Nonholonomic and Constrained Variational Methods Applied to a Rolling Disc

Julian Christopher[1]

2019/28/05

[1]Graduate student, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
Email: julian.christopher@queensu.ca

Chapter 1

Introduction

In this project we will discuss some the results from *Nonholonomic and constrained variational mechanics* by Andrew Lewis, 2018 in a concrete fashion using the example of a disc rolling on a flat plane. This example provides us with a mechanical system that features nonholonomic constraints, while being simple enough for the equations of motion to be written and manipulated within the body of the report. I will also attempt to discuss the geometric context of the objects featured whenever possible, and try to build a strong link between the geometry of the mathematical system and the physicality of the example. It is in this way that I think we can make the best use of the example in building our intuition about the results of Lewis [2018] that we wish to discuss.

Nonholonomic and constrained variational mechanics features a number of new results presented with a high level of generality and in a completely coordinate free fashion. It makes a comparison of the equations of motion for systems featuring nonholonomic constraints acquired through the physical principles of Newtonian mechanics and a variational principle; these two sets of equations do not in general agree, nor is one a subset of the other. The criterion for all constrained variational trajectories of a mechanical system to agree with the Newtonian trajectories is well known to be exactly when the constraint distribution is integrable; a complete characterization of the reverse situation, when all Newtonian trajectories are constrained variational trajectories, is an original contribution of *Nonholonomic and constrained variational mechanics*, and is presented as the main result of the paper. Also original to the paper is a characterization of the constrained variational equations of motion that we will explore through the rolling disc example.

1.1. Literature Review

The question of when the physically correct equations of motion for a nonholonomic mechincal system agree with those satisfying a variational principle has been contended with in a body of literature spanning at least the last three decades. The modern approach seems to originate with Kozlov [1992] and Kharlomov [1992]. For other formulations of the problem we look to [Borisov, Mamaev, and Bizyaev 2017, Cardin and Favretti 1996, Favretti 1998, Gràcia, Marin–Solano, and Muñoz–Lecanda 2003, Kupka and Oliva 2001, Lewis and Murray 1995, Vershik and Gershkovich 1990, Zampieri 2000]. The comparison of the two equations for a subset of nonholonomic systems known as Chaplygin systems is addressed in [Crampin and Mestdag 2010, Favretti 1998, Fernendez and Bloch 2008, Jóźwikowski and Respondek 2018, Rumiantsev 1978 and Terra 2018]. In [Cortés, de Léon, Martín de Diego and Martínez 2002] an algorithmic approach to the comparison problem is presented. The constrained variational approach to the rolling disc was worked out in a non-geometric fashion in Bloch [2003, example 7.1.4], and the nonholonomic equations are worked out in a geometric fashion in Lewis and Bullo [2005, chapter 4].

1.2. Overview

In Section 2 we introduce, and have a discussion about, the various objects required for computing the equations of motion for the rolling disc using the theorems from *Nonholonomic and constrained variational mechanics*, including the required notions of curvature. In section 3 we give some background on the optimization of functions on manifolds with domain constraints, define the spaces of curves that will model the motions and velocities of the rolling disc, and discuss the notions of constrained optimization used in the theorems characterizing the nonholonomic and constrained variational equations of motion. In section 4 we set up the necessary objects for analysis the rolling disc. In section 5 we compute the equations of motion using the two methods. In Section 6.1 we prove that all Newtonian nonholonomic trajectories are also variational trajectories for the rolling disc and make observations. In Section 7 we summarize the results of our exploration and discuss potentially interesting next steps.

1.3. Notation

The following conventions will be adopted throughout the paper unless otherwise stated.

- M refers to a generic smooth manifold.
- TM refers to the tangent bundle of a generic smooth manifold.
- T_xM refers to the tangent space at $x \in M$.
- I refers to a closed interval in $\mathbb{R}$.
- $\mathcal{D}$ is a distribution on a vector bundle, for our purposes all distributions will smooth subbundles.
- $\mathcal{P}_{\mathcal{D}}$ and $\mathcal{P}_{\mathcal{D}^\perp}$ refer to the projection mappings from a vector bundle, to a distribution $\mathcal{D}$ on that vector bundle, and to it's $\mathbb{G}$-orthogonal complement, respectively.
- $[a, b]$ is the closed interval in $\mathbb{R}$ from a to b.
- (a, b) is the open interval in $\mathbb{R}$ from a to b.
- $[a, b)$ is the half open interval containing a but not b.
- C^∞ means infinitely differentiable.
- σ is a variation of a curve.

- $\Gamma^r(E)$ is the set of C^r sections of the vector bundle E.
- For $f \in C^1(M;N)$ $D(f)$ is the Jacobian matrix of f.
- For a curve $\gamma : I \longrightarrow M$ that is written in a co-ordinate chart as $(x^1(t), x^2(t), \ldots, x^n(t))$, $\dot{x}^i(t) = \frac{d}{dt}x(t)$.
- $\overset{\mathbb{G}}{\nabla}$ is the Levi-Civita affine connection for the Riemannian metric $\mathbb{G}$.

Chapter 2

Geometric Constructions Used in Trajectory Characterization Theorems

The theorems characterizing constrained nonholonomic and constrained variational trajectories make use of four geometric constructions. In this section we will introduce each of them and breifly discuss their geometric intuition.

2.1. Constrained Connections and the Second Fundamental Form

In this exploration we are often interested in the geodesics of an affine connection having the property that their tangent vectors lie in a subbundle of the tangent space, denote this subbundle $\mathcal{D}$. The following observation from Lewis [1997] motivates some related constructions.

Theorem 2.1.1 (Lewis 1997, Section 5.1)**.** *Let $(M, \mathbb{G})$ be a C^∞ Riemannian manifold with $\overset{\mathbb{G}}{\nabla}$ the Levi-Civita affine connection on TM. Let $\mathcal{D}$ be a distribution on M. If $\gamma : [t_0, t_1] \to M$ is a geodesic of $\overset{\mathbb{G}}{\nabla}$ such that $\gamma'(t) \in \mathcal{D}$ for all $t \in [t_0, t_1]$, then there exists a section λ of $\mathcal{D}^\perp$ along γ such that*

$$\begin{aligned} \overset{\mathbb{G}}{\nabla}_{\gamma'(t)} \gamma'(t) &= \lambda(t), \ \textit{and} \\ \mathcal{P}_{\mathcal{D}^\perp}\left(\gamma'(t)\right) &= 0. \end{aligned} \tag{2.1.0.1}$$

Taking the covariant derivative of (2.1.0.1) gives

$$\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} \left(\mathcal{P}_{\mathcal{D}^\perp}\left(\gamma'(t)\right)\right) = \left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} \mathcal{P}_{\mathcal{D}^\perp}\right)(\gamma') + \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} \gamma'(t)\right) = 0; \tag{2.1.0.2}$$

from this we can see that

$$\mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} \gamma'(t)\right) = -\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} \mathcal{P}_{\mathcal{D}^\perp}\right)(\gamma').$$

an object hinted at in this discussion is of much interest and has been named and studied in it's own right.

Definition 2.1.1. *Let $(M, \mathbb{G})$ be a C^∞ Riemannian manifold with $\overset{\mathbb{G}}{\nabla}$ the Levi-Civita affine connection on TM, $X \in \Gamma^\infty(TM)$, and $Y \in \Gamma^\infty(\mathcal{D})$. The* ***second fundamental form*** *on $\mathcal{D}$ is the tensor field $S_{\mathcal{D}}(X,Y) \in \Gamma^\infty(\mathcal{D}^\perp \otimes T^*M \otimes \mathcal{D}^*)$ defined by*

$$S_{\mathcal{D}}(X,Y) = -\left(\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}^\perp}\right)(Y). \tag{2.1.0.3}$$

Curves satisfying a velocity constraint $\mathcal{D}$ can be defined as the geodesics of the constriained connection, which is defined as follows.

Definition 2.1.2. *Let $(M, \mathbb{G})$ be a C^∞ Riemannian manifold with $\overset{\mathbb{G}}{\nabla}$ the Levi-Civita affine connection on TM, $X \in \Gamma^\infty(TM)$, and $Y \in \Gamma^\infty(\mathcal{D})$.The* ***constrained connection*** *is the affine connection $\overset{\mathcal{D}}{\nabla}$ on $\mathcal{D}$ defined by*

$$\overset{\mathcal{D}}{\nabla}_X Y = \overset{\mathbb{G}}{\nabla}_X Y + \left(\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}^\perp}\right)(Y).$$

For the computations in this project, the constrained connection is more user friendly when stated as

$$\overset{\mathcal{D}}{\nabla}_X Y = \overset{\mathbb{G}}{\nabla}_X Y + \left(\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}^\perp}(Y)\right) - \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_X Y\right). \tag{2.1.0.4}$$

2.2. The Frobenius Curvature

This construction, as it is used in our equations, is related to the non-overlap of our constrained variational trajectories with the nonholonomic ones; with the curvature being zero when all variational trajectories are also nonholonomic. To describe this relationship in a more rigorous fashion, we will need a few definitions.

Definition 2.2.1. *Let M be a manifold and let $\mathcal{D} \subset TM$ be a distribution on M. The distribution $\mathcal{D}$ is* ***involutive*** *if, for any two vector fields X and Y defined on open sets of M and taking values in $\mathcal{D}$, $[X,Y]$ takes values in $\mathcal{D}$.*

Definition 2.2.2. *Let M be a manifold and let $\mathcal{D} \subset TM$ be a distribution on M. The distribution $\mathcal{D}$ is* ***integrable*** *if, for any $m \in M$, there is a a (local) submanifold $N \subset M$ such that tangent bundle of N is exactly $\mathcal{D}$ restricted to N.*

These two definitions are equivalent by the Frobenius theorem:

Theorem 2.2.1 (Frobenius)**.** *A distribution is involutive if and only if it is integrable.*

With these ideas stated we define the Frobenius curvature.

Definition 2.2.3. *The* ***Frobenius curvature*** *of a distribution $\mathcal{D}$ is the tensor field $F_{\mathcal{D}} \in \Gamma^\infty\left(\mathcal{D}^\perp \otimes \Lambda^2(\mathcal{D}^*)\right)$ given by*

$$F_{\mathcal{D}}(X,Y) = S_{\mathcal{D}}(X,Y) - S_{\mathcal{D}}(Y,X).$$

Lemma 2.2.1. *The Frobenius curvature of $\mathcal{D}$ vanishes if and only if $\mathcal{D}$ is involutive.*

Proof.

$$\begin{aligned} F_{\mathcal{D}}(X,Y) &= S_{\mathcal{D}}(X,Y) - S_{\mathcal{D}}(Y,X) \\ &= -\left(\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}^\perp}\right)(Y) + \left(\overset{\mathbb{G}}{\nabla}_Y \mathcal{P}_{\mathcal{D}^\perp}\right)(X) \end{aligned}$$

$\mathcal{P}_{\mathcal{D}^\perp}(X)$, $\mathcal{P}_{\mathcal{D}^\perp}(Y)$ are zero because $X, Y \in \mathcal{D}$

$$\begin{aligned} &= -\left(\cancelto{0}{\overset{\mathbb{G}}{\nabla}_X \left(\mathcal{P}_{\mathcal{D}^\perp}(Y)\right)} - \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_X Y\right)\right) + \cancelto{0}{\overset{\mathbb{G}}{\nabla}_Y \left(\mathcal{P}_{\mathcal{D}^\perp}(X)\right)} - \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_Y X\right) \\ &= \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_X Y\right) - \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_Y X\right) \\ &= \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_X Y - \overset{\mathbb{G}}{\nabla}_Y X\right) \\ &= \mathcal{P}_{\mathcal{D}^\perp}\left([X,Y]\right) \qquad \text{because } \overset{\mathbb{G}}{\nabla} \text{ is torsion-free} \end{aligned}$$

$\mathcal{P}_{\mathcal{D}^\perp}([X,Y]) = 0$ if and only if $[X,Y] \in \mathcal{D}$. □

2.3. Geodesic Curvature

In the same way that $F_{\mathcal{D}}$ gives information about the involutivity of $\mathcal{D}$, so too does the geodeisic curvature, $G_{\mathcal{D}}$, give information about the geodesic invariance of $\mathcal{D}$. Geodesic curvature and geodesic invariance are defined as follows.

Definition 2.3.1. *The* ***geodesic curvature*** *for a distribution $\mathcal{D}$ is the tensor field $G_{\mathcal{D}} \in \Gamma^\infty(\mathcal{D}^\perp \otimes S^2(\mathcal{D}^*))$ defined by*

$$G_{\mathcal{D}}(X,Y) = S_{\mathcal{D}}(X,Y) + S_{\mathcal{D}}(Y,X).$$

Definition 2.3.2. *A distribution $\mathcal{D}$ on a manifold M with affine connection ∇ is* ***geodesically invariant*** *if, for every geodesic $\gamma : [a,b] \to M$ of ∇, $\dot\gamma(a) \in \mathcal{D}_{\gamma(a)}$ implies that $\dot\gamma(t) \in \mathcal{D}_{\gamma(t)}$ for every $t \in (a,b]$.*

The implications of geodesic curvature are characterized by the behaviour of the symmetric product of vector fields, this object is defined as follows.

Definition 2.3.3. *Let M be a C^∞ manifold with affine connection ∇ and let $X, Y \in \Gamma^\infty(TM)$. The* ***symmetric product*** *of vector fields is defined as*

$$\langle X : Y \rangle = \nabla_X Y + \nabla_Y X. \tag{2.3.0.1}$$

The link between the symmetric product and geodesic invariance is proven in the following lemma.

Lemma 2.3.1 (Lewis 2018, Lemma 2.23 (v))**.**

$$G_{\mathcal{D}}(X,Y) = \mathcal{P}_{\mathcal{D}^{\perp}}(\langle X : Y\rangle).$$

The following theorem from Lewis [1997, Theorem 4.4] shows how these definitions fit together.

Theorem 2.3.1 (Lewis 1997, Theorem 4.4)**.** *Let $\mathcal{D}$ be a distribution on a manifold M with affine connection ∇. Let $\mathscr{D}$ be the set of smooth sections of a distribution $\mathcal{D}$. The following are equivalent:*

(i) *$\mathcal{D}$ is geodesically invariant;*

(ii) *$\langle X : Y\rangle \in \mathscr{D}$ for every $X, Y \in \mathscr{D}$;*

(iii) *$\nabla_X X \in \mathscr{D}$ for every $X \in \mathcal{D}$.*

From these results it is clear that a geodesic curvature of 0 implies geodesic invariance of $\mathcal{D}$.

2.4. $\mathbb{G}$-Transposes of Curvature Tensors

For $Y \in \Gamma^{\infty}(\mathcal{D})$ let $F_{\mathcal{D}}(Y) \in \mathcal{D}^* \otimes \mathcal{D}$ be the tensor field defined by $F_{\mathcal{D}}(X,Y),\ X \in \Gamma^{\infty}(\mathcal{D})$; define similar constructions $S_{\mathcal{D}}(Y)$ and $G_{\mathcal{D}}(Y)$ for $S_{\mathcal{D}}$ and $G_{\mathcal{D}}$. We will make use of the $\mathbb{G}$-transposes of $S_{\mathcal{D}}$, $F_{\mathcal{D}}$ and $G_{\mathcal{D}}$.

Definition 2.4.1. *Let $X, Y \in \Gamma^r(\mathcal{D})$ and $\alpha \in \Gamma^r(\mathcal{D}^{\perp})$.*

(i) *The $\mathbb{G}$-**transposes** $S^*_{\mathcal{D}}$ and $S^{\star}_{D}$ of $S_{\mathcal{D}}$ are defined as*

$$\mathbb{G}(S_{\mathcal{D}}(X,Y),\alpha) = \mathbb{G}(S^*_{\mathcal{D}}(Y)(\alpha),X) = \mathbb{G}(S^{\star}_{\mathcal{D}}(X)(\alpha),Y).$$

(ii) *The $\mathbb{G}$-**transposes** $F^*_{\mathcal{D}}$ and $F^{\star}_{D}$ of $F_{\mathcal{D}}$ are defined as*

$$\mathbb{G}(F_{\mathcal{D}}(X,Y),\alpha) = \mathbb{G}(F^*_{\mathcal{D}}(Y)(\alpha),X) = \mathbb{G}(F^{\star}_{\mathcal{D}}(X)(\alpha),Y).$$

(iii) *The $\mathbb{G}$-**transposes** $G^*_{\mathcal{D}}$ and $G^{\star}_{D}$ of $G_{\mathcal{D}}$ are defined as*

$$\mathbb{G}(G_{\mathcal{D}}(X,Y),\alpha) = \mathbb{G}(G^*_{\mathcal{D}}(Y)(\alpha),X) = \mathbb{G}(G^{\star}_{\mathcal{D}}(X)(\alpha),Y).$$

Chapter 3

Background on Lagrange Multipliers and Constrained Optimization

3.1. Constrained Optimization for Functions on $\mathbb{R}^n$ via the Method of Lagrange Multipliers

Defined loosely, the method of Lagrange multipliers allows us to optimize functions with a restricted domain provided that the domain can be defined as a level set of a continuous mapping $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. More precisely, for the case of a function on $\mathbb{R}^n$, let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be C^1 functions. Let $U \subset \mathbb{R}^n$ be such that $\mathtt{x} \in U$ when $g(\mathtt{x}) = c$ for a constant $c \in \mathbb{R}^m$. The critical points of $f\big|_U$ occur exactly when $\lambda_0 D(f) + D(\lambda g) = 0$ with $\lambda_0 = \{0, 1\}$ and $\lambda \in \mathbb{R}^m$. We call the solutions for $\lambda_0 = 1$ the **regular case** and those for $\lambda_0 = 0$ the **singular case**. For either the regular case or the singular case, if g is of full rank then we will end up with $n+m$ equations in the $n+m$ variables $x^1, \ldots, x^n, \lambda_1, \lambda_2, \ldots, \lambda_m$; the $n+m$ equations are

$$\begin{aligned} &D(\lambda_0 f(\mathtt{x}) + \lambda_i g^i(\mathtt{x})) = 0; \\ &g^i(\mathtt{x}) = c^i. \end{aligned}$$

If $x \in U$ minimizes the restriction of f to U, and if $Dg(x)$ is surjective, then one can take $\lambda_0 = 1$. However, as can be seen from the example with $n = 2$ and $m = 1$ and

$$f(x_1, x_2) = x_1 + x_2, \quad g(x_1, x_2) = x_1^2 + x_2^2 - 1,$$

it may happen that λ_0 must be 0.

3.2. Spaces of Curves on Manifolds

In this section we will develop a system of notation for referring to various classes of curves on manifolds. For an interval $I \subseteq \mathbb{R}$, $A \subseteq \mathbb{R}$, and for $p \in [1, \infty)$, we denote by $L^p(I; A)$ the set of measurable A-valued functions f on I for which

$$\int_I |f(t)|^p dt < \infty. \tag{3.2.0.1}$$

For $s \in \mathbb{Z}_{\geq 0}$, by $\mathrm{H}^s(I, \mathbb{R})$ we denote the set of measurable functions whose first s distributional derivatives are in $L^2(I; \mathbb{R})$.

Let $t_0, t_1 \in \mathbb{R}$ satisfy $t_0 < t_1$ and denote, for $s \in \mathbb{Z}_{\geq 0}$,

$$\mathrm{H}^s([t_0, t_1]; M) = \{\gamma : [t_0, t_1] \to M \mid \forall f \in C^\infty(M), f \circ \gamma \in \mathrm{H}^s([t_0, t_1]; \mathbb{R})\}.$$

Given $x_0, x_1 \in M$, denote

$$\mathrm{H}^s([t_0, t_1]; M; x_0, x_1) = \{\gamma \in \mathrm{H}^s([t_0, t_1]; M) \mid \gamma(t_0) = x_0, \gamma(t_1) = x_1\}.$$

Now suppose that $\mathcal{D} \subseteq TM$ is a smooth subbundle and that $s \in \mathbb{Z}_{>0}$. Denote

$$\mathrm{H}^s([t_0, t_1]; M; \mathcal{D}) = \{\gamma \in \mathrm{H}^s([t_0, t_1]; M) \mid \gamma'(t) \in \mathcal{D}_{\gamma(t)} \ a.e. \ t \in [t_0, t_1]\}$$

and

$$\mathrm{H}^s([t_0, t_1]; M; \mathcal{D}; x_0, x_1) = \mathrm{H}^s([t_0, t_1]; M; \mathcal{D}) \cap \mathrm{H}^s([t_0, t_1]; M; x_0, x_1).$$

Now we consider sections along the curves in M that we have just defined. Let $\pi : E \longrightarrow M$ be a C^∞-vector bundle, denote

$$\mathrm{Aff}^\infty(E) = \{F \in C^\infty(E) \mid F|E_x \text{ is affine for each } x \in M\}. \tag{3.2.0.2}$$

Lewis [2018, Lemma 3.1] shows that

$$\mathrm{H}^s([t_0, t_1]; E) = \{\xi : [t_0, t_1] \to E \mid F \circ \xi \in \mathrm{H}^s([t_0, t_1]; \mathbb{R}), f \in \ \mathrm{Aff}^\infty(E)\}.$$

For a fixed $\gamma : [t_0, t_1] \longrightarrow M$ denote

$$\gamma^* E = \{(t, e) \in [t_0, t_1] \times E \mid \gamma(t) = \pi(e)\}.$$

Then we can define

$$\mathrm{H}^s([t_0, t_1]; \gamma^* E) = \{\xi : [t_0, t_1] \to E \mid \pi \circ \xi = \gamma, F \circ \xi \in \mathrm{H}^s([t_0, t_1]; \mathbb{R}), f \in \ \mathrm{Aff}^\infty(E)\},$$

which simplifies to

$$\mathrm{H}^s([t_0, t_1]; \gamma^* E) = \{\xi \in \mathrm{H}^s([t_0, t_1]; E) \mid \pi \circ \xi = \gamma\}.$$

Lewis [2018, Lemma 3.2] shows that a curve covered by a regular section is regular, i.e., if $\gamma : [t_0, t_1] \longrightarrow M$ and $\xi \in \mathrm{H}^s([t_0, t_1]; \gamma^* E)$ then $\gamma \in \mathrm{H}^s([t_0, t_1]; M)$.

The construction of smooth sections along curves gives rise to a natural definition for tangent vector and tangent space to curves $\gamma \in \mathrm{H}^1([t_0, t_1]; M)$. We will mainly be considering spaces of curves of order 1 from this point forward.

Definition 3.2.1.

(i) A ***tangent vector*** *to* $\gamma \in \mathrm{H}^1([t_0, t_1]; M)$ *is an element of* $\mathrm{H}^1([t_0, t_1]; \gamma^* TM)$.

(ii) The union of all tangent vectors at γ *is the* ***tangent space*** *to* $\mathrm{H}^1([t_0, t_1]; M)$*, and we denote this by* $T_\gamma \mathrm{H}^1([t_0, t_1]; M)$.

Section 5.1 of Lewis [2018] discusses in detail the submanifold structure of classes of curves with endpoint and velocity constraints respective to $\mathrm{H}^1([t_0,t_1];M)$. The results relating to our discussion are:

(i) $\mathrm{H}^1([t_0,t_1];M;x_0)$ and $\mathrm{H}^1([t_0,t_1];M;x_0,x_1)$ are submanifolds of $\mathrm{H}^1([t_0,t_1];M)$, this is covered in Lewis [2018, section 5.1.1];

(ii) $\mathrm{H}^1([t_0,t_1];M;\mathcal{D})$ is a submanifold of $H^1([t_0,t_1];M)$ and its tangent space is

$$T_\gamma \mathrm{H}^1([t_0,t_1];M;\mathcal{D}) = \left\{\xi \in T_\gamma \mathrm{H}^1([t_0,t_1];M) \mid \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'}\xi\right) - S_{\mathcal{D}}(\xi,\gamma') = 0\right\}$$

(Lewis [2018, section 5.1.2]);

(iii) $\mathrm{H}^1([t_0,t_1];M;\mathcal{D};x_0)$ is a submanifold of $\mathrm{H}^1([t_0,t_1];M;\mathcal{D})$ and its tangent space is

$$T_\gamma \mathrm{H}^1([t_0,t_1];M;D,x_0) = \left\{\xi \in T_\gamma \mathrm{H}^1([t_0,t_1];M) \mid \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'}\xi\right) - S_{\mathcal{D}}(\xi,\gamma') = 0, \xi(t_0) = 0\right\}$$

(Lewis [2018, section 5.1.3]);

(iv) $\mathrm{H}^1([t_0,t_1];M;\mathcal{D};x_0,x_1)$ is a submanifold of $\mathrm{H}^1([t_0,t_1];M;\mathcal{D})$ only when γ is a **regular curve**. For this discussion we will take $\mathrm{H}^1([t_0,t_1];M;\mathcal{D};x_0,x_1)$ being a submanifold as a definition for regular curve; this is not the definition used in Lewis [2018, section 5.1.3] but it will function without the need to introduce layers of complexity that we do not otherwise need. Curves that are not regular are called **singular curves**. For a regular curve γ,

$$T_\gamma \mathrm{H}^1([t_0,t_1];M;\mathcal{D};x_0,x_1) = \left\{\xi \in T_\gamma \mathrm{H}^1([t_0,t_1];M;\mathcal{D},x_0) \mid \xi(t_1) = 0\right\}$$

(Lewis [2018, section 5.1.3]).

It will be useful to have the language of variations and infinitesimal variations at our disposal, so I will present the definitions of those terms here.

Definition 3.2.2. *Let M be a smooth manifold, let $t_0,t_1 \in \mathbb{R}$ satisfy $t_0 < t_1$ and let $\gamma \in \mathrm{H}^1([t_0,t_1];M)$.*

*(i) A **variation** of γ is a mapping $\sigma : (-a,a) \to \mathrm{H}^1([t_0,t_1];M)$, where*

(a) $a \in \mathbb{R}_{>0}$,

(b) $\sigma(0) = \gamma$, and

(c) σ is continuously differentiable.

*(ii) An **infinitesimal variation** ,δ, of γ is an element of $\mathrm{H}^1([t_0,t_1];\gamma^*TM)$.*

Infinitesimal variations are usefully thought of as, and often defined as, derivatives of variations; this is made explicit in the following lemma.

Lemma 3.2.1 (Variations and infinitesimal variations Lewis 2018, Lemma 3.12)**.** *Let M be a smooth manifold, let $t_0,t_1 \in \mathbb{R}$ satisfy $t_0 < t_1$ and let $\gamma \in \mathrm{H}^1([t_0,t_1];M)$. Then the following statements hold:*

(i) if $\sigma : (-a, a) \to \mathrm{H}^1([t_0, t_1]; M)$ is a variation of γ then $\delta\sigma(0)$ is an infinitesimal variation of γ;

(ii) if $\delta \in \mathrm{H}^1([t_0, t_1]; \gamma^*TM)$, then there exists a variation σ of γ such that $\delta = \delta\sigma(0)$.

Notice, that if a variation of γ lies in $\mathrm{H}^1([t_0, t_1]; M)$, then its corresponding infinitesimal variation $\delta\sigma(0)$ is an element of $T_\gamma\mathrm{H}^1([t_0, t_1]; M)$, and that if $\sigma \in \mathrm{H}^1([t_0, t_1]; M; x_0, x_1)$, then $\delta\sigma(0)$ vanishes at the endpoints.

3.3. Constrained Optimization of Curves and its Application to Mechanics

The methods that we will use to determine constrained variational trajectories are analogous to the method of Lagrange multipliers outlined on section 3.1 in that we want to optimize a function $F : \mathrm{H}^1([t_0, t_1]; M) \longrightarrow \mathbb{R}$ on a submanifold of $\mathrm{H}^1([t_0, t_1]; M)$. Since $\mathrm{H}^1([t_0, t_1]; M)$is an infinite dimensional space, the characterization of the constraint submanifold as level sets of C^1 functions is no longer valid. For restriction to curves that have tangent vectors in a subbundle $\mathcal{D}$ of TM, we restrict the domain of F to one of the four submanifolds discussed in section 3.2. The function from spaces of curves to manifolds that we wish to minimize in classical mechanics is the kinetic energy action, which we will now define.

Definition 3.3.1 (Kinetic Energy Action and Constrained Kinetic Energy Action)**.** *Let $(M, \mathbb{G})$ be a smooth Riemannian manifold, let $t_0, t_1 \in \mathbb{R}$ satisfy $t_0 < t_1$.*

*(i) The **kinetic energy function** is the mapping*

$$\begin{aligned} \mathrm{K}_{\mathbb{G}} : TM &\longrightarrow \mathbb{R} \\ v_x &\mapsto \frac{1}{2}\mathbb{G}(v_x, v_x). \end{aligned}$$

*(ii) The **kinetic energy action** is the mapping*

$$\begin{aligned} A_{\mathbb{G}} : \mathrm{H}^1([t_0, t_1]; \mathrm{M}) &\longrightarrow \mathbb{R} \\ \gamma &\mapsto \frac{1}{2}\int_a^b \mathrm{K}_{\mathbb{G}}(\gamma'(t))dt. \end{aligned}$$

*(iii) The **constrained kinetic energy action** is the mapping*

$$\begin{aligned} A_{\mathbb{G},\mathcal{D}} : \mathrm{H}^1([t_0, t_1]; \mathrm{M}; \mathcal{D}) &\longrightarrow \mathbb{R} \\ \gamma &\mapsto A_{\mathbb{G}}(\gamma(t)). \end{aligned}$$

Optimization, of course, requires the notion of a derivative. Much care is taken in Section 3 of Lewis [2018] to make sure the derivative of the constrained kinetic energy action is well defined and behaves as it ought to; that discussion culminates in a lemma which I will restate here:

Lemma 3.3.1 (Lewis 2018 Lemma 5.13)**.** *Let* $(M, \mathbb{G})$ *be a smooth Riemannian manifold, let* $t_0, t_1 \in \mathbb{R}$ *satisfy* $t_0 < t_1$*, let* $\gamma \in \mathrm{H}^1([t_0, t_1]; \mathrm{M})$ *and let* $\xi \in \mathrm{T}_\gamma \mathrm{H}^1([t_0, t_1]; \mathrm{M})$*. Then* $A_{\mathbb{G}}$ *is differentiable at* γ *and*

$$\langle dA_{\mathbb{G}}; \xi \rangle = \int_{t_0}^{t_1} \mathbb{G}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\xi(t), \gamma'(t)\right) dt.$$

Now we are ready to make the distinction between the nonholonomic and constrained variational approaches to generating trajectories for a mechanical system. We will define the trajectories that satisfy each approach side by side and make comparisons between them. These are the definitions that are given in Lewis [2018, section 5.3 and 5.4] except that here we omit incorporation of a potential function V.

Definition 3.3.2 (Nonholonomic Trajectory)**.** *Let* $\Sigma = (M, \mathbb{G}, \mathcal{D})$ *be a* C^∞*-constrained simple mechanical system. Let* $t_0, t_1 \in \mathbb{R}$ *with* $t_0 < t_1$ *and let* $x_0, x_1 \in M$*. A curve* $\gamma \in \mathrm{H}^1([t_0, t_1]; M; \mathcal{D}; x_0, x_1)$ *is a* ***nonholonomic trajectory*** *for* Σ *if*

$$\langle dA_{\mathbb{G}}; \delta \rangle = 0, \qquad \delta \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}; x_0, x_1).$$

Definition 3.3.3 (Constrained Variational Trajectory)**.** *Let* $\Sigma = (M, \mathbb{G}, \mathcal{D})$ *be a* C^∞*-constrained simple mechanical system. Let* $t_0, t_1 \in \mathbb{R}$ *with* $t_0 < t_1$ *and let* $x_0, x_1 \in M$*. A* ***constrained variational trajectory*** *for* Σ *is a curve* $\gamma \in \mathrm{H}^1([t_0, t_1]; \mathrm{M}; \mathcal{D}; x_0, x_1)$ *such that:*

$$\langle d(A_{\mathbb{G},\mathcal{D}}); \delta \rangle = 0 \qquad \delta \in \boldsymbol{T}_\gamma \mathrm{H}^1([t_0, t_1]; \mathrm{M}; \mathcal{D}; x_0, x_1).$$

The first noticeable difference is that the differential of the constrained kinetic energy action is used in the constrained variational trajectory, while the differential of the unconstrained kinetic energy action is used in the nonholonomic trajectory. The second is that δ is simply an element of a $\mathcal{D}$-section over γ in the nonholonomic case and it is an element of the tangent space to $\mathrm{H}^1([t_0, t_1]; \mathrm{M}; \mathcal{D}; x_0, x_1)$ in the constrained variational case. These differences are motivated by exactly when we constrain the allowable trajectories.

In the constrained variational case, the constraint is made conceptually from the outset; the domain of $A_{\mathbb{G}}$ is constrained to a submanifold like it is in the constrained optimization of functions on $\mathbb{R}^n$ that we discussed in section 3.1. By only accepting infinitesimal variations from this constrained domain, $T_\gamma \mathrm{H}^1([t_0, t_1]; \mathrm{M}; \mathcal{D}; x_0, x_1)$, we have that all variations corresponding to them satisfy the constraints.

In the nonholonomic case the restriction is conceptually imposed at the end; we can make this statement clear by applying definition 3.3.2 to lemma 3.3.1; first we need another lemma.

Lemma 3.3.2 (Lewis and Bullo 2005, Section 4.3.4)**.**

$$\begin{aligned}\frac{d}{dt}\mathbb{G}(\xi(t), \gamma'(t)) &= \mathscr{L}_{\gamma'}\mathbb{G}(\xi(t), \gamma'(t)) \\ &= \mathbb{G}(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\xi(t), \gamma'(t)) + \mathbb{G}(\xi(t), \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)),\end{aligned}$$

which rearranges to

$$\mathbb{G}(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\xi(t), \gamma'(t)) = \frac{d}{dt}\mathbb{G}(\xi(t), \gamma'(t)) - \mathbb{G}(\xi(t).\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)).$$

Now plug definition 3.3.2 into lemma 3.3.1:

$$
\begin{aligned}
T_\gamma A_{\mathbb{G}}(\xi) &= \int_{t_0}^{t_1} \mathbb{G}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\xi(t), \gamma'(t)\right) dt, \\
&= \int_a^b \left(\frac{d}{dt}\mathbb{G}(\xi(t), \gamma'(t)) - \mathbb{G}(\xi(t), \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t))\right) dt \\
&= \cancel{\mathbb{G}(\xi(t), \gamma'(t))}\Big|_{t=a}^{t=b} - \int_a^b \mathbb{G}(\xi(t), \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)) dt.
\end{aligned} \tag{3.3.0.1}
$$

The final step uses the fact that ξ vanishes at the endpoints. We can see that definition 3.3.2 constrains the final trajectories to $\mathrm{H}^1([t_0, t_1]; M; \mathcal{D})$, not by constraining the domain of the action, but by using $\xi \in \mathrm{H}^1(t_0, t_1; \gamma^*\mathcal{D}, x_0, x_1)$ in equation (3.3.0.1) that will cause it to be 0 if and only if $\overset{\mathbb{G}}{\nabla}_{\gamma'}\gamma' \in \mathcal{D}^\perp$, that is, only if γ is a geodesic of the constrained connection.

Chapter 4

Setup for the Rolling Disc

4.1. The Rolling Disc

The example mechanical system that we will consider throughout this discussion is a disc rolling on a flat plane. The "rolling" part is captured by requiring that the edge of the disc that is in contact with the plane does not slip in either direction, i.e., that the velocity of the part of the disc that contacts the plane is zero. We will assume that the disc is uniform in it's composition and symmetrical in the axis along which it rolls. We also assume that the disc is very thin so that there are no frictional effects to consider when "spinning" the disc so as to change the direction in which it rolls.

4.2. Basic Objects in the Rolling Disc Mechanical System

The rolling disc problem is the simple mechanical system $\Sigma = (M, \mathbb{G}, V, \mathcal{D})$, the elements of which are as follows:

- $M = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ with coordinates $(x, y, \theta, \phi) \in \mathbb{R}^2 \times (-\pi, \pi)^2$;
- $\mathbb{G} = m(dx \otimes dx + dy \otimes dy) + J_{\text{spin}} d\theta \otimes d\theta + J_{\text{roll}} d\phi \otimes d\phi$, where m is the mass of the disc, and J_{spin} and J_{roll} are the principal inertias about the θ and ϕ axes respectively;
- $V = 0$;
- $\mathcal{D}$ is characterized by the equations $\dot{x} = \rho\dot{\phi}\cos\theta$ and $\dot{y} = \rho\dot{\phi}\cos\theta$.

The physical interpretation of the coordinate system is shown in figure 4.1.

4.3. Orthonormal Frame Fields for The Constraint Distribution

It is possible to find a global orthonormal basis of vector fields for the tangent space to Q such that $\mathcal{D}$ and $\mathcal{D}^{\perp}$ can be generated using disjoint sets of these basis vector fields, making projection of tangent vectors onto $\mathcal{D}$ or $\mathcal{D}^{\perp}$ much easier. In this section we will find those vector fields explicitly. As previously stated the constraint distribution is characterized by

$$\dot{x} = \rho\dot{\phi}\cos\theta, \tag{4.3.0.1}$$

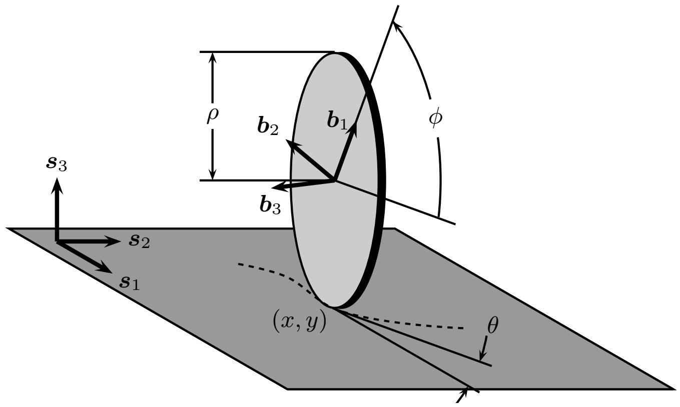

Figure 4.1: Rolling disc setup, diagram from Lewis and Bullo [2005]

$$\dot{y} = \rho\dot{\phi}\sin\theta, \tag{4.3.0.2}$$

which can be written as

$$\begin{bmatrix} 1 & 0 & 0 & -\rho\cos\theta \\ 0 & 1 & 0 & -\rho\sin\theta \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{4.3.0.3}$$

Equation (4.3.0.3) has $\mathbb{G}$-orthogonal solution vectors

$$\tilde{X}_1 = \rho\cos\theta\frac{\partial}{\partial x} + \rho\sin\theta\frac{\partial}{\partial y} + \frac{\partial}{\partial\phi},\ \tilde{X}_2 = \frac{\partial}{\partial\theta}$$

so $\mathcal{D} = \operatorname{span}_{\mathbb{R}}\{X_1, X_2\}$.

If we can find two vector fields that are $\mathbb{G}$-orthogonal to $\mathcal{D}$, the span of these would equal $\mathcal{D}^{\perp}$. The matrix in equation (4.3.0.3) has as its rows two covectors, let's call them X_3^* and X_4^*, with the property that they map both X_1 and X_2 to 0. We can use the sharp map on X_3^* and X_4^* to give us two vector fields that are $\mathbb{G}$-orthogonal to $\mathcal{D}$, and we can make them orthogonal to each other using the Gram–Schmidt algorithm:

$$\begin{aligned} \tilde{X}_3 &= \mathbb{G}^{\#}(X_3^*) = \frac{1}{m}\frac{\partial}{\partial x} - \frac{\rho}{J_{\text{roll}}}\cos\theta\frac{\partial}{\partial\phi}, \\ \tilde{X}_4 &= \mathbb{G}^{\#}(X_4^*) - \frac{\mathbb{G}(X_3, \mathbb{G}^{\#}(X_4^*))}{\mathbb{G}(X_3, X_3)}X_3 \\ &= \frac{-\rho^2\cos\theta\sin\theta}{J_{\text{roll}} + m(\rho\cos\theta)^2}\frac{\partial}{\partial x} + \frac{1}{m}\frac{\partial}{\partial y} - \frac{\rho\sin\theta}{J_{\text{roll}} + +m(\rho\cos\theta)^2}\frac{\partial}{\partial\phi}. \end{aligned}$$

We will also normalize the frame fields to make computations easier later on:

$$\mathscr{X} = \{X_1, X_2, X_3, X_4\} = \left\{\frac{\tilde{X}_1}{||\tilde{X}_1||_{\mathbb{G}}}, \frac{\tilde{X}_2}{||\tilde{X}_2||_{\mathbb{G}}}, \frac{\tilde{X}_3}{||\tilde{X}_3||_{\mathbb{G}}}, \frac{\tilde{X}_4}{||\tilde{X}_4||_{\mathbb{G}}}\right\}.$$

4.4. Christoffel Symbols in the Orthonormal Frame Field

The Christoffel symbols for $\overset{\mathbb{G}}{\nabla}$ in coordinates natural to the chart we defined for Q are all equal to zero, this is because all components of $\mathbb{G}$ are constant in that chart. We will need the Christoffel symbols for $\mathbb{G}$ in the basis $\mathscr{X}$, these we will determine in the following lemma:

Lemma 4.4.1. *For any $X_i, X_j \in \mathscr{X}$, $\overset{\mathbb{G}}{\nabla}_{X_i} X_j = 0$ when $i \neq 2$ and $\overset{\mathbb{G}}{\nabla}_{X_2} X_j = \frac{\partial}{\partial \theta} X_j$.*

Proof. Define p_i^j such that $X_i = p_i^j \frac{\partial}{\partial x^j}$ then

$$\begin{aligned}
\overset{\mathbb{G}}{\nabla}_{X_i} X_j &= \overset{\mathbb{G}}{\nabla}_{p_i^l \frac{\partial}{\partial x^l}} p_j^k \frac{\partial}{\partial x^k} \\
&= p_i^l \left[\left(\overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^l}} p_j^k \right) \frac{\partial}{\partial x^k} + \cancelto{0}{p_i^l p_j^k \overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}} \right] \\
&= p_i^l \frac{\partial}{\partial x^l} \left(p_j^k \right) \frac{\partial}{\partial x^k} \\
&= \begin{cases} \frac{\partial}{\partial \theta} X_j & \text{when } l = 2; \\ 0 & \text{when } l = 1, 3, 4. \end{cases}
\end{aligned}$$

□

Chapter 5

Equations of Motion for the Rolling Disc by Two Methods of Calculation

5.1. Theorems Characterizing Equations of Motion

Now that we have all of the pieces, let's look at the two theorems of Lewis [2018] that characterize nonholonomic and variational trajectories of a mechanical system.

Theorem 5.1.1 (Characterization of Nonholonomic Trajectories, Lewis 2018 Theorem 5.18)**.** *Let $\Sigma = (M, \mathbb{G}, V, \mathcal{D})$ be a C^∞-constrained simple mechanical system. Let $t_0, t_1 \in \mathbb{R}$ with $t_0 < t_1$ and let $x_0, x_1 \in M$. For $\gamma \in \boldsymbol{H}^1([t_0, t_1]; M; \mathcal{D}; x_0, x_1)$ The following are equivalent:*

(i) γ is a nonholonomic trajectory;

(ii) $\gamma \in \mathrm{H}^2([t_0, t_1]; \mathrm{M})$ and there exists a $\lambda \in L^2([t_0, t_1]; \gamma^* \mathcal{D}^\perp)$ such that

$$\overset{\mathbb{G}}{\nabla}_{\gamma'} \gamma' + \mathrm{grad} V \circ \gamma = \lambda;$$

(iii) γ satisfies

$$\overset{\mathcal{D}}{\nabla}_{\gamma'} \gamma' + \mathcal{P}_{\mathcal{D}} \circ \mathrm{grad} V \circ \gamma = 0.$$

Theorem 5.1.2 (Characterization of Constrained Variational Trajectories, Lewis 2018, Theorem 5.22)**.** *Let $\sigma = (M, \mathbb{G}, V, \mathcal{D})$ be a C^∞-constrained simple mechanical system. Let $t_0, t_1 \in \mathbb{R}$ with $t_0 < t_1$ and let $x_0, x_1 \in M$. For $\gamma \in \mathrm{H}^1([t_0, t_1]; M; \mathcal{D}; x_0, x_1)$, the following are equivalent:*

(i) γ is a constrained variational trajectory;

(ii) at least one of the following conditions holds:

(a) there exists a nowhere zero $\lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^* \mathcal{D}^\perp)$ such that

$$\overset{\mathbb{G}}{\nabla}_{\gamma'} \lambda + S^*_{\mathcal{D}}(\gamma')(\lambda) = 0,$$

(b) $\gamma \in \mathrm{H}^2([t_0, t_1]; M)$ *and there exists* $\lambda \in \mathrm{H}^1([T_0, t_1]; \gamma^*\mathcal{D}^\perp)$ *such that*

$$\overset{\mathbb{G}}{\nabla}_{\gamma'}\gamma' + \mathrm{grad}V \circ \gamma - \overset{\mathbb{G}}{\nabla}_{\gamma'}\lambda - S^*_{\mathcal{D}}(\gamma')(\lambda) = 0;$$

(iii) *at least one of the following conditions holds:*

(a) *there exists a nowhere zero* $\lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}^\perp)$ *such that*

$$\begin{aligned} &F^*_{\mathcal{D}}(\gamma')(\lambda) = 0, \\ &\overset{\mathcal{D}^\perp}{\nabla}_{\gamma'}\lambda = \frac{1}{2}G^*_{\mathcal{D}^\perp}(\gamma')(\lambda) + \frac{1}{2}F^*_{\mathcal{D}^\perp}(\gamma')(\lambda), \end{aligned}$$

(b) $\gamma \in \mathrm{H}^2([t_0, t_1]; M)$ *and there exists* $\lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}^\perp)$ *such that*

$$\overset{\mathcal{D}}{\nabla}_{\gamma'}\gamma' + \mathcal{P}_{\mathcal{D}} \circ \mathrm{grad}V \circ \gamma = F^*_{\mathcal{D}}(\gamma')(\lambda) \tag{5.1.0.1}$$

$$\overset{\mathcal{D}^\perp}{\nabla}_{\gamma'}\lambda = \frac{1}{2}G_{\mathcal{D}}(\gamma', \gamma') + \mathcal{P}_{\mathcal{D}} \circ \mathrm{grad}V \circ \gamma + \frac{1}{2}G^\star_{\mathcal{D}^\perp}(\gamma')(\lambda) + \frac{1}{2}F^\star_{\mathcal{D}^\perp}(\gamma')(\lambda). \tag{5.1.0.2}$$

In our exploration of the rolling disc system, we will focus on the regular constrained variational trajectories only. That is, we will work only with trajectories satisfying theorem 5.1.2 $(ii)(b)$ and $(iii)(b)$.

The significance of equation (5.1.0.1) is seen with the following theorem, which has been proved many times, possibly most recently as Theorem 7.6 of [Lewis, 2018].

Theorem 5.1.3 (When $\mathcal{D}$-regular constrained variational trajectories are also nonholonomic trajectories)**.** *Let* $\Sigma = (M, \mathbb{G}, V, \mathcal{D})$ *be a* C^∞*-constrained simple mechanical system. The following are equivalent:*

(i) *every* $\mathcal{D}$*-regular constrained variational trajectory is a nonholonomic trajectory;*

(ii) $\mathcal{D}$ *is integrable.*

We can now see that every $\mathcal{D}$-regular constrained variational trajectory is a nonholonomic trajectory if and only if $F_{\mathcal{D}} = 0$; and that in calculating the Frobenius curvature tensor for a distribution generated by a nonholonomic constraint, we can deduce under what circumstances variational trajectories are nonholonomic trajectories.

5.2. Equations of Motion for the Rolling Disc Obtained From the Constrained Nonholonomic Approach

In this section we will determine the nonholonomic equations of motion for the rolling disc; those that satisfy the Newton–Euler equations of motion and describe the physical trajectories for the system. We will use theorem 5.1.1 (iii) to find these equations. Our system has no potential function so $\mathcal{P}_{\mathcal{D}} \circ \mathrm{grad}V \circ \gamma = 0$.

The orthonormal frame fields computed in section 4.4 make a good basis in which to work for computation of the trajectories. We know that, for $v^1, v^2 : Q \longrightarrow \mathbb{R}$,

$$\gamma' = v^1 X_1 + v^2 X_2. \tag{5.2.0.1}$$

We will find $\overset{\mathcal{D}}{\nabla}_{\gamma'}\gamma'$:

$$\begin{aligned}
\overset{\mathbb{G}}{\nabla}_{\gamma'}\gamma' &= \overset{\mathbb{G}}{\nabla}_{\dot{x}^i \frac{\partial}{\partial x^i}} v^1 X_1 + v^2 X_2 \\
&= \dot{x}^i \left(\overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^i}} v^1 X^1 + \overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^i}} v^2 X^2 \right) \\
&= \dot{x} \left(\frac{\partial v^1}{\partial x^i} X_1 + \frac{\partial v^2}{\partial x^i} X_2 + v^1 \overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^i}} X^1 + \cancelto{0}{v^2 \overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^i}} X_2} \right) \\
&= \dot{x} \left(\frac{\partial v^1}{\partial x^i} X_1 + \frac{\partial v^2}{\partial x^i} X_2 \right) + v^1 \dot{\theta} \frac{\partial}{\partial \theta} X_1 \\
&= \dot{v}^1 X_1 + \dot{v}^2 X_2 + v^1 \dot{X}_1;
\end{aligned} \tag{5.2.0.2}$$

$$\overset{\mathbb{G}}{\nabla}_{\gamma'} \left(\mathcal{P}_{\mathcal{D}\perp}(\gamma') \right) = 0;$$

$$\mathcal{P}_{\mathcal{D}^\perp} \left(\overset{\mathbb{G}}{\nabla}_{\gamma'}\gamma' \right) = v^1 \mathcal{P}_{\mathcal{D}^\perp} \left(\dot{X}_1 \right). \tag{5.2.0.3}$$

Putting it all together

$$\begin{aligned}
\overset{\mathcal{D}}{\nabla}_{\gamma'}\gamma' &= \overset{\mathbb{G}}{\nabla}_{\gamma'}\gamma' + \overset{\mathbb{G}}{\nabla}_{\gamma'} \left(\mathcal{P}_{\mathcal{D}^\perp}(\gamma') \right) - \mathcal{P}_{\mathcal{D}^\perp} \left(\overset{\mathbb{G}}{\nabla}_{\gamma'}\gamma' \right) \\
&= \dot{v}^1 X_1 + \dot{v}^2 X_2 + v^1 \left(\dot{X}_1 - \mathcal{P}_{\mathcal{D}^\perp}(\dot{X}_1) \right).
\end{aligned}$$

The notation $\dot{X}_1$ is a notational convenience and is defined only as written in the context of this computation. We find through brute fore computation that $\dot{X}_1 \in \mathcal{D}^\perp$ so $\mathcal{P}_{\mathcal{D}^\perp}(\dot{X}_1) = \dot{X}_1$, this makes our final expression

$$\overset{\mathcal{D}}{\nabla}_{\gamma'}\gamma' = \dot{v}^1 X_1 + \dot{v}^2 X_2, \tag{5.2.0.4}$$

so $\dot{v}^1 = 0$ and $\dot{v}^2 = 0$. Transforming (5.2.0.1) into the standard basis gives

$$\dot{x} = \frac{\rho \cos(\theta) v^1}{\sqrt{j_{\text{roll}} + m\rho^2}}, \qquad \dot{y} = \frac{\rho \sin(\theta) v^1}{\sqrt{j_{\text{roll}} + m\rho^2}}, \tag{5.2.0.5}$$

$$\dot{\theta} = \frac{v^2}{\sqrt{J_{\text{spin}}}}, \qquad \dot{\phi} = \frac{v^1}{\sqrt{j_{\text{roll}} + m\rho^2}}. \tag{5.2.0.6}$$

We see from (5.2.0.6) that $v^1 = \dot{\phi}\sqrt{j_{\text{roll}} + m\rho^2}$ and $v^2 = \dot{\theta}\sqrt{J_{\text{spin}}}$; these can be substituted into equations (5.2.0.5) to obtain $\dot{x} = \rho\dot{\phi}\cos\theta$ and $\dot{y} = \rho\dot{\phi}\sin\theta$, which are simply our constraint equations (4.3.0.1) and (4.3.0.2). If we differentiate equations (5.2.0.6) we get $\ddot{\theta} = \frac{\dot{v}^2}{\sqrt{J_{\text{spin}}}} = 0$ and $\ddot{\phi} = \frac{\dot{v}^1}{\sqrt{j_{\text{roll}} + m\rho^2}} = 0$. The equations of motion can then be written as

$$
\begin{aligned}
\dot{x} &= \rho\dot{\phi}\cos\theta, & \ddot{\theta} &= 0, \\
\dot{y} &= \rho\dot{\phi}\sin\theta, & \ddot{\phi} &= 0.
\end{aligned}
\tag{5.2.0.7}
$$

These equations make sense for a rolling disc absent external forces: the initial angular velocity in the roll and spin direction are conserved which is why $\ddot{\theta}$ and $\ddot{\phi}$ are zero. The rate of progress in the planar directions is equal to the length of arc traversed by the edge of the disc: $\rho\dot{\phi}$, $\dot{x}$ and $\dot{y}$ are simply components of $\rho\dot{\phi}$ in an orthonormal coordinate system.

5.3. Equations of Motion for the Rolling Disc Obtained From Constrained Variational Approach

5.3.1. Covariant Derivatives Used in Calculating Variational Trajectories

We will need six covariant derivatives for the calculations, let $X = w^i X_i$ be a vector field in $\mathrm{H}^1([t_0, t_1]; \gamma^* TQ)$. When we use these calculations it will be necessary to assume that X is either in $\mathrm{H}^1([t_0, t_1]; \gamma^* \mathcal{D})$ or $\mathrm{H}^1([t_0, t_1]; \gamma^* \mathcal{D}^{\perp})$, but as we will see, the final forms of the calculations will allow us to switch between these cases without much trouble. Recalling our calculation of the Christoffel symbols for $\mathscr{X}$ in lemma 4.4.1, we compute

$$
\begin{aligned}
\overset{\mathbb{G}}{\nabla}_{\gamma'} X &= \overset{\mathbb{G}}{\nabla}_{v^1 X_1 + v^2 X_2} w^i X_i \\
&= v^1 \left(\overset{\mathbb{G}}{\nabla}_{X_1} w^i X_i \right) + v^2 \left(\overset{\mathbb{G}}{\nabla}_{X_2} w^i X_i \right) \\
&= v^1 \left((\mathscr{L}_{X_1} w^i) X_i + w^i \cancel{\overset{\mathbb{G}}{\nabla}_{X_1} X_i} \right) + v^2 \left((\mathscr{L}_{X_2} w^i) X_i + w^i \overset{\mathbb{G}}{\nabla}_{X_2} X_i \right) \\
&= \left(v^1 \mathscr{L}_{X_1} w^i + v^2 \mathscr{L}_{X_2} w^i \right) X_i + v^2 \left(w^i \frac{\partial}{\partial\theta} X_i \right),
\end{aligned}
$$

$$
\begin{aligned}
\overset{\mathbb{G}}{\nabla}_{X} \gamma' &= \overset{\mathbb{G}}{\nabla}_{w^i X_i} (v^1 X_1 + v^2 X_2) \\
&= w^i \left(\overset{\mathbb{G}}{\nabla}_{X_i} v^1 X_1 + \overset{\mathbb{G}}{\nabla}_{X_i} v^2 X_2 \right) \\
&= w^i \left(\left(\mathscr{L}_{X_i} v^1 \right) X_1 + v^1 \overset{\mathbb{G}}{\nabla}_{X_i} X_1 + \left(\mathscr{L}_{X_i} v^2 \right) X_2 + v^2 \overset{\mathbb{G}}{\nabla}_{X_i} X_2 \right) \\
&= w^i \left(\left(\mathscr{L}_{X_i} v^1 \right) X_1 + \left(\mathscr{L}_{X_i} v^2 \right) X_2 \right) + w^2 v^1 \frac{\partial}{\partial\theta} X_1,
\end{aligned}
$$

$$
\begin{aligned}
\overset{\mathbb{G}}{\nabla}_{\lambda} X &= \overset{\mathbb{G}}{\nabla}_{u^1 X_3 + u^2 X_3} w^i X_i \\
&= u^1 \left((\mathscr{L}_{X_3} w^i) X_i + \cancel{\overset{\mathbb{G}}{\nabla}_{X_3} X_i} \right) + u^2 \left((\mathscr{L}_{X_4} w^i) X_i + \cancel{\overset{\mathbb{G}}{\nabla}_{X_4} X_i} \right)
\end{aligned}
$$

$$= \left(u^1 \mathscr{L}_{X_3} w^i + u^2 \mathscr{L}_{X_4} w^i\right) X_i,$$

$$\begin{aligned}
\overset{\mathbb{G}}{\nabla}_X \lambda &= \overset{\mathbb{G}}{\nabla}_{w^i X_i}(u^1 X_3 + u^2 X_3) \\
&= w^i \left(\left(\mathscr{L}_{X_i} u^1\right) X_3 + u^1 \overset{\mathbb{G}}{\nabla}_{X_i} X_3 + \left(\mathscr{L}_{X_i} u^2\right) X_4 + u^2 \overset{\mathbb{G}}{\nabla}_{X_i} X_4 \right) \\
&= w^i \left(\left(\mathscr{L}_{X_i} u^1\right) X_3 + \left(\mathscr{L}_{X_i} u^2\right) X_4 \right) + w^2 u^1 \frac{\partial}{\partial \theta} X_3 + w^2 u^2 \frac{\partial}{\partial \theta} X_4.
\end{aligned}$$

For the calculation of $\overset{\mathbb{G}}{\nabla}_{\gamma'} \lambda$, it is easiest to use $\gamma' = \dot{x}^i \frac{\partial}{\partial x^i}$ and recall that we used $X_i = p_i^k \frac{\partial}{\partial x^k}$ to denote the co-ordinate functions of the basis vectors in the frame field $\mathscr{X}$. We calculate

$$\begin{aligned}
\overset{\mathbb{G}}{\nabla}_{\gamma'} \lambda &= \overset{\mathbb{G}}{\nabla}_{\dot{x}^i \frac{\partial}{\partial x^i}} (u^1 X_3 + u^2 X_4) \\
&= \dot{x}^i \left(\overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^i}} u^1 X_3 + \overset{\mathbb{G}}{\nabla}_{\frac{\partial}{\partial x^i}} u^2 X_4 \right) \\
&= \dot{x}^i \left(\frac{\partial u^1}{\partial x^i} X_3 + u^1 \left(\frac{\partial p_3^k}{\partial x^i} \frac{\partial}{\partial x^k} + \cancel{p_3^k \Gamma_{ik}^j \frac{\partial}{\partial x^j}} \right) \right. \\
&\qquad \left. + \frac{\partial u^2}{\partial x^i} X_4 + u^2 \left(\frac{\partial p_4^k}{\partial x^i} \frac{\partial}{\partial x^k} + \cancel{p_4^k \Gamma_{ik}^j \frac{\partial}{\partial x^j}} \right) \right) \\
&= \dot{u}^1 X_3 + u^1 \dot{X}_3 + \dot{u}^2 X_4 + u^2 \dot{X}_4 \\
&= \dot{\lambda}.
\end{aligned} \tag{5.3.1.1}$$

Using the above result

$$\begin{aligned}
\overset{\mathcal{D}^\perp}{\nabla}_{\gamma'} \lambda &= \overset{\mathbb{G}}{\nabla}_{\gamma'} \lambda + \left(\cancel{\overset{\mathbb{G}}{\nabla}_{\gamma'} (\mathcal{P}_{\mathcal{D}}(\lambda))} - \mathcal{P}_{\mathcal{D}} \left(\overset{\mathbb{G}}{\nabla}_{\gamma'} \lambda \right) \right) \\
&= \dot{\lambda} - \mathcal{P}_{\mathcal{D}} \dot{\lambda}.
\end{aligned}$$

and from equation (5.2.0.2)

$$\overset{\mathbb{G}}{\nabla}_{\gamma'} \gamma' = \dot{v}^1 X_1 + \dot{v}^2 X_2 + v^1 \dot{X}_1. \tag{5.3.1.2}$$

5.3.2. Frobenius Curvature and Geodesic Curvature of $\mathcal{D}$ and $\mathcal{D}^\perp$ for the Rolling Disc

We will begin by calculating the Frobenius curvature and the geodesic curvature of $\mathcal{D}^\perp$, starting with the definitions from section 2. For these calculations $X \in \mathcal{D}$, i.e., $w^3, w^4 = 0$. We calculate

$$\begin{aligned}
F_{\mathcal{D}}(X, \gamma') &= S_D(\gamma', X) - S_{\mathcal{D}}(X, \gamma') \\
&= -\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}^\perp}(\gamma') + \mathcal{P}_{\mathcal{D}^\perp} \left(\overset{\mathbb{G}}{\nabla}_X \gamma' \right) + \overset{\mathbb{G}}{\nabla}_{\gamma'} \mathcal{P}_{\mathcal{D}^\perp}(X) - \mathcal{P}_{\mathcal{D}^\perp} \left(\overset{\mathbb{G}}{\nabla}_{\gamma'} X \right).
\end{aligned}$$

We notice immediately that both X and γ' are elements in $\mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D})$, this means that their projections onto $\mathcal{D}^\perp$ are zero. Thus

$$
\begin{aligned}
F_{\mathcal{D}}(X, \gamma') &= -\cancel{\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}^\perp}(\gamma')} + \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_X \gamma'\right) + \cancel{\overset{\mathbb{G}}{\nabla}_{\gamma'} \mathcal{P}_{\mathcal{D}^\perp}(X)} - \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'} X\right) \\
&= w^2 v^1 \mathcal{P}_{\mathcal{D}^\perp}\left(\frac{\partial}{\partial\theta} X_1\right) - v^2 w^1 \mathcal{P}_{\mathcal{D}^\perp}\left(\frac{\partial}{\partial\theta} X_1\right),
\end{aligned}
$$

$$
\begin{aligned}
G_{\mathcal{D}}(\gamma', \gamma') &= 2S_{\mathcal{D}}(\gamma', \gamma') \\
&= -2\left(\cancel{\overset{\mathbb{G}}{\nabla}_{\gamma'} (\mathcal{P}_{\mathcal{D}^\perp}(\gamma'))} - \mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'} \gamma'\right)\right) \\
&= 2\left(\mathcal{P}_{\mathcal{D}^\perp}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'} \gamma'\right)\right) \\
&= 2\mathcal{P}_{\mathcal{D}^\perp}\left(\dot{v}^1 X_1 + \dot{v}^2 X_2 + v^1 \dot{X}_1\right) \\
&= 2v^1 \mathcal{P}_{\mathcal{D}^\perp}(\dot{X}_1).
\end{aligned}
$$

Now we will compute the geodesic curvature of $\mathcal{D}^\perp$ applied to $X, \lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}^\perp)$,

$$
\begin{aligned}
G_{\mathcal{D}^\perp}(X, \lambda) &= S_{\mathcal{D}^\perp}(X, \lambda) + S_{\mathcal{D}^\perp}(X, \lambda) \\
&= -\cancel{\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}}(\lambda)} + \mathcal{P}_{\mathcal{D}}\left(\overset{\mathbb{G}}{\nabla}_X \lambda\right) - \cancel{\overset{\mathbb{G}}{\nabla}_\lambda \mathcal{P}_{\mathcal{D}}(X)} + \mathcal{P}_{\mathcal{D}}\left(\overset{\mathbb{G}}{\nabla}_\lambda X\right).
\end{aligned}
$$

From section 5.3.1, $\overset{\mathbb{G}}{\nabla}_X \lambda$ and $\overset{\mathbb{G}}{\nabla}_\lambda X \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}^\perp)$ so

$$
G_{\mathcal{D}^\perp}(X, \lambda) = 0.
$$

The calculation for the Frobenius curvature of $\mathcal{D}^\perp$ applied to $X, \lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}^\perp)$ is similar to the calculation of geodesic curvature above,

$$
\begin{aligned}
F_{\mathcal{D}^\perp}(X, \lambda) &= S_{\mathcal{D}^\perp}(X, \lambda) - S_{\mathcal{D}^\perp}(X, \lambda) \\
&= -\cancel{\overset{\mathbb{G}}{\nabla}_X \mathcal{P}_{\mathcal{D}}(\lambda)} + \cancel{\mathcal{P}_{\mathcal{D}}\left(\overset{\mathbb{G}}{\nabla}_X \lambda\right)} + \cancel{\overset{\mathbb{G}}{\nabla}_\lambda \mathcal{P}_{\mathcal{D}}(X)} - \cancel{\mathcal{P}_{\mathcal{D}}\left(\overset{\mathbb{G}}{\nabla}_\lambda X\right)} \\
&= 0.
\end{aligned}
$$

The $\mathbb{G}$-transpose of $F_{\mathcal{D}}(X, \gamma')$, $F^*_{\mathcal{D}}(\gamma')(\lambda)$, from definition 2.4.1 We will present the calculation as a lemma.

Lemma 5.3.1. *Let $\Sigma = (M, \mathbb{G}, V, \mathcal{D})$ be the rolling disc mechanical system. The $\mathbb{G}$-transpose, $F^*_{\mathcal{D}}$, of the Frobenius curvature tensor is given by*

$$
F^{*\,j}_{\mathcal{D}\,lk} = F_{\mathcal{D}\,jk}^{\;\;l}.
$$

Proof. Note that, for this calculation, $X \in \mathcal{D}$ and all indices are summed over $\{1,2\}$. Starting with definition 2.4.1,

$$\begin{aligned}
\mathbb{G}(\lambda, F_{\mathcal{D}}(X,\gamma')) &= \mathbb{G}(F^*_{\mathcal{D}}(\gamma')(\lambda), X)\\
\mathbb{G}(u^l X_{l+2}, F_{\mathcal{D}}(w^j X_j, v^k X_k)) &= \mathbb{G}(F^*_{\mathcal{D}}(v^k X_k)(u^l X_{l+2}), w^j X_j)\\
\mathbb{G}(u^l X_{l+2}, F_{\mathcal{D}}{}^i_{jk} w^j v^k X_i) &= \mathbb{G}(F^*_{\mathcal{D}}{}^i_{lk} v^k u^l X_i, w^j X_j)\\
F_{\mathcal{D}}{}^i_{jk} \cancel{w^j v^k u^l} \mathbb{G}(X_{l+2}, X_i) &= F^*_{\mathcal{D}}{}^i_{lk} \cancel{v^k u^l w^j} \mathbb{G}(X_i, X_j)\\
F_{\mathcal{D}}{}^i_{jk} \mathbb{G}_{(l+2)i} \underset{i(l+2)}{\delta} &= F^*_{\mathcal{D}}{}^i_{lk} \mathbb{G}_{ij} \underset{ij}{\delta}\\
F_{\mathcal{D}}{}^l_{jk} \cancel{\mathbb{G}_{ii}} &= F^*_{\mathcal{D}}{}^j_{lk} \cancel{\mathbb{G}_{ii}}.
\end{aligned}$$

$\square$

To write the final equations in a more reader friendly format, calculate

$$\begin{aligned}
\frac{\partial}{\partial\theta}(X_1) &= \frac{-2J_{\text{roll}} m\rho\sqrt{\frac{1}{m} + \frac{\rho^2\cos^2\theta}{J_{\text{roll}}}}\sin\theta}{\sqrt{J_{\text{roll}} + m\rho^2}(2J_{\text{roll}} + m\rho^2\cos^2\theta)} X_3 + \frac{\sqrt{2} m\rho\cos\theta\sqrt{\frac{J_{\text{roll}}+m\rho^2}{m(2J_{\text{roll}}+m\rho^2\cos^2\theta)}}}{\sqrt{J_{\text{roll}} + m\rho^2}} X_4,\\
\frac{\partial}{\partial\theta}(X_2) &= 0,\\
\frac{\partial}{\partial\theta}(X_3) &= \frac{\rho\sin(\theta)}{\sqrt{J_{\text{roll}} + m\rho^2}\sqrt{\frac{1}{m} + \frac{\rho^2\cos^2\theta}{J_{\text{roll}}}}} X_1 - \frac{2m\rho^2\sqrt{\frac{J_{\text{roll}}+m\rho^2}{m(2J_{\text{roll}}+m\rho^2\cos^2\theta)}}\sin^2\theta}{(J_{\text{roll}} + m\rho^2)\sqrt{\frac{2J_{\text{roll}}+m\rho^2\cos^2\theta}{J_{\text{roll}} m}}} X_4,\\
\frac{\partial}{\partial\theta}(X_4) &= \frac{-m\rho\cos\theta}{\sqrt{J_{\text{roll}} m + m^2\rho^2\cos^2\theta}} X_1 + \frac{J_{\text{roll}} m\rho^2\sqrt{\frac{1}{m} + \frac{\rho^2\cos^2\theta}{J_{\text{roll}}}}\sin^2\theta}{(J_{\text{roll}} + m\rho^2\cos^2\theta)^2\sqrt{\frac{J_{\text{roll}}+m\rho^2}{J_{\text{roll}} m+m^2\rho^2\cos^2\theta}}} X_3,
\end{aligned}$$

and define $f^1(\theta),\ f^2(\theta),\ f^3(\theta),\ f^4(\theta)$ so that

$$\begin{aligned}
\frac{\partial}{\partial\theta}(X_1) &= f^1(\theta)X_3 + f^2(\theta)X_4,\\
\mathcal{P}_{\mathcal{D}^\perp}\left(\frac{\partial}{\partial\theta}(X_3)\right) &= f^3(\theta)X_4,\\
\mathcal{P}_{\mathcal{D}^\perp}\left(\frac{\partial}{\partial\theta}(X_4)\right) &= f^4(\theta)X_3.
\end{aligned}$$

We can now calculate $F^*_{\mathcal{D}}(\gamma')(\lambda)$ using Lemma 5.3.1:

$$\begin{aligned}
F_D &= dx^2 \otimes dx^1(f^1(\theta)X_3 + f^2(\theta)X_4) - dx^1 \otimes dx^2(f^1(\theta)X_3 + f^2(\theta)X_4)\\
&= dx^2 \otimes dx^1 f^1(\theta)X_3 + dx^2 \otimes dx^1 f^2(\theta)X_4 - dx^1 \otimes dx^2 f^1(\theta)X_3 - dx^1 \otimes dx^2 f^2(\theta)X_4,\\
F^*_{\mathcal{D}} &= dx^3 \otimes dx^1 f^1(\theta)X_2 + dx^4 \otimes dx^1 f^2(\theta)X_2 - dx^3 \otimes dx^2 f^1(\theta)X_1 - dx^4 \otimes dx^2 f^2(\theta)X_1,\\
F^*_{\mathcal{D}}(\gamma')(\lambda) &= u^1v^1f^1(\theta)X_2 + u^2v^1f^2(\theta)X_2 - u^1v^2f^1(\theta)X_1 - u^2v^2f^2(\theta)X_1\\
&= -(u^1v^2f^1(\theta) + u^2v^2f^2(\theta))X_1 + (u^1v^1f^1(\theta) + u^2v^1f^2(\theta))X_2.
\end{aligned}$$

5.3.3. Differential Equations Characterizing Constrained Variational Trajectories for the Rolling Disc

With the objects that we have computed in the previous two subsections, we can write equation (5.1.0.2) explicitly in terms of u^1, u^2, v^1 and v^2, the components of γ' and λ in the orthonormal basis. Equation (5.1.0.2) becomes

$$\begin{aligned} \dot{v}^1 &= -(u^1 v^2 f^1(\theta) + u^2 v^2 f^2(\theta)), \\ \dot{v}^2 &= (u^1 v^1 f^1(\theta) + u^2 v^1 f^2(\theta)). \end{aligned}$$

In matrix form this becomes

$$\begin{bmatrix} \dot{v}^1 \\ \dot{v}^2 \end{bmatrix} = \begin{bmatrix} -v^2 f^1(\theta) & -v^2 f^2(\theta) \\ v^1 f^1(\theta) & v^1 f^2(\theta) \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}. \tag{5.3.3.1}$$

We can also write equation (5.1.0.1) explicitly. First we write $\overset{\mathcal{D}^\perp}{\nabla}_{\gamma'}\lambda$ as linear combinations in $\mathscr{X}$:

$$\begin{aligned} \overset{\mathcal{D}^\perp}{\nabla}_{\gamma'}\lambda &= \dot{u}^1 X_3 + u^1\dot{\theta}\frac{\partial}{\partial\theta}(X_3) + \dot{u}^2 X_4 + u^2\dot{\theta}\frac{\partial}{\partial\theta}(X_4) - \mathcal{P}_{\mathcal{D}}\left(u^1\dot{\theta}\frac{\partial}{\partial\theta}(X_3) + u^2\dot{\theta}\frac{\partial}{\partial\theta}(X_4)\right) \\ &= \dot{u}^1 X_3 + \dot{u}^2 X_4 + u^1\dot{\theta}\left(\frac{\partial}{\partial\theta}(X_3) - \mathcal{P}_{\mathcal{D}}\left(\frac{\partial}{\partial\theta}(X_3)\right)\right) + u^2\dot{\theta}\left(\frac{\partial}{\partial\theta}(X_4) - \mathcal{P}_{\mathcal{D}}\left(\frac{\partial}{\partial\theta}(X_4)\right)\right) \\ &= \dot{u}^1 X_3 + \dot{u}^2 X_4 + u^1\dot{\theta}\mathcal{P}_{\mathcal{D}^\perp}\left(\frac{\partial}{\partial\theta}(X_3)\right) + u^2\dot{\theta}\mathcal{P}_{\mathcal{D}^\perp}\left(\frac{\partial}{\partial\theta}(X_4)\right) \\ &= \dot{u}^1 X_3 + \dot{u}^2 X_4 + u^1\dot{\theta} f^3(\theta) X_4 + u^2\dot{\theta} f^4(\theta) X_3. \end{aligned}$$

Substituting the above calculation and the necessary objects from section 5.3.2 into equation (5.1.0.1), we get

$$\begin{aligned} \dot{u}^1 X_3 + \dot{u}^2 X_4 + u^1 f^3(\theta) X_4 + u^2 f^4(\theta) X_3 &= v^1\dot{\theta}\left(f^1(\theta) X_3 + f^2(\theta) X_4\right) \\ \dot{u}^1 X_3 + \dot{u}^2 X_4 &= \dot{\theta}\left(v^1 f^1(\theta) - u^2 f^4(\theta)\right) X_3 + \left(v^1 f^2(\theta) - u^1 f^3(\theta)\right) X_4, \end{aligned}$$

which simplifies to the differential equations

$$\begin{aligned} \dot{u}^1 &= \dot{\theta}\left(v^1 f^1(\theta) - u^2 f^4(\theta)\right), \\ \dot{u}^2 &= \dot{\theta}\left(v^1 f^2(\theta) - u^1 f^3(\theta)\right). \end{aligned} \tag{5.3.3.2}$$

Chapter 6

Comparing Variational and Non-Holonomic Trajectories for the Rolling Disc

6.1. When Do Nonholonomic and Variational Trajectories Agree for the Rolling Disc?

In the previous section we derived two sets of equations describing the variational trajectories of the rolling disc. We will now compare those to equation (5.2.0.4),

$$\dot{v}^1 X_1 + \dot{v}^2 X_2,$$

to determine when variational trajectories of the rolling disc are also nonholonomic trajectories and to show some of their characteristics. We summarize the results in the following theorem.

Theorem 6.1.1. *Let $\Sigma = (M, \mathbb{G}, V, \mathcal{D})$ be the rolling disc mechanical system. Let $t_0, t_1 \in \mathbb{R}$ with $t_0 < t_1$ and let $x_0, x_1 \in M$. For $\gamma \in \mathrm{H}^1([t_0, t_1]; M; \mathcal{D}; x_0, x_1)$, the following statements are true:*

(i) If γ is a nonholonomic trajectory then it is a variational trajectory;

(ii) Given $v_{x_0} \in TM$, there exists a one-dimensional affine subspace $B \subset D^{\perp}_{x_0}$ such that, if $\lambda(t_0) \in B$, then the constrained variational trajectory $t \mapsto (\gamma(t), \lambda(t))$ is such that γ is a nonholonomic trajectory.

Proof. (i) Suppose that γ is a nonholonomic trajectory of Σ. It is a variational trajectory if there exists a $\lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}^{\perp})$ satisfying equations (5.3.3.1) and (5.3.3.2). Write $\gamma' = v^1 X_1 + v^2 X_2$ and $\lambda = u^1 X_3 + u^2 X_4$. Equation (5.3.3.1) is satisfied if λ is in the kernel of

$$\begin{bmatrix} -v^2 f^1(\theta) & -v^2 f^2(\theta) \\ v^1 f^1(\theta) & v^1 f^2(\theta) \end{bmatrix}.$$

It is clear upon inspection that λ is in the kernel of this matrix only when

$$u^1 = -u^2 \frac{f^2(\theta)}{f^1(\theta)}. \tag{6.1.0.1}$$

To determine when a λ satisfying equation (6.1.0.1) also satisfies equation (5.3.3.2), we will substitute (6.1.0.1) into (5.3.3.2) to get

$$\dot{u}^1 = u^1 \frac{f^1(\theta)}{f^2(\theta)} f^4(\theta)\dot{\theta} + v^1 f^1(\theta)\dot{\theta}, \tag{6.1.0.2}$$

$$\dot{u}^2 = u^2 \frac{f^2(\theta)}{f^1(\theta)} f^3(\theta)\dot{\theta} + v^1 f^2(\theta)\dot{\theta}. \tag{6.1.0.3}$$

We will now solve these equations for u^1 and u^2. Let $a_1 = \frac{f^1(\theta)}{f^2(\theta)} f^4(\theta)\dot{\theta}, a^2 = \frac{f^2(\theta)}{f^1(\theta)} f^3(\theta)\dot{\theta}$ and let $\hat{a}^1$ and $\hat{a}^2$ be the antiderrivatives of a^1 and a^2. We calculate

$$\begin{aligned}
\left(\dot{u}^1 - u^1 a_1\right) e^{-\hat{a}_1} &= v^1 f^1 \dot{\theta} e^{-\hat{a}_1} \\
\frac{d}{dt}\left(u^1 e^{-\hat{a}_1}\right) &= v^1 f^1 \dot{\theta} e^{-\hat{a}_1} \\
\int \frac{d}{dt}\left(u^1 e^{\hat{a}_1}\right) dt &= \int v^1 f^1 \dot{\theta} e^{\hat{a}_1} dt \\
u^1 &= e^{\hat{a}_1} \int v^1 f^1 \dot{\theta} e^{-\hat{a}_1} dt.
\end{aligned}$$

Note that we have assumed $\dot{v}^1 = 0$, so v^1 is a constant and

$$\begin{aligned}
u^1 &= e^{\hat{a}_1} v^1 \int f^1 \dot{\theta} e^{-\hat{a}_1} dt. \\
&= e^{\hat{a}_1} v^1 \int f^1 e^{-\hat{a}_1} d\theta.
\end{aligned}$$

Now, let $F^1(\theta) + C_1$ be the antiderrivative of $f^1 e^{-\hat{a}_1}$. We integrate to get

$$u^1 = e^{\hat{a}_1} v^1 (F^1(\theta) + C_1). \tag{6.1.0.4}$$

By the same process

$$u^2 = e^{\hat{a}_2} v^1 \left(F^2(\theta) + C_2\right), \tag{6.1.0.5}$$

with $F^2(\theta)$ defined similarly to $F^1(\theta)$.

In co-ordinates equations (6.1.0.4) and (6.1.0.5) look like

$$\begin{aligned}
u^1 &= v^1 \frac{\cos\theta^{\frac{m\rho^2}{J_{\mathrm{roll}}+m\rho^2}} C_1 + \rho\sqrt{\frac{2m(J_{\mathrm{roll}}+m\rho^2)}{J_{\mathrm{roll}}}}\cos\theta}{\sqrt{2J_{\mathrm{roll}} + m\rho^2\cos^2\theta}}, \text{ and} \\
u^2 &= v^1 \frac{C_2 + \rho\sqrt{2m}\sin\theta}{\sqrt{2J_{\mathrm{roll}} + m\rho^2\cos^2\theta}}.
\end{aligned} \tag{6.1.0.6}$$

We can see from these equations that there is a well defined λ for any nonholonomic trajectory γ. Thus, all nonholonomic trajectories of Σ are constrained variational trajectories of Σ.

(ii) Suppose that γ is a constrained variational trajectory of Σ. In order for it to be a nonholonomic trajectory, there must be a $\lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D})$ that satisfies equations (6.1.0.4), (6.1.0.5) and (6.1.0.1). Plugging equations (6.1.0.4) and (6.1.0.5) into equation (6.1.0.1) gives us

$$\begin{aligned} e^{\hat{a}_2}\left(F^2(\theta)+C_2\right) &= -e^{\hat{a}_1}\left(F^1(\theta)+C_1\right)\frac{f^1(\theta)}{f^2(\theta)} \\ F^2(\theta)+C_2 &= -e^{\hat{a}_1-\hat{a}_2}\left(F^1(\theta)+C_1\right)\frac{f^1(\theta)}{f^2(\theta)} \\ C_2 &= -e^{\hat{a}_2-\hat{a}_1}\left(F^1(\theta)+C_1\right)\frac{f^1(\theta)}{f^2(\theta)} - F^2(\theta). \end{aligned} \tag{6.1.0.7}$$

Similarly

$$C_1 = -e^{\hat{a}_1-\hat{a}_2}\left(F^2(\theta)+C_2\right)\frac{f^2(\theta)}{f^1(\theta)} - F^2(\theta). \tag{6.1.0.8}$$

This form hints toward a situation that we will need to be cognizant of, namely that C_1 and C_2 may be undefined for x_0 such that either f^1 or f^2 are zero; this happens when $\theta(t_0) = 0$ and $\theta(t_0) = \pm\frac{\pi}{2}$, respectively. We cannot say for sure whether or not these initial conditions pose a problem without inspecting equations (6.1.0.7) and (6.1.0.8) in co-ordinate form. We see that

$$\begin{aligned} C_1 = {} & \frac{-2\rho\sqrt{J_{\mathrm{roll}}m(J_{\mathrm{roll}}+m\rho^2)}\cos\theta^{\frac{J_{\mathrm{roll}}}{J_{\mathrm{roll}}+m\rho^2}}}{\sqrt{J_{\mathrm{roll}}+m\rho^2\cos^2\theta}} \\ & - C^2\frac{\sqrt{J_{\mathrm{roll}}+m\rho^2}\cos\theta^{\frac{J_{\mathrm{roll}}}{J_{\mathrm{roll}}+m\rho^2}}}{\sqrt{2}J_{\mathrm{roll}}\sqrt{\frac{J_{\mathrm{roll}}+m\rho^2\cos^2\theta}{J_{\mathrm{roll}}(2J_{\mathrm{roll}}+\cos^2\theta)}}}, \\ C_2 = {} & \frac{1}{\frac{1}{\sqrt{2}}\sqrt{\frac{J_{\mathrm{roll}}+m\rho^2}{m(2J_{\mathrm{roll}}+m\rho^2 cos^2\theta)}}}\Bigg(2m\rho^3\sqrt{J_{\mathrm{roll}}+m\rho^2}\cos^3\theta \\ & + C^1 J_{\mathrm{roll}}\cos\theta^{\frac{m\rho^2}{J_{\mathrm{roll}}+m\rho^2}}\sqrt{\frac{1}{m}+\frac{\rho^2\cos^2\theta}{J_{\mathrm{roll}}}}\sqrt{2J_{\mathrm{roll}}+m\rho^2\cos^2\theta} \\ & + \rho\sqrt{J_{\mathrm{roll}}+m\rho^2}\cos\theta(4J_{\mathrm{roll}}+m\rho^2\cos^2\theta)\tan\theta\Bigg). \end{aligned}$$

As hideous as these expressions are, we can tell through some careful inspection that, as we suspected above, they are defined everywhere except for $\theta(t_0) = 0$ in the case of C_1; and $\theta(t_0) = \pm\frac{\pi}{2}$ in the case of C_2. From this we see that we must choose our constant based on the initial conditions for γ. For any initial condition such that $\theta(t_0) = 0$, $C_1 \to 0$ and we leave C_2 as a free parameter; similarly if $\theta(t_0) = \pm\frac{\pi}{2}$ then $C_1 \to 0$ and C_2 is left free. For any other x_0 either constant can be used. The purpose of all of these observations is to show that, when the correct constant is chosen and substituted into equations (6.1.0.6) we get a trajectory for λ such that γ is a nonholonomic trajectory for all time $t \in [t_0, t_1]$. We can then see that $\lambda(t_0)$ is an element of an affine subspace B of $\mathcal{D}^{\perp}_{x_0}$.

$\square$

Chapter 7

Conclusion and Interesting Next Steps

We have explored the difference between nonholonomic and constrained variational trajectories in detail from a geometric perspective through the example of a rolling disc. We proved that all nonholonomic trajectories for the rolling disc are also constrained variational trajectories; and we characterized which constrained variational trajectories are nonholonomic trajectories.

The rolling disc, while it is a satisfying and illustrative example of a mechanical system with nonholonomic constraints, is well trodden ground in the nonholonomic mechaincs literature and it is likely that all that is novel to say about it has been said. The true power of the theorems in *Nonholonomic and constrained variational mechanics* will be shown with less standard examples. One place to start looking for novel examples may be with systems that do not have defined $\lambda \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathcal{D}^\perp)$, satisfying Theorem 5.1.2, for every nonholonomic trajectory. Some previously explored examples of such systems are the nonholonomic free particle and the Chaplygin sphere, analysed in in Fernandez and Bloch [2008, section 6]; it would be interesting to generate other examples that are not Chaplygin. The characterization of nonholonomic trajectories that are also constrained variational trajectories in Lewis [2018, Theorem 7.7], may give clues as to how such examples may be found. It would also be interesting to try and generate examples such that there is no overlap between the nonholonomic trajectories and the constrained variational trajectories.

Bibliography

[1] Borisov, A. V., Mamaev, I. S., and Bizayev, I. A. [2017] *Dynamical systems with non-integrable constraints, vakanomic mechanics, sub-Riemannian geometry, and nonholonomic mechanics,* Russian Mathematical Surveys. **72**(5), pages 783-840.

[2] Boyd, C. [2018] *Nonholonomic and variational equations of motion for mechanical systems*, Masters thesis, Queen's University, Kingston, August 2018.

[3] Bullo, F. and Lewis, A. D. [2005] *Geometric Control of Mechanical Systems, Modelling, Analysis, and Design for Simple Mechanical Systems*, number 49 in Texts in Applied Mathematics, Springer Science+Business Media, Inc.

[4] Cardin, F. and Favretti, M. [1996] *On nonholonomic and vakanomic dynamics of mechanical systems with nonintegrable constraints,* Journal of Geometry and Physics, **18**(4), pages 295-235.

[5] Cortés, J., de León, M., Martín de Diego, D., and Martínez, S. [2002] *Geometric description of vakanomic and nonholonomic dynamics. Comparison of solutions,* SIAM Journal on Control and Optimization, **41**(5), pages 1389-1412.

[6] Crampin, M. and Mestdag, T. [2010] *Anholonomic frames in constrained dynamics,* Dynamical Systems, An International Journal, **25**(2), pages 159-187.

[7] Favretti, M. [1998] *Equivalence of dynamics for nonholonomic systems with transverse constraints,* Journal of Dynamics and Differential Equations, **10**(4), pages 511-536.

[8] Fernandez, O. E. and Bloch, A. M. [2008] *Equivalence of the dynamics of nonholonomic and variational nonholonomic systems for certain initial data,* Journal of Physics. A. Mathematical and Theoretical. **41**(34), page 344005.

[9] Gràcia, X., Marin–Solano, J., and Muñoz–Lecanda, M.C.[2003] *Some geometric aspects of variational calculus in constrained systems,* Reports on Mathematical Physics, **51**(1), pages 127-148.

[10] Jóźwikowski, M. and Respondek, W. [2018] *A comparison of vakonomic and nonholonomic dynamics with applications to non-invariant Chaplygin systems,* Journal of Geometric Mechanics, to appear.

[11] Kharlomov, P. V. [1992] *A critique of some mathematical models of mechanical systems with differential constraints,* Journal of Applied Mathematics and Mechanics, **56**(4), pages 584-594.

[12] Kozlov,V. V. [1992] *The problem of realizing constraints in dynamics,* Journal of Applied Mathematics and Mechanics, **56**(4), pages 594-600.

[13] Kupka, I. and Oliva, W. M. [2001] *The nonholonomic mechanics,* Journal of Differential Equations, **169**(1), pages 169-189.

[14] Lewis, A. D. [2018] *Nonholonomic and constrained variational mechanics,* preprint, available online at `https://mast.queensu.ca/ andrew/papers/abstracts/2018a.html`

[15] Lewis, A. D. and Murray, R. M. [1995] *Variational principles for constrained systems: Theory and experiment,* International Journal of Non-Linear Mechanics, **30**(6), pages 793-815.

[16] Rumiantsev, V. V. [1978] *On Hamilton's principle for nonholonomic systems,* Journal of Applied Mathematics and Mechanics, **42**(3), pages387-399.

[17] Terra, G. [2018] *Vakonomic versus nonholonomic mechanics revisited,* São Paulo Journal of Mathematical Science, University of São Paulo, São Paulo, **12**(1), pages 136-145.

[18] Vershik, A. M. and Gershkovich, V. Y. [1990] *Nonholonomic dynamical systems, geometry of distributions and variational problems*, in *Dynamical Systems,* edited by V. I. Arnol'd and S. Novikov, Volume 7. Encyclopedia of Mathematical Sciences, pages 4-79, Springer-Verlag: New York/Heidelberg/Berlin.

[19] Zampieri, G. [2000] *Nonholonomic versus vakonomic mechanics,* Journal of Differential Equations, **163**(2), pages 335-347.