belong to $\mathcal{D}(C^2, x_a, x_b)$. Also, given any two curves γ_1, $\gamma_2 \in \mathcal{D}(C^2, x_a, x_b)$, $\sigma(s, t) = \gamma_1(t)(1 - s) + \gamma_2(t)s$ defines a C^2-variation of γ_1 satisfying $\sigma(1, t) = \gamma_2(t)$. That is, given any two curves γ_1, $\gamma_2 \in \mathcal{D}(C^2, x_a, x_b)$ there exists a C^2-variation of γ_1 such that $\gamma_2 = \sigma_1$.

In order for γ_0 to minimize A_L, we must have

$$A_L[\gamma_0] \leq A_L[\gamma] \text{ for all } \gamma \in \mathcal{D}(C^2, x_a, x_b). \tag{3.1}$$

(3.1) will hold, if for all C^2-variations of γ_0, we have,

$$A_L[\gamma_0] \leq A_L[\sigma_s] \text{ for any fixed } s \in I.$$

Fixing the particular variation, $A_L[\sigma_s]$ becomes a C^2-function of s. From elementary calculus, if $\gamma_0 = \sigma_s|_{s=0}$ minimizes the function $A_L[\sigma_s] : I \to \mathbb{R}$ then

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} A_L[\sigma_s] = 0. \tag{3.2}$$

3.2 Remark: (3.2) is only a necessary condition for γ_0 to be a minimizer. A curve satisfying this condition is called stationary. $\bullet$

Since $\sigma_s \in \mathcal{D}(C^2, x_a, x_b)$ for all C^2-variations, if γ_0 minimizes with respect to all $\gamma \in \mathcal{D}(C^2, x_a, x_b)$, then (3.2) must hold for all C^2-variations of γ_0. Expanding the left hand side of (3.2),

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} A[\sigma(s,t)] &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \int_a^b L(\dot\sigma(s,t))\mathrm{d}t \\ &= \int_a^b \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} L(\dot\sigma(s,t))\mathrm{d}t \\ &= \int_a^b \langle dL, \tfrac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} \dot\sigma(s,t)\rangle \mathrm{d}t.\end{aligned}$$

The order of integration and differentiation in the second equality can be exchanged since $L \in C^2(TM, \mathbb{R})$.

Thus, if γ_0 is stationary with respect to all $\gamma \in \mathcal{D}(C^2, x_a, x_b)$, then the following equation must hold for all C^2-variations σ of γ_0:

$$\int_a^b \langle dL, \tfrac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} \dot\sigma(s,t)\rangle \mathrm{d}t = 0. \tag{3.3}$$

3.3 Remark: (3.3) is a local condition in the sense that it depends only on the values of $\sigma(s,t)$ and $\frac{d}{\mathrm{d}t}\sigma(s,t)$ in a neighbourhood of $s=0$. •

The next step is to write (3.3) in terms of vector fields defined on TM. This is done in the following proposition.

3.4 Proposition: *Let X and Z be vector fields defined in an open neighbourhood of* image(γ) *which satisfy*

$$\begin{aligned} X(\sigma(s,t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\sigma(s,t), \\ Z(\sigma(s,t)) &= \frac{\mathrm{d}}{\mathrm{d}s}\sigma(s,t). \end{aligned}$$

Then, $\int_a^b \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} L(\dot\sigma(s,t))\mathrm{d}t = 0$ *if and only if*

$$\int_a^b (\mathcal{L}_{Z^T}L)\dot\gamma(t)\mathrm{d}t = 0. \tag{3.4}$$

Proof: To prove Proposition 3.4 we will use the following lemma.

1 Lemma: *Let X and Z be as defined in Proposition 3.4, then* $\Phi_s^{Z^T}(\dot\sigma(s,t)) = X(\sigma(s,t))$.

Proof: Consider $Z^T(\dot\sigma(s,t))$:

$$\begin{aligned} Z^T(\dot\sigma(s,t)) &= \frac{\mathrm{d}}{\mathrm{d}\xi}\Big|_{\xi=0} T\Phi_\xi^Z(\dot\sigma(s,t)) \\ &= \frac{\mathrm{d}}{\mathrm{d}\xi}\Big|_{\xi=0} \frac{\mathrm{d}}{\mathrm{d}w}\Big|_{w=0} \Phi_\xi^Z \circ \Phi_w^X(\sigma(s,t)) \\ &= \frac{\mathrm{d}}{\mathrm{d}\xi}\Big|_{\xi=0} \frac{\mathrm{d}}{\mathrm{d}w}\Big|_{w=0} \Phi_\xi^Z(\sigma(s,t+w)) \\ &= \frac{\mathrm{d}}{\mathrm{d}\xi}\Big|_{\xi=0} \frac{\mathrm{d}}{\mathrm{d}w}\Big|_{w=0} \sigma(s+\xi,t+w) \\ &= \frac{\mathrm{d}}{\mathrm{d}\xi}\Big|_{\xi=0} \frac{\mathrm{d}}{\mathrm{d}t}\sigma(s+\xi,t) \\ &= \frac{\mathrm{d}}{\mathrm{d}\xi}\Big|_{\xi=0} \dot\sigma(s+\xi,t) = \frac{\mathrm{d}}{\mathrm{d}s}\dot\sigma(s,t), \end{aligned}$$

which implies

$$\Phi_s^{Z^T}(\dot\sigma(s,t)) = X(\sigma(s,t)). \quad ▼$$

Now, using the definition of the Lie derivative and Lemma 1 we have,

$$\begin{aligned} (\mathcal{L}_{Z^T}L)(\dot\gamma(t)) &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \Phi_s^{Z^T*}L(\dot\gamma(t)) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L\circ\Phi_s^{Z^T}(\dot\gamma(t)) \\ &= dL \circ \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} X(\sigma(s,t)) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L(\dot\sigma(s,t)). \end{aligned}$$

Integrating,

$$\int_a^b (\mathcal{L}_{Z^T}L)(\dot\gamma(t))\mathrm{d}t = \int_a^b \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L(\dot\sigma(s,t))\mathrm{d}t. \quad ■$$

Now consider the following coordinate calculation:

$$\begin{aligned}((\mathcal{L}_{Z^T}L)(\dot\gamma(t)))_\phi &= \frac{\partial L(\gamma(t),\gamma'(t))}{\partial x^i}Z^i(\gamma(t)) + \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\partial Z^i(\gamma(t))}{\partial x^k}\frac{\mathrm{d}\gamma^k(t)}{\mathrm{d}t} \\ &= \frac{\partial L(\gamma(t),\gamma'(t))}{\partial x^i}Z^i(\gamma(t)) + \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\mathrm{d}Z^i(\gamma(t))}{\mathrm{d}t}.\end{aligned}$$

This calculation depends only on how Z is defined along image(γ), that is, (3.4) depends only on how $\delta\sigma$ is defined. Therefore, $\gamma \in \mathcal{D}(C^2, x_a, x_b)$ is stationary with respect to all curves belonging to $\mathcal{D}(C^2, x_a, x_b)$ if and only if

$$\int_a^b (\mathcal{L}_{Z^T}L)\dot\gamma(t)\mathrm{d}t = 0$$

for all Z satisfying that there exists some C^2-variation such that $Z(\gamma(t)) = \delta\sigma(t)$ for all $t \in [a,b]$.

It only remains to show that we can in fact extend $\delta\sigma$ to a neighbourhood of image(γ).

3.2. Construction of Vector Fields on TM. Given any C^2-variation σ with corresponding vector fields $\dot\gamma$ and $\delta\sigma \in \Gamma^1(TM)$ defined along image(γ), we will show that there exist vector fields X, $Z \in \Gamma^1(TM)$ satisfying,

1. $X(\gamma(t)) = \dot\gamma(t)$,
2. $Z(\gamma(t)) = \delta\sigma(t)$,
3. X and Z are defined in a neighbourhood of image(γ).

3.5 Remark: To this point it is only evident that Z needs to be defined in a neighbourhood of image(γ). In Chapter 4, however, the derivation which we discuss will also require that the extension of $\dot\gamma$ exists. Therefore, we will explicitly construct extensions for both $\delta\sigma$ and $\dot\gamma$. •

We will present two different methods for constructing the vector fields X and Z. The first method we discuss uses the idea of the push-forward of a vector field. Our construction will use the following proposition.

3.6 Proposition: *Let M be a C^∞-Riemannian manifold. Let $\gamma : [a,b] \to M$ be a C^r-embedding, $r > 1$. Let $X_0 :$ image$(\gamma) \to \cup_{t\in[a,b]} T_{\gamma(t)}M$ be a C^k-vector field ($k \le r$) defined along* image(γ). *Then there exists a C^k-vector field X in an open neighbourhood of* image(γ) *such that $X_0(\gamma(t)) = X(\gamma(t))$ for all $t \in [a,b]$.*

Proof: (Proof by construction) We use the push-forward with respect to the exponential map to construct the vector field X. Let $W : \cup_{t\in[a,b]} T_{\gamma(t)}M \to \cup_{t\in[a,b]} T(T_{\gamma(t)}M)$ be the vertical vector field defined for all $t \in [a,b]$ given by $W(v_{\gamma(t)}) = \mathrm{vlft}_{v_{\gamma(t)}} X_0(\gamma(t))$. Note that, since the vertical lift is a pointwise construction, W is well-defined. Let image$(\gamma)^\perp$ denote the normal bundle to $T($image$(\gamma))$ in TM. Then for some open set $V \subset$ image$(\gamma)^\perp$,

$\exp_\gamma : V \to U$ defines a diffeomorphism from V to the tubular neighbourhood U. X is then defined in U as,

$$\begin{aligned} X(x) &= \exp_{\gamma *} W(x) = T\exp_\gamma \circ W \circ \exp_\gamma^{-1}(x) \\ &= T\exp_\gamma \circ W(u_{\gamma(t)}) = T\exp_\gamma(\mathrm{vlft}_{u_{\gamma(t)}} X_0(\gamma(t))) \end{aligned}$$

where, $u_{\gamma(t)} \in \mathrm{image}(\gamma)^\perp$ and $\exp_\gamma(u_{\gamma(t)}) = x$.

Since $\exp_\gamma : V \to U$ is a C^∞-diffeomorphism, both $\exp^{-1}$ and $T\exp \in C^\infty$. Therefore, since X is a composition of two C^∞-maps and one C^k-map, it is itself of class C^k.

From (2.2),

$$T_{0_{\gamma(t)}} \exp_\gamma(\mathrm{vlft}_{0_{\gamma(t)}} X_0(\gamma(t))) = X_0(\gamma(t)). \tag{3.5}$$

(3.5), combined with the fact that $\exp_\gamma^{-1}(\gamma(t)) = 0_{\gamma(t)}$ results in

$$X(\gamma(t)) = X_0(\gamma(t)). \qquad \blacksquare$$

Define the following maps $W_1 : \cup_{t\in[a,b]} T_{\gamma(t)}M \to \cup_{t\in[a,b]} T(T_{\gamma(t)}M)$ and $W_2 : \cup_{t\in[a,b]} T_{\gamma(t)}M \to \cup_{t\in[a,b]} T(T_{\gamma(t)}M)$ as, $W_1(v_{\gamma(t)}) = \mathrm{vlft}_{v_{\gamma(t)}}(\dot\gamma(t))$ and $W_2(v_{\gamma(t)}) = \mathrm{vlft}_{v_{\gamma(t)}}(\delta\sigma(t))$. Using Proposition 3.6, construct the following two vector fields on the tubular neighbourhood U,

$$X(x) = \exp_{\gamma *} W_1(x) \text{ and } Z(x) = \exp_{\gamma *} W_2(x).$$

To construct the appropriate vector fields on TM take the complete lifts of X and Z.

The second method we discuss uses partitions of unity. Our construction will use the following proposition.

3.7 Proposition: *Let M be a C^∞-manifold admitting C^∞-partitions of unity, $f : [a,b] \to M$ a C^r-embedding, $r > 1$. Then there exists a vector field $X \in \Gamma^{r-1}(TM)$ on a neighbourhood of $f([a,b])$ in M which satisfies $X(f(t)) = T_t f \cdot 1$ for all $t \in [a,b]$.*

Proof: Since $f([a,b])$ is a submanifold of M, there exist admissible charts (U_i, ϕ_i) covering $f([a,b])$ such that $\phi_i(U_i \cap \mathrm{image}(f)) = \phi_i(U_i) \cap (\mathbb{R} \times \{0\})$. Now, the coordinate representation of the tangent vector $T_t f \cdot 1$ in the chart $(TU_i, T\phi_i)$ is given by $((t, 0, \dots, 0), (1, 0, \dots, 0))$. This vector field agrees with the constant vector field e_1 (where $e_1 = (1, 0, 0..., 0)$) defined on the open neighbourhood $\phi_i(U_i)$. Since M admits a C^∞-partition of unity, there exists a set of pairs $\{(U_{i(j)}, g_{i(j)})\}$ such that $U_{i(j)} \subset U_i$ for all j, $\{U_{i(j)}\}$ is a locally finite cover at each $x \in \cup_i U_i$ and the $g_{i(j)} \in C^\infty(M, \mathbb{R})$ satisfy $\sum_{i(j)} g_{i(j)}(x) = 1$. Therefore, make the following definition:

$$X(x) = \sum_{i(j)} g_{i(j)} T_x \phi_{i(j)}^{-1} \circ e_1 \circ \phi_{i(j)}(x), \tag{3.6}$$

where $\phi_{i(j)}$ is the restriction of ϕ_i to $U_{i(j)}$. Now, since the sum in (3.6) is finite at any point $x \in \cup_i U_i$, X is well defined in $\cup_i U_i$.

We must also show that $X(f(t)) = T_t f \cdot 1$. Since $f(t) = x$, $x \in \cup_i U_i$ and $T_{f(t)}\phi_i(T_t f \cdot 1) = e_1$ we have,

$$T_{f(t)} \phi_{i(j)}^{-1} \circ e_1 \circ \phi_{i(j)}(f(t)) = T_t f \cdot 1$$

and, therefore, since $\sum_{i(j)} g_{i(j)}(x) = 1$,

$$X(f(t)) = \sum_{i(j)} g_{i(j)}(T_t f \cdot 1) = T_t f \cdot 1.$$

■

Choosing $\{(U_i, \phi_i)\}$ to be a set of submanifold charts for image(γ) and $\{(U_{i(j)}, g_{i(j)})\}$ a partition of unity subordinate to it, direct application of Proposition 3.7 allows us to construct the following vector field:

$$X(x) = \sum_{i(j)} g_{i(j)} T_x \phi_{i(j)}^{-1} \circ e_1 \circ \phi_{i(j)}(x).$$

Using the partition of unity that was just defined, let $W : \cup_{t\in[a,b]} T_{\gamma(t)}M \to \cup_{t\in[a,b]} T(T_{\gamma(t)}M)$ be the vertical vector field defined by,

$$W(v_{\gamma(t)}) = \mathrm{vlft}_{v_{\gamma(t)}} \delta\sigma(t).$$

Then $W_{\phi_{i(j)}}$ is a vector field defined at all points $(t, 0, \ldots, 0) \in \phi_{i(j)}(U_{i(j)})$, where $t \in [a, b]$. Let $(\exp_\gamma)_{\phi_{i(j)}}$ be the coordinate representation of $\exp_\gamma$. Using Proposition 3.6, define the following vector field in $\phi_{i(j)}(U_{i(j)})$,

$$Z_{i(j)}(\phi_{i(j)}(x)) = (\exp_\gamma)_{\phi_{i(j)}*} W_{\phi_{i(j)}}(\phi_{i(j)}(x)),$$

where $x \in U_{i(j)}$. Now use the partition of unity $\{(U_{i(j)}, g_{i(j)})\}$ to define the vector field Z in M. That is,

$$Z(x) = \sum_{i(j)} g_{i(j)} T_x \phi_{i(j)}^{-1} \circ Z_{i(j)} \circ \phi_{i(j)}(x).$$

To construct the appropriate vector fields on TM take the complete lifts of X and Z.

Before proceeding to the coordinate-free derivation of the Euler–Lagrange equations, we first show that the vector fields X and Z commute on the C^2-variation σ. This property will be used in the coordinate-free derivation.

3.8 Proposition: *Let X, $Z \in \Gamma^1(TM)$ be vector fields defined on open neighbourhoods of* image(γ) *satisfying,*

$$X(\sigma(s,t)) = \frac{\mathrm{d}}{\mathrm{d}t}\sigma(s,t),$$
$$Z(\sigma(s,t)) = \frac{\mathrm{d}}{\mathrm{d}s}\sigma(s,t).$$

Then $[X, Z](\sigma(s,t)) = 0$.

Proof: Consider the following,

$$T\sigma \circ \frac{\partial}{\partial s}(s,t) = T_1\sigma \circ ((s,1),t) = \frac{\mathrm{d}}{\mathrm{d}s}\sigma(s,t) = Z \circ \sigma(s,t).$$

Similar calculations for $T\sigma \circ \frac{\partial}{\partial t}(s,t)$ result in,

$$T\sigma \circ \frac{\partial}{\partial t}(s,t) = X \circ \sigma(s,t).$$

Therefore $\frac{\partial}{\partial t} \sim_\sigma X$ and $\frac{\partial}{\partial s} \sim_\sigma Z$. Theorem 2.15 implies that $[\frac{\partial}{\partial t}, \frac{\partial}{\partial s}] \sim_\sigma [X, Z]$, which in turn implies that $[X, Z]\sigma(s,t) = 0$. ■

4. Main Results

In this section we give characterizing properties of a minimizing curve γ on a manifold M. We begin by giving the coordinate free expression for the Euler–Lagrange equation and then proceed with its coordinate-free derivation. Next we characterize a minimizing curve in terms of its local behavior. In short we prove that a curve on a manifold is a minimizer if and only if it solves a corresponding minimization problem in all of its local coordinate charts.

4.1. Coordinate Free Derivation of Euler–Lagrange Equations.

4.1 Definition: Let $L : TM \to \mathbb{R}$ be a C^2-function and $\gamma : I \to M$ a curve of class C^2. Then,

$$E_L := i_{\ddot{\gamma}} d i_J dL + d i_V dL - dL$$

defines a one-form on TM. •

4.2 Theorem: *Let L and γ be as defined in Definition 4.1, then $\gamma \in \mathcal{D}(C^2, x_a, x_b)$ is stationary with respect to all $\tilde{\gamma} \in \mathcal{D}(C^2, x_a, x_b)$ if and only if*

$$E_L(\dot{\gamma}(t)) = 0 \text{ for all } t \in [a, b].$$

Proof: Referring to (3.2), γ is stationary if and only if for all C^2-variations of γ

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} A_L[\sigma_s] = 0.$$

Therefore, by (3.4) γ is stationary if and only if

$$\int_a^b (\mathcal{L}_{Z^T} L)(\dot{\gamma}(t))\mathrm{d}t = 0, \tag{4.1}$$

where, Z satisfies $Z(\gamma(t)) = \delta\sigma(t)$ for all $t \in [a, b]$.

The derivation found in [Nester 1988] uses the two identities

$$\mathrm{d}\omega(\tilde{X}, \tilde{Z}) = \mathcal{L}_{\tilde{X}}\langle\omega, \tilde{Z}\rangle - \mathcal{L}_{\tilde{Z}}\langle\omega, \tilde{X}\rangle - \langle\omega, [\tilde{X}, \tilde{Z}]\rangle \tag{4.2}$$

and

$$\mathcal{L}_{\tilde{X}}\langle\omega, \tilde{Z}\rangle = \langle\mathcal{L}_{\tilde{X}}\omega, \tilde{Z}\rangle + \langle\omega, \mathcal{L}_{\tilde{X}}\tilde{Z}\rangle \tag{4.3}$$

where, ω is a differentiable one-form and $\tilde{X}$ and $\tilde{Z}$ are vector fields on TTM. When the substitutions $\omega = i_J dL$, $\tilde{X} = X^T$ and $\tilde{Z} = Z^T$ are made, (4.2) and (4.3) involve the object $[X^T, Z^T]$. This necessitates that X^T and Z^T are of class C^1 or equivalently that γ is of class C^3. The following two lemmas will be used to preserve our original specification that γ is only required to be of class C^2.

4.3 Lemma: *Let $\tilde{X}, \tilde{Z} \in \Gamma^0(TTM)$ satisfying $T\pi(\tilde{X}), T\pi(\tilde{Z}) \in \Gamma^1(TM)$. Let $L : TM \to \mathbb{R}$ be a C^2-function. Then,*

$$\begin{aligned}
&\mathrm{d}(i_J dL)(\tilde{X}, \tilde{Z}) = \\
&\qquad \mathcal{L}_{\tilde{X}}\langle dL, \mathrm{vlft}(T\pi(\tilde{Z}))\rangle - \mathcal{L}_{\tilde{Z}}\langle dL, \mathrm{vlft}(T\pi(\tilde{X}))\rangle - \langle dL, \mathrm{vlft}[T\pi(\tilde{X}), T\pi(\tilde{Z})]\rangle. \quad (4.4)
\end{aligned}$$

Furthermore, if $\tilde{X}$ and $\tilde{Z}$ are chosen to be the complete lifts of arbitrary extensions of $\dot{\gamma}$ and $\delta\sigma$ then $(\mathrm{d}(i_J dL)(X^T, Z^T))(\dot{\gamma}(t))$ is independent of choice of extension for all $t \in [a, b]$.

Proof: In coordinates the value of the one form $i_J dL$ acting on any vector $\tilde{Y} \in \Gamma(TTM)$ is $\frac{\partial L}{\partial v^i} Y^i$ where $Y^i \frac{\partial}{\partial x^i}$ is the coordinate representative of the projection $T\pi(\tilde{Y})$. Therefore the coordinate representation of $i_J dL$ is $\frac{\partial L}{\partial v^i} dx^i$. Applying the definition of the exterior derivative in coordinates we have,

$$(\mathrm{d}(i_J dL))_\phi = \frac{\partial^2 L}{\partial x^j \partial v^i} dx^j \wedge dx^i + \frac{\partial^2 L}{\partial v^j \partial v^i} dv^j \wedge dx^i.$$

Letting $X^i \frac{\partial}{\partial x^i} + S^i \frac{\partial}{\partial v^i}$ and $Z^i \frac{\partial}{\partial x^i} + R^i \frac{\partial}{\partial v^i}$ denote the coordinate representations of $\tilde{X}$ and $\tilde{Z}$ respectively,

$$(\mathrm{d}(i_J dL)(\tilde{X}, \tilde{Z}))_\phi = \frac{\partial^2 L}{\partial x^j \partial v^i}(X^j Z^i - Z^j X^i) + \frac{\partial^2 L}{\partial v^j \partial v^i}(S^j Z^i - R^j X^i).$$

Now consider the coordinate expression of the first term on the right hand side of (4.4):

$$\begin{aligned}
\Big(\mathcal{L}_{\tilde{X}}\langle dL, \mathrm{vlft}(T\pi(\tilde{Z}))\rangle\Big)_\phi &= X^j \frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial v^i} Z^i\right) + S^j \frac{\partial}{\partial v^j}\left(\frac{\partial L}{\partial v^i} Z^i\right) \\
&= X^j \frac{\partial^2 L}{\partial x^j \partial v^i} Z^i + X^j \frac{\partial L}{\partial v^i}\frac{\partial Z^i}{\partial x^j} + S^j \frac{\partial^2 L}{\partial v^j \partial v^i} Z^i + S^j \frac{\partial L}{\partial v^i}\frac{\partial Z^i}{\partial v^j}.
\end{aligned}$$

A similar calculation for the second term results in

$$\Big(\mathcal{L}_{\tilde{Z}}\langle dL, \mathrm{vlft}(T\pi(\tilde{X}))\rangle\Big)_\phi = Z^j \frac{\partial^2 L}{\partial x^j \partial v^i} X^i + Z^j \frac{\partial L}{\partial v^i}\frac{\partial X^i}{\partial x^j} + R^j \frac{\partial^2 L}{\partial v^j \partial v^i} X^i + R^j \frac{\partial L}{\partial v^i}\frac{\partial X^i}{\partial v^j}$$

and for the final term,

$$\Big(\langle dL, \mathrm{vlft}[T\pi(\tilde{X}), T\pi(\tilde{Z})]\rangle\Big)_\phi = \frac{\partial L}{\partial v^i}\frac{\partial Z^i}{\partial x^j} X^j - \frac{\partial L}{\partial v^i}\frac{\partial X^i}{\partial x^j} Z^j.$$

Combining terms,

$$\begin{aligned}
&\Big(\mathcal{L}_{\tilde{X}}\langle dL, \mathrm{vlft}(T\pi(\tilde{Z}))\rangle - \mathcal{L}_{\tilde{Z}}\langle dL, \mathrm{vlft}(T\pi(\tilde{X}))\rangle - \langle dL, \mathrm{vlft}[T\pi(\tilde{X}), T\pi(\tilde{Z})]\rangle\Big)_\phi \\
&= X^j \frac{\partial^2 L}{\partial x^j \partial v^i} Z^i + S^j \frac{\partial^2 L}{\partial v^j \partial v^i} Z^i s - Z^j \frac{\partial^2 L}{\partial x^j \partial v^i} X^i - R^j \frac{\partial^2 L}{\partial v^j \partial v^i} X^i \\
&= \Big(\mathrm{d}(i_J dL)(\tilde{X}, \tilde{Z})\Big)_\phi.
\end{aligned}$$

To show the last part of Lemma 4.3, let $\tilde{X} = X^T$ and $\tilde{Z} = Z^T$ and evaluate (4.4) at $\dot{\gamma}(t)$. By the first part of Lemma 4.3 we have

$$\mathrm{d}(i_J dL)(X^T, Z^T) = \mathcal{L}_{X^T}\langle dL, \mathrm{vlft}(Z)\rangle - \mathcal{L}_{Z^T}\langle dL, \mathrm{vlft}(X)\rangle - \langle dL, \mathrm{vlft}[X, Z]\rangle. \quad (4.5)$$

Writing $\mathcal{L}_{Z^T}\langle dL, JX^T\rangle\dot{\gamma}(t)$ in coordinates,

$$\begin{aligned}
&(\mathcal{L}_{Z^T}\langle dL, JX^T\rangle\dot{\gamma}(t))_\varphi \\
&= \frac{\partial^2 L(\gamma(t),\gamma'(t))}{\partial x^j \partial v^i} X^i(\gamma(t)) Z^j(\gamma(t)) + \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\partial X^i(\gamma(t))}{\partial x^j} Z^j(\gamma(t)) \\
&\quad + \frac{\partial^2 L(\gamma(t),\gamma'(t))}{\partial v^j \partial v^i} X^i(\gamma(t)) \frac{\partial Z^j(\gamma(t))}{\partial x^k} v^k + \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\partial X^i(\gamma(t))}{\partial v^j}\frac{\partial Z^j(\gamma(t))}{\partial x^k} v^k.
\end{aligned} \tag{4.6}$$

Since X is a vector field on M, the last term in (4.6) is zero. The first term is obviously independent of extension off of image(γ). This leaves only the second and third term of (4.6) to consider. Writing the second term, we have,

$$\begin{aligned}
\frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\partial X^i(\gamma(t))}{\partial x^j} Z^j(\gamma(t)) &= \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\partial Z^i(\gamma(t))}{\partial x^j} X^j(\gamma(t)) \\
&= \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{d}{\mathrm{d}t}(Z^i(\gamma(t))),
\end{aligned} \tag{4.7}$$

where the first step uses the fact that $[X,Z](\gamma(t)) = 0$ and the second step uses the fact that $X^j(\gamma(t)) = v^j$. Since all extensions of Z must agree along γ, (4.7) is independent of extension.

Similarly for the third term of (4.6),

$$\frac{\partial^2 L(\gamma(t),\gamma'(t))}{\partial v^j \partial v^i} X^i(\gamma(t)) \frac{\partial Z^j(\gamma(t))}{\partial x^k} v^k = \frac{\partial^2 L(\gamma(t),\gamma'(t))}{\partial v^j \partial v^i} X^i(\gamma(t)) \frac{d}{\mathrm{d}t}(Z^i(\gamma(t))),$$

which is also independent of extension.

Now consider the second term in (4.5). Writing $\mathcal{L}_{X^T}\langle dL, JZ^T\rangle\dot{\gamma}(t)$ in coordinates,

$$\begin{aligned}
&(\mathcal{L}_{X^T}\langle dL, JZ^T\rangle\dot{\gamma}(t))_\phi \\
&= \frac{\partial^2 L(\gamma(t),\gamma'(t))}{\partial x^j \partial v^i} Z^i(\gamma(t)) X^j(\gamma(t)) + \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\partial Z^i(\gamma(t))}{\partial x^j} X^j(\gamma(t)) \\
&\quad + \frac{\partial^2 L(\gamma(t),\gamma'(t))}{\partial v^j \partial v^i} Z^i(\gamma(t)) \frac{\partial X^j(\gamma(t))}{\partial x^k} v^k + \frac{\partial L(\gamma(t),\gamma'(t))}{\partial v^i}\frac{\partial Z^i(\gamma(t))}{\partial v^j}\frac{\partial X^j(\gamma(t))}{\partial x^k} v^k.
\end{aligned}$$

Using similar arguments, we have that $\mathcal{L}_{X^T}\langle dL, JZ^T\rangle\dot{\gamma}(t)$ is also independent of vector field extension. The third term in (4.5), $\langle dL, \mathrm{vlft}[X,Z]\rangle(\dot{\gamma}(t))$, is independent of extension since all extensions of X and Z satisfies, $[X,Z](\gamma(t)) = 0$. Combining these results leads to the desired conclusion that $\langle i_{X^T} d i_J dL, Z^T\rangle(\dot{\gamma}(t))$ is independent of vector field extension. ■

4.4 Lemma: *Let X, $Z \in \Gamma^1(TM)$. Let $\tilde{X} \in \Gamma(TTM)$ be any vector field satisfying $T\pi(\tilde{X}) = X$. Then for any $v_x \in TM$ satisfying $[X,Z](x) = 0$ and $X(x) = v_x$, we have,*

$$(\mathcal{L}_{Z^T}\langle dL, V - J\tilde{X}\rangle)(v_x) = (\langle \mathcal{L}_{Z^T} dL, V - J\tilde{X}\rangle)(v_x) + (\langle dL, \mathrm{vlft}[Z,X]\rangle)(v_x). \tag{4.8}$$

Proof: We begin by writing the coordinate expression for each term in (4.8). Let $X^i \frac{\partial}{\partial x^i} + R^i \frac{\partial}{\partial v^i}$ be the coordinate expression for Z^T. Then

$$\Big((\mathcal{L}_{Z^T}\langle dL, V - J\tilde{X}\rangle)(v_x)\Big)_\phi = \frac{\partial^2 L(x,v)}{\partial x^j \partial v^i}(v^i - X^i(x))Z^j(x) + \frac{\partial^2 L(x,v)}{\partial v^j \partial v^i}(v^i - X^i(x))R^j(v_x)$$
$$+ \frac{\partial L(x,v)}{\partial v^i}\frac{\partial}{\partial x^j}(v^i - X^i(x))Z^j(x) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial}{\partial v^j}(v^i - X^i(x))R^j(v_x),$$
$$\Big((\langle di_{Z^T}dL, V - J\tilde{X}\rangle)(v_x)\Big)_\phi = \frac{\partial^2 L(x,v)}{\partial x^i \partial v^j}(v^j - X^j(x))Z^i(x) + \frac{\partial^2 L(x,v)}{\partial v^j \partial v^i}(v^j - X^j(x))R^i(v_x)$$
$$+ \frac{\partial L(x,v)}{\partial x^i}\frac{\partial Z^i(x)}{\partial v^j}(v^j - X^j(x)) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial R^i(v_x)}{\partial v^j}(v^j - X^j(x)),$$
$$((\langle dL, \mathrm{vlft}[Z,X]\rangle)(v_x))_\phi = \frac{\partial L(x,v)}{\partial v^i}\frac{\partial X^i(x)}{\partial x^j}Z^j(x) - \frac{\partial L(x,v)}{\partial v^i}\frac{\partial Z^i(x)}{\partial x^j}X^j(x).$$

From these expressions we see that (4.8) is equivalent to requiring

$$\frac{\partial L(x,v)}{\partial v^i}\frac{\partial}{\partial x^j}(v^i - X^i(x))Z^j(x) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial}{\partial v^j}(v^i - X^i(x))R^j(v_x)$$
$$= \frac{\partial L(x,v)}{\partial x^i}\frac{\partial Z^i(x)}{\partial v^j}(v^j - X^j(x)) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial R^i(v_x)}{\partial v^j}(v^j - X^j(x))$$
$$-\frac{\partial L(x,v)}{\partial v^i}\frac{\partial Z^i(x)}{\partial x^j}X^j(x) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial X^i(x)}{\partial x^j}Z^j(x),$$

which simplifies to

$$-\frac{\partial L(x,v)}{\partial v^i}\frac{\partial X^i(x)}{\partial x^j}Z^j(x) + \frac{\partial L(x,v)}{\partial v^i}R^i(v_x) = \frac{\partial L(x,v)}{\partial x^i}\frac{\partial Z^i(x)}{\partial v^j}(v^j - X^j(x))$$
$$+\frac{\partial L(x,v)}{\partial v^i}\frac{\partial R^i(v_x)}{\partial v^j}(v^j - X^j(x)) - \frac{\partial L(x,v)}{\partial v^i}\frac{\partial Z^i(v_x)}{\partial x^j}X^j(v_x) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial X^i(x)}{\partial x^j}Z^j(x).$$

Therefore, by the definition of complete lift we have

$$-\frac{\partial L(x,v)}{\partial v^i}\frac{\partial X^i(x)}{\partial x^j}Z^j(x) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial Z^i(x)}{\partial x^k}v^k = \frac{\partial L(x,v)}{\partial x^i}\frac{\partial Z^i(x)}{\partial v^j}(v^j - X^j(x))$$
$$+\frac{\partial L(x,v)}{\partial v^i}\frac{\partial R^i(v_x)}{\partial v^j}(v^j - X^j(x)) - \frac{\partial L(x,v)}{\partial v^i}\frac{\partial Z^i(x)}{\partial x^j}X^j(x) + \frac{\partial L(x,v)}{\partial v^i}\frac{\partial X^i(x)}{\partial x^j}Z^j(x).$$

Now, since $v_x = X(x)$,

$$0 = 2\frac{\partial L(x,v)}{\partial v^i}\left(-\frac{\partial Z^i(x)}{\partial x^j}X^j(x) + \frac{\partial X^i(x)}{\partial x^j}Z^j(x)\right),$$

which holds since $[X,Z](x) = 0$. ■

By Lemma 4.4 we have,

$$(\mathcal{L}_{Z^T}\langle dL, V - J(X^T)\rangle)(\dot\gamma(t)) = \langle(\mathcal{L}_{Z^T}dL, V - J(X^T)\rangle)(\dot\gamma(t)) + (\langle dL, \mathrm{vlft}[Z,X]\rangle)(\dot\gamma(t)). \quad (4.9)$$

Now, using (4.5) and (4.9), the integrand in (4.1) can be written as

$$\begin{aligned}
(\mathcal{L}_{Z^T}L)(\dot\gamma(t)) &= (\mathcal{L}_{Z^T}\langle dL, V\rangle)(\dot\gamma(t)) - (\langle di_V dL - dL, Z^T\rangle)(\dot\gamma(t)) \\
&= (\mathcal{L}_{Z^T}\langle dL, V\rangle)(\dot\gamma(t)) - (\langle E_L, Z^T\rangle)(\dot\gamma(t)) + (\langle i_{X^T} di_J dL, Z^T\rangle)(\dot\gamma(t)) \\
&= (\mathcal{L}_{Z^T}\langle dL, V\rangle)(\dot\gamma(t)) - (\langle E_L, Z^T\rangle)(\dot\gamma(t)) + (\mathcal{L}_{X^T}\langle dL, \mathrm{vlft}(Z)\rangle)(\dot\gamma(t)) \\
&\quad -(\mathcal{L}_{Z^T}\langle dL, \mathrm{vlft}(X)\rangle)(\dot\gamma(t)) - (\langle dL, \mathrm{vlft}([X,Z])\rangle)(\dot\gamma(t)) \\
&= -(\langle E_L, Z^T\rangle)(\dot\gamma(t)) + (\mathcal{L}_{X^T}\langle dL, \mathrm{vlft}(Z)\rangle)(\dot\gamma(t)) \\
&\quad +(\mathcal{L}_{Z^T}\langle dL, V - \mathrm{vlft}(X)\rangle)(\dot\gamma(t)) - (\langle dL, \mathrm{vlft}([X,Z])\rangle)(\dot\gamma(t)) \\
&= -(\langle E_L, Z^T\rangle)(\dot\gamma(t)) + (\mathcal{L}_{X^T}\langle dL, \mathrm{vlft}(Z)\rangle)(\dot\gamma(t)) \\
&\quad +(\langle L_{Z^T} dL, V - \mathrm{vlft}(X)\rangle)(\dot\gamma(t)) + (\langle dL, \mathrm{vlft}([Z,X])\rangle)(\dot\gamma(t)) \\
&\quad -(\langle dL, \mathrm{vlft}([X,Z])\rangle)(\dot\gamma(t)) \\
&= -(\langle E_L, Z^T\rangle)(\dot\gamma(t)) + (\mathcal{L}_{X^T}\langle dL, \mathrm{vlft}(Z)\rangle)(\dot\gamma(t)). \qquad (4.10)
\end{aligned}$$

Using the flow interpretation of the Lie derivative the second term in (4.10) becomes

$$(\mathcal{L}_{X^T}\langle dL, JZ^T\rangle)(\dot\gamma(t)) = (\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^{X^T *}\langle dL, JZ^T\rangle)(\dot\gamma(t)),$$

which, due to the condition $JZ^T\big|_{\sigma(a,s)} = JZ^T\big|_{\sigma(a,s)} = 0$, integrates to zero. This leads to the equation

$$\int_a^b (\mathcal{L}_{Z^T}L)(\dot\gamma(t))\mathrm{d}t = -\int_a^b \langle E_L, Z^T\rangle(\dot\gamma(t))\mathrm{d}t.$$

Therefore $E_L(\dot\gamma(t)) = 0$ is a sufficient condition for γ to be stationary. It only remains to show that $E_L(\dot\gamma(t)) = 0$ is also a necessary condition. This is done in [Nester 1988] by showing that, for all $t \in [a,b]$, the value of $E_L(\dot\gamma(t))$ acting on any vertical vector is zero:

4.5 Lemma: $i_J E_L(\dot\gamma(t)) = 0$.

Proof: First note that $i_J dL = d_J L$ and $E_L = i_{X^T} di_J dL + di_V dL - dL$. Therefore, using the relations from Section 2.5, we have,

$$\begin{aligned}
i_J i_{X^T} di_J dL &= [i_J, i_X] di_J dL + i_{X^T} i_J di_J dL \\
&= -i_{JX^T} dd_J L + i_{X^T} d_J^2 L + i_{X^T} di_J d_J L = -i_{JX^T} di_J dL
\end{aligned}$$

and

$$i_J di_V dL - i_J dL = d_J i_V dL - d_J L = [d_J, i_V]dL - d_J L - i_V d_J dL = -i_V d_J dL = i_V dd_J L.$$

Therefore, $i_J E_L = i_{V - JX^T} di_J dL$ and since $V(\dot\gamma(t)) = JX^T(\dot\gamma(t))$ for all $t \in [a,b]$ we have,

$$i_J E_L(\dot\gamma(t)) = 0$$

■

Therefore, $E_L(\dot\gamma(t)) = 0$ for all $t \in [a,b]$ is a necessary and sufficient condition for γ to be a stationary curve. ■

4.2. Local Minimizers from Global Minimizers. The object of this section is to prove that, if γ_0 is a solution to the minimization problem on some time interval $[a,b]$, then it also solves a minimization problem on any subinterval of $[a,b]$. What follows uses techniques found in [Bullo and Lewis 2004] and [Troutman 1996].

4.6 Theorem: *Let $L : TM \to \mathbb{R}$ be a C^0-function. Suppose $a < \tilde{a} < \tilde{b} < b$. If $\gamma_0 \in \mathcal{D}(C^2, x_a, x_b)$ minimizes $A_L[\gamma]$ with respect to all $\gamma \in \mathcal{D}(C^2, x_a, x_b)$, then γ_0 minimizes $\tilde{A}_L[\gamma]$ with respect to all $\gamma \in \mathcal{D}(C^2, \gamma_0(\tilde{a}), \gamma_0(\tilde{b}))$. Where $\tilde{A}_L[\gamma]$ is the restriction of $A_L[\gamma]$ to the interval $[\tilde{a}, \tilde{b}]$ and $\mathcal{D}(C^2, \gamma_0(\tilde{a}), \gamma_0(\tilde{b})) = \{\gamma \in C^2([\tilde{a}, \tilde{b}]) \mid \gamma(\tilde{a}) = \gamma_0(\tilde{a}), \gamma(\tilde{b}) = \gamma_0(\tilde{b})\}$.*

Proof: We prove the contra-positive. Suppose that γ_0 does not minimize $\tilde{A}_L[\gamma]$ with respect to all $\gamma \in \mathcal{D}(C^2, \gamma_0(\tilde{a}), \gamma_0(\tilde{b}))$. Then there exists a curve $\tilde{\gamma} \in \mathcal{D}(C^2, \gamma_0(\tilde{a}), \gamma_0(\tilde{b}))$ satisfying,

$$\tilde{A}_L[\gamma_0] - \tilde{A}_L[\tilde{\gamma}] = \delta \tag{4.11}$$

for some $\delta > 0$.

Now, if the curve

$$\gamma_*(t) = \begin{cases} \gamma_0(t) \text{ for all } t \in [a, \tilde{a}] \\ \tilde{\gamma}(t) \text{ for all } t \in (\tilde{a}, \tilde{b}) \\ \gamma_0(t) \text{ for all } t \in [\tilde{b}, b] \end{cases} \tag{4.12}$$

is of class C^2, then $A_L[\gamma_*] < A_L[\gamma_0]$ and then certainly γ_0 is not a minimizer. Unfortunately, it is very likely that γ_* is not of class C^2. In the next lemma we show that there exists a curve $\hat{\gamma}$ satisfying

$$\left|\tilde{A}_L[\tilde{\gamma}] - \tilde{A}_L[\hat{\gamma}]\right| = \frac{1}{2}\delta$$

and whose concatenation

$$\hat{\gamma}_0(t) = \begin{cases} \gamma_0(t) \text{ for all } t \in [a, \tilde{a}] \\ \hat{\gamma}(t) \text{ for all } t \in (\tilde{a}, \tilde{b}) \\ \gamma_0(t) \text{ for all } t \in [\tilde{b}, b] \end{cases} \tag{4.13}$$

is of class C^2.

4.7 Lemma: *For any $\epsilon > 0$, sufficiently small, there exists a curve $\hat{\gamma} \in \mathcal{D}(C^2, \gamma_0(\tilde{a}), \gamma_0(\tilde{b}))$ such that $\hat{\gamma}(t) = \tilde{\gamma}(t)$ for all $t \in (\tilde{a}+\epsilon, \tilde{b}-\epsilon)$, $\mid \tilde{A}_L[\tilde{\gamma}] - \tilde{A}_L[\hat{\gamma}] \mid = \frac{1}{2}\delta$ (where δ satisfies (4.11)) and the concatenated curve in (4.13) is of class C^2.*

Proof: By hypothesis, $\hat{\gamma}(t)$ is specified for all t except for $t \in [\tilde{a}, \tilde{a}+\epsilon]$ and $t \in [\tilde{b}-\epsilon, \tilde{b}]$. The construction of $\hat{\gamma}(t)$ is similar on both time intervals so we will only show the construction for $t \in [\tilde{a}, \tilde{a}+\epsilon]$. Let (U_1, ϕ_1) be an admissible chart with $\tilde{\gamma}(\tilde{a}) \in U_1$. Choose $\epsilon > 0$ such that $\tilde{\gamma}(\tilde{a}+\epsilon) \in U_1$. We will construct $\hat{\gamma}(t)$ such that, in the local coordinate chart, the following properties are satisfied:

1. $\hat{\gamma}_{\phi_1}(\tilde{a}) = \gamma_{\phi_1}(\tilde{a})$;
2. $\dot{\hat{\gamma}}_{\phi_1}(\tilde{a}) = \dot{\gamma}_{\phi_1}(\tilde{a})$;
3. $\ddot{\hat{\gamma}}_{\phi_1}(\tilde{a}) = \ddot{\gamma}_{\phi_1}(\tilde{a})$;
4. $\hat{\gamma}_{\phi_1}(\tilde{a}+\epsilon) = \tilde{\gamma}_{\phi_1}(\tilde{a}+\epsilon)$;
5. $\dot{\hat{\gamma}}_{\phi_1}(\tilde{a}+\epsilon) = \dot{\tilde{\gamma}}_{\phi_1}(\tilde{a}+\epsilon)$;
6. $\ddot{\hat{\gamma}}_{\phi_1}(\tilde{a}+\epsilon) = \ddot{\tilde{\gamma}}_{\phi_1}(\tilde{a}+\epsilon)$.

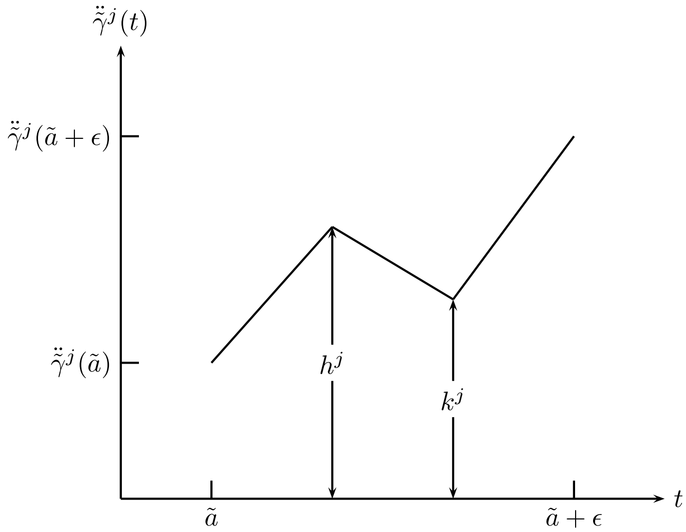


Figure 1. Proposed choice for a continuous $\ddot{\hat{\gamma}}$ on the time interval $t \in [\tilde{a}, \tilde{a}+\epsilon]$.

Let $\hat{\gamma}^j(t)$ denote the jth coordinate of $\hat{\gamma}_{\phi_1}(t)$. We begin by defining $\ddot{\hat{\gamma}}^j(t)$ for $t \in [\tilde{a}, \tilde{a}+\epsilon]$. Choose $\ddot{\hat{\gamma}}^j(t)$ to be the continuous sawtooth function depicted in Figure 1. As shown in Figure 1, $\ddot{\hat{\gamma}}^j(\tilde{a})$ and $\ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon)$ are chosen to satisfy Properties 3 and 4 and k^j and h^j are to be determined. The coordinate expression for $\ddot{\hat{\gamma}}^j(t)$ is given by,

$$\ddot{\hat{\gamma}}^j(t) = \begin{cases} \frac{3(k^j - \ddot{\gamma}^j(\tilde{a}))}{\epsilon} t + \ddot{\gamma}^j(\tilde{a}) - \frac{3(k^j - \ddot{\gamma}^j(\tilde{a}))\tilde{a}}{\epsilon}, & t \in [\tilde{a}, \tilde{a} + \frac{1}{3}\epsilon), \\ \frac{3(h^j - k^j)}{\epsilon} t + 2k^j - h^j + \frac{3\tilde{a}}{\epsilon}(k^j - h^j), & t \in [\tilde{a} + \frac{1}{3}\epsilon, \tilde{a} + \frac{2}{3}\epsilon], \\ \frac{3(-h^j + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon))}{\epsilon} t + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon) - \frac{3(-h^j + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon))(\tilde{a}+\epsilon)}{\epsilon}, & t \in (\tilde{a} + \frac{2}{3}\epsilon, \tilde{a}+\epsilon]. \end{cases} \tag{4.14}$$

To find the coordinate expressions for $\dot{\hat{\gamma}}^j(t)$, and $\hat{\gamma}^j(t)$ we integrate (4.14). This yields the following equations,

$$\dot{\hat{\gamma}}^j(t) = \begin{cases} \frac{3(k^j - \ddot{\gamma}^j(\tilde{a}))}{2\epsilon} t^2 + \ddot{\gamma}^j(\tilde{a})t - \frac{3(k^j - \ddot{\gamma}^j(\tilde{a}))\tilde{a}}{\epsilon} t + B^j, & t \in [\tilde{a}, \tilde{a} + \frac{1}{3}\epsilon), \\ \frac{3(h^j - k^j)}{2\epsilon} t^2 + 2k^j t - h^j t + \frac{3\tilde{a}}{\epsilon}(k^j - h^j)t + B^j, & t \in [\tilde{a} + \frac{1}{3}\epsilon, \tilde{a} + \frac{2}{3}\epsilon], \\ \frac{3(-h^j + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon))}{2\epsilon} t^2 + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon)t - \frac{3(-h^j + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon))(\tilde{a}+\epsilon)}{\epsilon} t + B^j, & t \in (\tilde{a} + \frac{2}{3}\epsilon, \tilde{a}+\epsilon], \end{cases}$$

$$\hat{\gamma}^j(t) = \begin{cases} \frac{(k^j - \ddot{\gamma}^j(\tilde{a}))}{2\epsilon} t^3 + \frac{1}{2}\ddot{\gamma}^j(\tilde{a})t^2 - \frac{3(k^j - \ddot{\gamma}^j(\tilde{a}))\tilde{a}}{2\epsilon} t^2 + B^j t + C^j, & t \in [\tilde{a}, \tilde{a} + \frac{1}{3}\epsilon), \\ \frac{(h^j - k^j)}{2\epsilon} t^3 + k^j t^2 - \frac{1}{2}h^j t^2 + \frac{3\tilde{a}}{2\epsilon}(k^j - h^j)t^2 + B^j t + C^j, & t \in [\tilde{a} + \frac{1}{3}\epsilon, \tilde{a} + \frac{2}{3}\epsilon], \\ \frac{(-h^j + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon))}{2\epsilon} t^3 + \frac{1}{2}\ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon)t^2 - \frac{3(-h^j + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon))(\tilde{a}+\epsilon)}{2\epsilon} t^2 + B^j t + C^j, & t \in (\tilde{a} + \frac{2}{3}\epsilon, \tilde{a}+\epsilon]. \end{cases}$$

B^j, C^j, h^j and k^j are chosen to satisfy Properties 1–5. A tedious computation gives

$$h^j = \frac{1}{3(\tilde{a}^2 + 2\tilde{a}\epsilon + \epsilon^2)} (\ddot{\gamma}^j(\tilde{a})\tilde{a}^2 + 4\dot{\hat{\gamma}}^j(\tilde{a}+\epsilon)\tilde{a} + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon)\tilde{a}^2 + \ddot{\hat{\gamma}}^j(\tilde{a}+\epsilon)\tilde{a}\epsilon + 6\gamma(\tilde{a})^j$$
$$- 6\tilde{\gamma}^j(\tilde{a}+\epsilon) - 4\dot{\gamma}^j(\tilde{a})\tilde{a} + 6\dot{\hat{\gamma}}^j(\tilde{a}+\epsilon)\epsilon)$$

and

$$k^j = \frac{1}{3\tilde{a}^2}(2\ddot{\gamma}^j(\tilde{a})\tilde{a}^2 + 2\ddot{\gamma}^j(\tilde{a})\tilde{a}\epsilon - 4\dot{\tilde{\gamma}}^j(\tilde{a}+\epsilon)\tilde{a} + \dddot{\tilde{\gamma}}^j(\tilde{a}+\epsilon)\tilde{a}^2 + 2\dddot{\tilde{\gamma}}^j(\tilde{a}+\epsilon)\epsilon\tilde{a} - 6\gamma(\tilde{a})^j$$
$$+ 6\tilde{\gamma}^j(\tilde{a}+\epsilon) + 4\dot{\gamma}^j(\tilde{a})\tilde{a} - 4\dot{\tilde{\gamma}}^j(\tilde{a}+\epsilon)\epsilon + \dddot{\tilde{\gamma}}^j(\tilde{a}+\epsilon)\epsilon^2 - 2\dot{\gamma}^j(\tilde{a})\epsilon).$$

Provided $0 < \epsilon < 1$ and letting

$$m^j = \max\{\left|\gamma^j(t)\right|, \left|\dot{\gamma}^j(t)\right|, \left|\ddot{\gamma}^j(t)\right|, \left|\tilde{\gamma}^j(t)\right|, \left|\dot{\tilde{\gamma}}^j(t)\right|, \left|\dddot{\tilde{\gamma}}^j(t)\right|\} \tag{4.15}$$

we have

$$\left|h^j\right| \leq \frac{1}{3\tilde{a}^2}(3m^j\tilde{a}^2 + 9m^j\left|\tilde{a}\right| + 18m^j)$$

and

$$\left|k^j\right| \leq \frac{1}{3\tilde{a}^2}(3m^j\tilde{a}^2 + 12m^j\left|\tilde{a}\right| + 19m^j).$$

Therefore, there exists a positive constant μ^j such that

$$\left|h^j\right|, \left|k^j\right| \leq \mu^j \tag{4.16}$$

Referring to Figure 1 we see that (4.15) and (4.16) imply

$$\left|\dddot{\tilde{\gamma}}^j\right| \leq 2m^j + 2\mu^j,$$

and therefore,

$$\begin{aligned} \left|\dot{\tilde{\gamma}}^j(t)\right| &= \left|\int_{\tilde{a}}^t \dddot{\tilde{\gamma}}^j(\tau)\mathrm{d}\tau + \dot{\gamma}^j(\tilde{a})\right| \leq \int_{\tilde{a}}^t \left|\dddot{\tilde{\gamma}}^j(\tau)\right|\mathrm{d}\tau + m^j \\ &\leq 2(m^j+\mu^j)(t-\tilde{a}) + m^j \leq 2(m^j+\mu^j)(\epsilon) + m^j \leq 3m^j + 2\mu^j. \end{aligned}$$

Also,

$$\begin{aligned} \left|\hat{\gamma}^j(t) - \tilde{\gamma}^j(t)\right| &= \left|\int_{\tilde{a}}^t (\dot{\hat{\gamma}}^j(\tau) - \dot{\tilde{\gamma}}^j(\tau))\mathrm{d}\tau + \hat{\gamma}^j(\tilde{a}) - \tilde{\gamma}^j(\tilde{a})\right| \leq \int_{\tilde{a}}^t \left|\dot{\hat{\gamma}}^j(\tau) - \dot{\tilde{\gamma}}^j(\tau)\right|\mathrm{d}\tau \\ &\leq \int_{\tilde{a}}^t \left|\dot{\hat{\gamma}}^j(\tau)\right| + \left|\dot{\tilde{\gamma}}^j(\tau)\right|\mathrm{d}\tau \leq \int_{\tilde{a}}^{\tilde{a}+\epsilon} \left|\dot{\hat{\gamma}}^j(\tau)\right| + \left|\dot{\tilde{\gamma}}^j(\tau)\right|\mathrm{d}\tau \leq (4m^j + 2\mu^j)\epsilon. \end{aligned}$$

Therefore, by appropriate choice of ϵ, we can construct a curve $\hat{\gamma}_\phi(t)$ which is arbitrarily close to $\tilde{\gamma}_\phi(t)$ for $t \in (\tilde{a}, \tilde{a}+\epsilon)$ and satisfies Properties 1–6. Applying ϕ_1^{-1} to $\hat{\gamma}_\phi(t)$ gives the desired curve in M.

In order to show that $\mid A_L[\hat{\gamma}] - A_L[\tilde{\gamma}] \mid \leq \frac{1}{2}\delta$, we first show that there exists a $\beta_a > 0$, independent of ϵ, satisfying

$$\left|L(\dot{\hat{\gamma}}(t))\right| < \beta_a \ \forall \ \ t \in [\tilde{a}, \tilde{a}+\epsilon].$$

To do this, first choose $0 < \epsilon_a < 1$ such that if $\hat{\gamma}^j(t) \in [\tilde{\gamma}^j(t) - (4m^j+2\mu^j)\epsilon_a, \tilde{\gamma}^j(t) + (4m^j+2\mu^j)\epsilon_a]$ for all $t \in [\tilde{a}, \tilde{a}+\epsilon_a]$ then $\hat{\gamma}(t) \in \phi_1(U_1)$ for all $t \in [\tilde{a}, \tilde{a}+\epsilon_a]$. For any $0 < \epsilon < \epsilon_a$, $\hat{\gamma}^j(t) \in [\tilde{\gamma}^j(t) - (4m^j+2\mu^j)\epsilon_a, \tilde{\gamma}^j(t) + (4m^j+2\mu^j)\epsilon_a]$, and $\dot{\hat{\gamma}}^j(t) \in [-(3m^j+2\mu^j), (3m^j+2\mu^j)]$. Since these intervals are compact there exists a constant β_a such that $\left|L(\dot{\hat{\gamma}}(t))\right| \leq \beta_a$ for all

$t \in [\tilde{a}, \tilde{a}+\epsilon_a]$. Using similar arguments we can show that there exist analogous constants, ϵ_b and β_b, for the time interval $[\tilde{b}-\epsilon, \tilde{b}]$. Let $\beta = \max\{\beta_a, \beta_b\}$ and choose $\epsilon < \max\{\epsilon_a, \epsilon_b\}$.

With the above constructions we have

$$\begin{aligned}
\left|\tilde{A}_L[\hat{\gamma}(t)] - \tilde{A}_L[\tilde{\gamma}(t)]\right| &= \left|\int_{\tilde{a}}^{\tilde{b}} L(\dot{\hat{\gamma}}(t))\mathrm{d}t - \int_{\tilde{a}}^{\tilde{b}} L(\dot{\tilde{\gamma}}(t))\mathrm{d}t\right| \\
&= \left|\int_{\tilde{a}}^{\tilde{b}} (L(\dot{\hat{\gamma}}(t)) - L(\dot{\tilde{\gamma}}(t)))\mathrm{d}t\right| \\
&= \left|\int_{\tilde{a}}^{\tilde{a}+\epsilon} (L(\dot{\hat{\gamma}}(t)) - L(\dot{\tilde{\gamma}}(t)))\mathrm{d}t + \int_{\tilde{b}-\epsilon}^{\tilde{b}} (L(\dot{\hat{\gamma}}(t)) - L(\dot{\tilde{\gamma}}(t)))\mathrm{d}t\right|.
\end{aligned}$$

Now since $L : TM \to \mathbb{R}$ is a continuous function with respect to elements in TM and $\tilde{\gamma} \in C^2([\tilde{a}, \tilde{b}], M)$ the composition $L\circ\dot{\tilde{\gamma}}$ belongs to $C([\tilde{a}, \tilde{b}], \mathbb{R})$. Since $[\tilde{a}, \tilde{b}]$ is a compact interval, this implies that $L \circ \dot{\tilde{\gamma}}$ achieves its maximum and minimum values. Let $M = \max\left|L \circ \dot{\tilde{\gamma}}\right|$ for all $t \in [\tilde{a}, \tilde{b}]$.

Therefore,

$$\begin{aligned}
\left|\tilde{A}_L[\hat{\gamma}(t)] - \tilde{A}_L[\tilde{\gamma}(t)]\right| &\leq \int_{\tilde{a}}^{\tilde{a}+\epsilon} (\beta + M)\mathrm{d}t + \int_{\tilde{b}-\epsilon}^{\tilde{b}} (\beta + M)\mathrm{d}t \\
&\leq 2(\beta + M)\epsilon. \qquad (4.17)
\end{aligned}$$

Choosing $\epsilon < \frac{1}{2}\frac{\delta}{2(\beta+M)}$ (where δ satisfies (4.11)), (4.17) results in,

$$\left|\tilde{A}_L[\hat{\gamma}(t)] - \tilde{A}_L[\tilde{\gamma}(t)]\right| < \frac{\delta}{2}.$$

Now, since $\gamma_0 \in C^2([a, b], M)$ and $\hat{\gamma} \in C^2([\tilde{a}, \tilde{b}], M)$ satisfies Properties (i) and (vi), the curve $\hat{\gamma}_0$ defined in (4.13) certainly belongs to $\mathcal{D}(C^2, x_a, x_b)$. ■

Since $\tilde{\gamma}$ is a minimizer of $\tilde{A}_L$ we have,

$$\begin{aligned}
\int_a^b L(\dot{\gamma}_0(t))\mathrm{d}t &= \int_a^{\tilde{a}} L(\dot{\gamma}_0(t))\mathrm{d}t + \int_{\tilde{a}}^{\tilde{b}} L(\dot{\gamma}_0)\mathrm{d}t + \int_{\tilde{b}}^{b} L(\dot{\gamma}_0)\mathrm{d}t \\
&= \int_a^{\tilde{a}} L(\dot{\gamma}_0(t))\mathrm{d}t + \int_{\tilde{a}}^{\tilde{b}} L(\dot{\tilde{\gamma}}(t))\mathrm{d}t + \delta + \int_{\tilde{b}}^{b} L(\dot{\gamma}_0(t))\mathrm{d}t \\
&\geq \int_a^{\tilde{a}} L(\dot{\gamma}_0)\mathrm{d}t + \int_{\tilde{a}}^{\tilde{b}} L(\dot{\hat{\gamma}}(t))\mathrm{d}t + \frac{1}{2}\delta + \int_{\tilde{b}}^{b} L(\dot{\gamma}_0(t))\mathrm{d}t \\
&\geq \int_a^{\tilde{a}} L(\dot{\gamma}_0)\mathrm{d}t + \int_{\tilde{a}}^{\tilde{b}} L(\dot{\hat{\gamma}}(t))\mathrm{d}t + \int_{\tilde{b}}^{b} L(\dot{\gamma}_0(t))\mathrm{d}t \\
&= \int_a^b L(\dot{\hat{\gamma}}_0(t))\mathrm{d}t.
\end{aligned}$$

Therefore, $\gamma_0(t)$ cannot minimize A_L. ■

An important consequence of Theorem 4.6 is that to check that a curve γ on a manifold M minimizes an action A_L, it is equivalent to check that for any set of charts $\{(\phi_i, U_i)\}$ covering γ, all of the coordinate representations γ_{ϕ_i} are solutions to corresponding local minimization problems.

5. Conclusion and Further Work

5.1. Conclusion. In this report we have elaborated on the previous work of Nester [Nester 1988], in order to provide a thorough coordinate-free derivation of the Euler–Lagrange equations. We began by using a variational approach to define two vector fields; $\dot{\gamma}$, the tangent vector field to the curve, and $\delta\sigma$, the vector field which defines the infinitesimal variation. We showed that the condition for stationarity can be expressed in terms of the Lie derivative of the Lagrangian with respect to the complete lift of an arbitrary extension of $\delta\sigma$.

After proving two identities which allowed us to preserve our original requirement that $\gamma(t)$ be only of class C^2, we presented a coordinate-free derivation of the Euler–Lagrange equations. Most importantly, we addressed the fact that using complete lifts in the derivation necessitated first extending $\dot{\gamma}(t)$ and $\delta\sigma$ to neighbourhoods of image(γ). In line with this we showed, using coordinate calculations, that the derivation was in fact independent of our choice of extension.

Lastly, we showed that, if a curve is a C^2-minimizer on a certain interval then it must also solve a corresponding minimization problem on any subinterval. A corollary of this result is that a C^2-curve on a manifold solves a minimization problem if and only if its coordinate expression solves a corresponding minimization problem in a set of charts covering the curve.

A summary of the results given in this report follows:

- An equivalent condition for stationarity is

$$\int_a^b (\mathcal{L}_{Z^T} L)\dot{\gamma}(t)\mathrm{d}t = 0,$$

  where Z^T is the complete lift of an arbitrary extension of $\delta\sigma$. Moreover this equation is independent of how $\delta\sigma$ is extended.

- If $\dot{\gamma}$ and $\delta\sigma$ are defined on a Riemannian manifold, then they can be extended to cover the tubular neighbourhood defined by the exponential map. This extension is constructed using the push-forward of certain vertical vector fields with respect to the exponential map.

- If $\dot{\gamma}$ and $\delta\sigma$ are defined on a manifold which admits partitions of unity, then they can be extended to cover the open neighbourhoods U_i of the submanifold charts $\{(U_i, \phi_i)\}$. This extension is constructed by extending the local vectors in the coordinate neighbourhoods given by the submanifold charts, and then patching together the resulting vector fields on M using a partition of unity.

- Given that $\tilde{X}$, $\tilde{Z} \in \Gamma^0(TTM)$ satisfy $T\pi(\tilde{X})$, $T\pi(\tilde{Z}) \in \Gamma^1(TM)$, then

$$\mathrm{d}(i_J dL)(\tilde{X}, \tilde{Z}) = \mathcal{L}_{\tilde{X}}\langle dL, \mathrm{vlft}(T\pi(\tilde{Z}))\rangle - \mathcal{L}_{\tilde{Z}}\langle dL, \mathrm{vlft}(T\pi(\tilde{X}))\rangle - \langle dL, \mathrm{vlft}[T\pi(\tilde{X}), T\pi(\tilde{Z})]\rangle.$$

- Let X, $Z \in \Gamma^1(TM)$. At points $v_x \in TM$ satisfying $[X, Z](x) = 0$ and $X(x) = v_x$ we have

$$(\mathcal{L}_{Z^T}\langle dL, V - J\tilde{X}\rangle)(v_x) = \langle \mathcal{L}_{Z^T} dL, V - J\tilde{X}\rangle(v_x) + \langle dL, \mathrm{vlft}([Z, X])\rangle(v_x).$$

- The coordinate-free derivation (4.10) of the Euler–Lagrange equation is independent of vector field extension.
- If a C^2-curve solves the minimization problem on a certain interval, then it solves a corresponding minimization problem on any subinterval.

5.2. Further Work. The results in this report admit further generalization and suggest further study.

- Time-dependent Lagrangian: All results in this report have been for a time independent Lagrangian. For a more general theory the results should be generalized to include time-dependent Lagrangians. This is straightforward.
- Coordinate-free derivations of Major Theorems of Calculus of Variations: In particular, it would be useful to give the coordinate-free derivations of Hilbert's criterion, Lagrange's Lemma and the Weierstrass–Erdmann Corner conditions. These would serve to further elucidate the issues involving differentiability constraints.
- Minimization over a more general class of curves: Armed with a coordinate-free versions of the major theorems of Calculus of Variations, the next step would be to extend the theory to a more general class of curves (C^1- or piecewise C^1-curves).
- Geometric study of identities from Section 4.1: The identities used to relax the differentiability constraints on X^T and Z^T were proven using coordinate calculations. It would be informative to both determine the geometric significance of these identities and to find their coordinate-free proofs.

Acknowledgment. Foremost I would like to thank my supervisor Andrew Lewis for always being available for invaluable discussion and for taking a vested interest in my education and work. Thanks are also due to John Chapman, David Tyner and Ajit Bhand; John for always having interesting answers to my questions, Dave for directing me towards numerous indispensable resources and Ajit for always being willing to discuss differential geometry over a cup of coffee.

List of Symbols

Below is a roughly alphabetical list of symbols that reoccur often throughout the text.

A_L	the action of L 10
$\tilde{A}_L$	restriction of A_L 21
d	the exterior derivative 10
d_J	the vertical derivative 10
$\mathcal{D}(C^2, \gamma_0(\tilde{a}), \gamma_0(\tilde{b}))$	a certain class of curves 21
$\mathcal{D}(C^2, x_a, x_b)$	a certain class of curves 10
$\delta\sigma$	an infinitesimal variation 11
E_L	the one-form used in expression for Euler–Lagrange equation . . 16
exp	the exponential map 7

References

Abraham, R., Marsden, J. E., and Ratiu, T. S. [1988] *Manifolds, Tensor Analysis, and Applications*, number 75 in Applied Mathematical Sciences, Springer-Verlag: New York/-Heidelberg/Berlin, ISBN: 978-0-387-96790-5.

Arnol′d, V. I. [1978] *Mathematical Methods of Classical Mechanics*, translated by K. Vogtmann and A. Weinstein, number 60 in Graduate Texts in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 0-387-96890-3, New edition: [Arnol′d 1989].

— [1989] *Mathematical Methods of Classical Mechanics*, translated by K. Vogtmann and A. Weinstein, 2nd edition, number 60 in Graduate Texts in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-96890-2, First edition: [Arnol′d 1978].

Bullo, F. and Lewis, A. D. [2004] *Geometric Control of Mechanical Systems, Modeling, Analysis, and Design for Simple Mechanical Systems*, number 49 in Texts in Applied Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-22195-3.

Crampin, M. and Pirani, F. A. E. [1986] *Applicable Differential Geometry*, number 59 in London Mathematical Society Lecture Note Series, Cambridge University Press: New York/Port Chester/Melbourne/Sydney, ISBN: 978-0-521-23190-9.

Edelen, D. G. B. [1985] *Applied Exterior Calculus*, John Wiley and Sons: NewYork, NY, ISBN: 0-471-80773-7, Reprint: [Edelen 2005].

— [2005] *Applied Exterior Calculus*, Dover Publications, Inc.: New York, NY, ISBN: 978-0-486-43871-9, Original: [Edelen 1985].

Gamboa Saraví, R. E. and Solomin, J. E. [2003] *On the global version of Euler–Lagrange equations*, Journal of Physics. A. Mathematical and Theoretical, **36**(26), pages 7301–7305, ISSN: 1751-8113, DOI: 10.1088/0305-4470/36/26/306.

Kolář, I., Michor, P. W., and Slovák, J. [1993] *Natural Operations in Differential Geometry*, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-3-540-56235-1.

Nester, J. M. [1988] *Invariant derivation of the Euler–Lagrange equation*, Journal of Physics. A. Mathematical and Theoretical, **21**(21), pages 1013–1017, ISSN: 1751-8113, DOI: 10.1088/0305-4470/21/21/003.

Troutman, J. L. [1996] *Variational Calculus and Optimal Control, Optimization with Elementary Convexity*, 2nd edition, Undergraduate Texts in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-94511-8.