Linearization and Stability of Nonholonomic Mechanical Systems

Steven D. Yang[1]

2015/06/15

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

Abstract

The stability of an equilibrium point of a nonlinear system is typically analyzed in two ways: (1) stability of its linearization, and (2) Lyapunov stability. An unconstrained simple mechanical system is a type of nonlinear system with a special structure, and so the methods for stability analysis can be specialized for this particular class of nonlinear systems. For a simple mechanical system subject to velocity constraints, the situation becomes more complicated. If the constraints are holonomic, then the problem can simply be reduced to that of an unconstrained simple mechanical system by restricting analysis to a certain submanifold of the configuration space. If the constraints are nonholonomic, this approach cannot be taken. In this report we study the differences and additional complexities that arise in these nonholonomic mechanical systems, and derive results with regards to linearization and stability of its equilibria.

Contents

Chapter 1

Introduction

In this report, we study the linearization and stability of nonholonomic mechanical systems at equilibria. Herein, we will refer to mechanical systems without constraints on velocity as *unconstrained mechanical systems*, and those with velocity constraints as *constrained mechanical systems*. A *nonholonomic mechanical system* is a constrained mechanical system in which the velocity constraints are not integrable (and so are called "nonholonomic constraints"). Nonholonomic constraints arise naturally, for example, in the rolling of a wheel, sliding of skates, or in cases where angular momentum is conserved, such as with a space robot or satellite in orbit.

Nonholonomic mechanical systems have been explored throughout the literature; see [12] for a classic overview of the kinematics, dynamics, linearization, and stability of these systems. Investigations into stability via linearization analysis can be found in [9], [13], and [5], and studies involving Lyapunov stability include [1] and [8]. With regards to linearization, there has been much confusion on how to approach the problem due to its complexity. In the early work of [14], Whittaker tackled the problem by linearizing the constraint equations in addition to the equations of motion containing unsolved Lagrange multipliers, i.e. the equations are linearized before the Lagrange multipliers are actually solved for. In doing so, he claimed that the nonholonomic constraints give rise to holonomic ones, and therefore concluded that the integrability of constraints plays no role in stability analysis. Some maintain that this approach is correct while others disagree. The works of [2] and [12], for example, attempt to fix this approach while other works such as [4] appear to argue otherwise. In [6], DeMarco presents a thorough approach to the correct linearization. It turns out that much of what we derive here has already been explored in [6]. One key difference is that we now incorporate everything into our geometric mechanics framework, following [3], and use differential geometric methods to obtain the results in a cleaner and more concise way. We also examine the concept of linearization more closely and attempt to apply Lyapunov methods for stability.

This report is organized as follows. In Chapter 2, we review the equations of motion, linearization, and stability of equilibria of unconstrained mechanical systems. Laying down these results first provides for a useful comparison when we later study constrained mechanical systems. In Chapter 3, we start looking at constrained mechanical systems. Our main interest is in mechanical systems in which the constraints are nonholonomic. We study the equations of motion and equilibria for these systems before introducing a more general framework, that of forced affine connection systems. In Chapter 4, we look at how to linearize the equations of motion for nonholonomic mechanical systems. We do this by first looking at the linearization of forced affine connection systems. We later discuss the alternative linearization approach which early researchers have used and point out its shortcomings. Finally, in Chapter 5, we attempt to make statements about the stability of equilibria of nonholonomic systems using the linearization approach as well as Lyapunov methods. Chapter 6 summarizes the report and discusses possible future work.

Chapter 2

Unconstrained Mechanical Systems

We begin by looking at unconstrained mechanical systems, i.e. simple mechanical systems without any constraints on velocity. Reviewing the results for linearization and stability of unconstrained mechanical systems here will be useful for comparison when we later study the constrained case.

Definition 2.1 (Forced Simple Mechanical System)**.** A C^∞**-forced simple mechanical system** is a 4-tuple $(Q, \mathbb{G}, V, F)$, where
(i) Q is a C^∞-manifold (called the **configuration manifold**),
(ii) $\mathbb{G}$ is a C^∞-Riemannian metric on Q (called the **kinetic energy metric**),
(iii) $V \in C^\infty(Q)$ (called the **potential function**), and
(iv) $F : TQ \to T^*Q$ is a C^∞-vector bundle map over id_Q (called the **Lagrangian force**).

Note that in our above definition, F is time-independent. We do not consider time-dependent Lagrangian forces here. Also note that this is our general notion of an unconstrained mechanical system; there are no constraints on velocity.

2.1. Equations of Motion and Equilibria

The main proposition in this section (Proposition 2.4) gives us the equations of motion for a forced simple mechanical system. We refer to [3] for a detailed treatment of the equations of motion, in particular how they relate to the Euler-Lagrange equations and the motion of interconnected mechanical systems.

In order to present the equations of motion precisely, we review the Lagrange-d'Alembert Principle.

Definition 2.2 (Variation of a Curve)**.** Let Q be a C^∞-manifold and consider a C^2-curve $\gamma : [a, b] \to Q$. A **variation of** γ is a C^2-map $\nu : J \times [a, b] \to Q$ such that
(i) $J \subseteq \mathbb{R}$ is an interval with $0 \in \mathrm{int}(J)$,
(ii) $\nu(0, t) = \gamma(t)$ for all $t \in [a, b]$,
(iii) $\nu(s, a) = \gamma(a)$ for all $s \in J$, and
(iv) $\nu(s, b) = \gamma(b)$ for all $s \in J$.
The **infinitesimal variation associated with** ν is the vector field $\delta\nu$ along γ such that for all $t \in [a, b]$, given a local chart (U, ϕ) of Q around $\gamma(t)$ and the corresponding chart

$(TU, T\phi)$ of TQ, we have

$$T\phi \circ \delta\nu(t) = \left(\phi \circ \nu(0,t), \frac{d}{ds}\Big|_{s=0} \phi \circ \nu(s,t)\right).$$

Note that this is well-defined. With the use of local charts implied, we can write this in a compact form

$$\delta\nu(t) = \frac{d}{ds}\Big|_{s=0} \nu(s,t).$$

Definition 2.3 (Lagrange-d'Alembert Principle)**.** Let L be a C^∞-Lagrangian on a C^∞-manifold Q, and let $F : TQ \to T^*Q$ be a (time-independent) C^∞-force. A C^2-curve $\gamma : [a,b] \to Q$ satisfies the **Lagrange-d'Alembert Principle for the force** F **and Lagrangian** L if, for every variation $\nu : J \times [a,b] \to Q$ of γ, we have

$$\frac{d}{ds}\Big|_{s=0} \int_a^b L\Big(t, \frac{d}{dt}\nu(s,t)\Big) dt + \int_a^b \langle F(\gamma'(t)); \delta\nu(t)\rangle dt = 0,$$

where the use of local charts is implied.

The equations of motion can now be summarized as follows.

Proposition 2.4 (Equations of Motion)**.** *Let $(Q, \mathbb{G}, V, F)$ be a C^∞-forced simple mechanical system. A curve $\gamma : I \to Q$ satisfies the Lagrange-d'Alembert Principle for the force F and the Lagrangian $L(v_q) = \frac{1}{2}\mathbb{G}(v_q, v_q) - V(q)$ if and only if γ satisfies the differential equation*

$$\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -\mathit{grad}V(\gamma(t)) + \mathbb{G}^\# \circ F(\gamma'(t)), \tag{2.1.1}$$

where $\mathit{grad}V \triangleq \mathbb{G}^\# \circ dV$.

Let us recall the coordinate expressions for the equations of motion. Consider a C^∞-forced simple mechanical system $(Q, \mathbb{G}, V, F)$. For a local chart $(U, \phi = (q^1, \ldots, q^n))$ of Q, the equations of motion (2.1.1) in coordinates are

$$\ddot{q}^i(t) + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q(t))\dot{q}^j(t)\dot{q}^k(t) = -G^{ij}(q(t))\frac{\partial V}{\partial q^j}(q(t)) + G^{ij}(q(t))F_j(q(t), \dot{q}(t)),$$

where $\phi \circ \gamma(t) = q(t) = (q^1(t), \ldots, q^n(t))$ for admissible $t \in I$, $\overset{\mathbb{G}}{\Gamma}{}^i_{jk}$ are the Christoffel symbols for $\overset{\mathbb{G}}{\nabla}$ in the local chart (U, ϕ), $\mathbb{G}(q) = G_{ij}(q)dq^i(q) \otimes dq^j(q)$, $\left[G^{ij}\right] = \left[G_{ij}\right]^{-1}$ and $F(v_q) = F_i(v_q)dq^i(q)$. With the time dependence and function arguments implied, this can be written in a compact form

$$\ddot{q}^i + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}\dot{q}^j\dot{q}^k = -G^{ij}\frac{\partial V}{\partial q^j} + G^{ij}F_j.$$

For studying the equilibria of forced simple mechanical systems, we start with a basic definition.

Definition 2.5 (Equilibrium Configuration)**.** Let $\Sigma = (Q, \mathbb{G}, V, F)$ be a C^∞-forced simple mechanical system. A point $q_0 \in Q$ is an **equilibrium configuration for** Σ if the trivial curve $\gamma(t) = q_0$ satisfies the equations of motion (2.1.1).

Let $\Sigma = (Q, \mathbb{G}, V, F)$ be a C^∞-forced simple mechanical system and consider a local chart $(U, \phi = (q^1, \ldots, q^n))$ of Q. For a curve $\gamma : I \to Q$ satisfiying the equations of motion (2.1.1), let $\phi \circ \gamma(t) = q(t) = (q^1(t), \ldots, q^n(t))$ (again, for admissible $t \in I$; this assumption will be made throughout). The equations of motion in coordinates are

$$\ddot{q}^i + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}\dot{q}^j\dot{q}^k = -G^{ij}\frac{\partial V}{\partial q^j} + G^{ij}F_j,$$

or

$$\begin{cases} \dot{q}^i = v^i \\ \dot{v}^i = -\overset{\mathbb{G}}{\Gamma}{}^i_{jk}v^j v^k - G^{ij}\dfrac{\partial V}{\partial q^j} + G^{ij}F_j, \end{cases}$$

where $v(t) = (v^1(t), \ldots, v^n(t)) = \dot{q}(t)$. It can be shown that this gives rise to a well-defined vector field on TQ, i.e. $X : TQ \to TTQ$ defined by

$$X(v_q) \triangleq \overset{\mathbb{G}}{S}(v_q) - \mathrm{vlft}(\mathrm{grad}V)(v_q) + \mathrm{vlft}(\mathbb{G}^\# \circ F)(v_q), \tag{2.1.2}$$

where the integral curves of X, projected onto Q (using the canonical projection), are curves satisfying the equations of motion (2.1.1). We will call X the **associated vector field for** Σ. In coordinates this can be written as

$$X\Big|_{TU}(v_q) = v^i\frac{\partial}{\partial q^i}(v_q) + \left(-\overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q)v^j v^k - G^{ij}(q)\frac{\partial V}{\partial q^j}(q) + G^{ij}(q)F_j(v_q)\right)\frac{\partial}{\partial v^i}(v_q),$$

where $(TU, T\phi = (q^1, \ldots, q^n, v^1, \ldots, v^n))$ is the corresponding local chart of TQ for (U, ϕ).

Now, an equilibrium configuration $q_0 \in Q$ for Σ gives rise to an equilibrium point 0_{q_0} for the vector field X. Stability notions can be made for the equilibrium configuration q_0 by reference to the equilibrium point 0_{q_0}. We will sometimes say "equilibrium" in reference to either an equilibrium configuration or an equilibrium point; to which should be clear by context.

The following lemma gives us simple, yet useful characterizations of equilibrium configurations. We denote the zero section of a vector bundle by $Z(\cdot)$ (see Appendix D).

Lemma 2.6 (Characterizations of Equilibrium Configurations)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F)$ be a C^∞-forced simple mechanical system. Then,*

(i) $q_0 \in Q$ is an equilibrium configuration for Σ if and only if $dV(q_0) = F(0_{q_0})$, and

(ii) if $F(Z(TQ)) = Z(T^*Q)$, the equilibrium configurations are exactly the critical points of V.

Proof. The proof is immediate by considering the equations of motion (2.1.1) and setting $\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = 0$. □

2.2. Linearization about Equilibria

Here we look at how to linearize the equations of motion (2.1.1) about an equilibrium configuration q_0. We do this by linearizing the associated vector field about the equilibrium point 0_{q_0}. Doing so will give us a forced linear mechanical system, defined as follows. We refer to [3] for additional details.

Definition 2.7 (Forced Linear Mechanical System)**.** A **forced linear mechanical system** is a 4-tuple $(V, M, K, (F_1, F_2))$, where

(i) V is a finite-dimensional $\mathbb{R}$-vector space,
(ii) M is an inner product on V,
(iii) $K \in \Sigma_2(V)$, and
(iv) $F_1, F_2 \in L(V; V^*)$.

Let $(V, M, K, (F_1, F_2))$ be a forced linear mechanical system. The state space of the system is $V \oplus V$ and the equations of motion are

$$\begin{pmatrix} \dot{q}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{id}_V \\ -M^\sharp \circ K^\flat + M^\# \circ F_1 & M^\# \circ F_2 \end{pmatrix} \begin{pmatrix} q(t) \\ v(t) \end{pmatrix},$$

or equivalently,

$$\ddot{q}(t) + M^\sharp \circ K^\flat(q(t)) = M^\# \circ F_1(q(t)) + M^\# \circ F_2(\dot{q}(t)).$$

Now we review a useful lemma for decomposing the state space of the linearization to a form matching that of a forced linear mechanical system.

Lemma 2.8 (Natural Isomorphism of $T_{0_{q_0}}TQ$)**.** *There is a natural isomorphism of $T_{0_{q_0}}TQ$ with $T_{q_0}Q \oplus T_{q_0}Q$.*

Proof. Passing through 0_{q_0} are two submanifolds of TQ: $Z(TQ)$ and $T_{q_0}Q$. We have $T_{0_{q_0}}(Z(TQ)) \cong T_{q_0}Q$ since $q \mapsto 0_q$ is a diffeomorphism of Q with $Z(TQ)$. Also, note that $v \mapsto \mathrm{vlft}_{0_{q_0}}(v)$ is an isomorphism of $T_{q_0}Q$ with $T_{0_{q_0}}(T_{q_0}Q)$. Now, since $T_{0_{q_0}}(Z(TQ)) \cap T_{0_{q_0}}(T_{q_0}Q) = \{0_{q_0}\}$, we conclude that $T_{0_{q_0}}TQ \cong T_{0_{q_0}}(Z(TQ)) \oplus T_{0_{q_0}}(T_{q_0}Q) \cong T_{q_0}Q \oplus T_{q_0}Q$. □

The tangent lift of a vector field is reviewed in Appendix C, but it will be useful to recall the coordinate expressions here as we will use it extensively. Let M be a C^∞-manifold and $X \in \Gamma^\infty(TM)$. Consider a local chart $(U, \phi = (x^1, \dots, x^n))$ of M and the corresponding local chart $(TU, T\phi = (x^1, \dots, x^n, v^1, \dots, v^n))$ of TM. Let $X\Big|_U(x) = X^i(x)\frac{\partial}{\partial x^i}(x)$. Then the tangent lift $X^T : TM \to TTM$ can be written locally as

$$X^T\Big|_{TU}(v_x) = X^i(x)\frac{\partial}{\partial x^i}(v_x) + \frac{\partial X^i}{\partial x^j}(x)v^j\frac{\partial}{\partial v^i}(v_x).$$

Essentially, the tangent lift gives us the desired linearization. We shall apply the tangent lift to (2.1.2) by splitting the calculation into various lemmas.

Lemma 2.9 (Tangent Lift of the Geodesic Spray)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F)$ be a C^∞-forced simple mechanical system and consider an equilibrium configuration $q_0 \in Q$ for Σ. Then, for all $v_1 \oplus v_2 \in T_{q_0}Q \oplus T_{q_0}Q \cong T_{0_{q_0}}TQ$, we have*

$$\overset{\mathbb{G}}{S}{}^T(v_1 \oplus v_2) = (0 \oplus 0 \oplus v_2 \oplus 0)_{v_1 \oplus v_2}.$$

Proof. Consider $(U, \phi = (q^1, \dots, q^n))$, a local chart of Q around q_0. The geodesic spray $\overset{\mathbb{G}}{S} : TQ \to TTQ$ in coordinates $(TU, T\phi = (q^1, \dots, q^n, v^1, \dots, v^n))$ is

$$\overset{\mathbb{G}}{S}\Big|_{TU}(v_q) = v^i \frac{\partial}{\partial q^i}(v_q) - \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q) v^j v^k \frac{\partial}{\partial v^i}(v_q).$$

The tangent lift $\overset{\mathbb{G}}{S}{}^T : TTQ \to TTTQ$ in coordinates $(TTU, TT\phi = (q^1, \dots, q^n, v^1, \dots, v^n, \dot{q}^1, \dots, \dot{q}^n, \dot{v}^1, \dots, \dot{v}^n))$ is

$$\begin{aligned}\overset{\mathbb{G}}{S}{}^T\Big|_{TTU}(w_{v_q}) &= v^i \frac{\partial}{\partial q^i}(w_{v_q}) - \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q) v^j v^k \frac{\partial}{\partial v^i}(w_{v_q}) + \frac{\partial X^i}{\partial q^j}(v_q)\dot{q}^j \frac{\partial}{\partial \dot{q}^i}(w_{v_q}) \\ &\quad + \frac{\partial X^i}{\partial v^j}(v_q)\dot{v}^j \frac{\partial}{\partial \dot{q}^i}(w_{v_q}) + \frac{\partial Y^i}{\partial q^j}(v_q)\dot{q}^j \frac{\partial}{\partial \dot{v}^i}(w_{v_q}) + \frac{\partial Y^i}{\partial v^j}(v_q)\dot{v}^j \frac{\partial}{\partial \dot{v}^i}(w_{v_q}),\end{aligned}$$

where $X^i(v_q) = v^i$ and $Y^i(v_q) = -\overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q) v^j v^k$. Hence, given $v_1 \oplus v_2 \in T_{q_0}Q \oplus T_{q_0}Q \cong T_{0_{q_0}}TQ$,

$$\begin{aligned}\overset{\mathbb{G}}{S}{}^T(v_1 \oplus v_2) &= \overset{\mathbb{G}}{S}{}^T\Big|_{TTU}(v_1 \oplus v_2) \\ &= 0 - 0 + 0 + v_2^i \frac{\partial}{\partial \dot{q}^i}(v_1 \oplus v_2) + 0 + 0 \\ &= v_2^i \frac{\partial}{\partial \dot{q}^i}(v_1 \oplus v_2) \\ &= (0 \oplus 0 \oplus v_2 \oplus 0)_{v_1 \oplus v_2},\end{aligned}$$

where $v_1 = v_1^i \frac{\partial}{\partial q^i}(q_0)$ and $v_2 = v_2^i \frac{\partial}{\partial q^i}(q_0)$. $\square$

Lemma 2.10 (Tangent Lift of the Potential Force)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F)$ be a C^∞-forced simple mechanical system and consider an equilibrium configuration $q_0 \in Q$ for Σ. If q_0 is a critical point for V (i.e. $dV(q_0) = 0$), then, for all $v_1 \oplus v_2 \in T_{q_0}Q \oplus T_{q_0}Q \cong T_{0_{q_0}}TQ$, we have*

$$(\mathit{vlft}(\mathbb{G}^\# \circ dV))^T(v_1 \oplus v_2) = (0 \oplus 0 \oplus 0 \oplus (\mathbb{G}(q_0)^\# \circ \mathit{Hess}V(q_0)^\flat(v_1)))_{v_1 \oplus v_2}.$$

Proof. Consider $(U, \phi = (q^1, \dots, q^n))$, a local chart of Q around q_0. The differential of V in coordinates $(U, \phi = (q^1, \dots, q^n))$ is

$$dV\Big|_U(q) = \frac{\partial V}{\partial q^i}(q) dq^i(q).$$

The gradient of V in coordinates $(U, \phi = (q^1, \dots, q^n))$ is

$$\mathbb{G}^\# \circ dV\Big|_U(q) = G^{ij}(q) \frac{\partial V}{\partial q^j}(q) \frac{\partial}{\partial q^i}(q).$$

The vertical lift $\mathrm{vlft}(\mathbb{G}^\# \circ dV) : TQ \to TTQ$ in coordinates $(TU, T\phi = (q^1, \dots, q^n, v^1, \dots, v^n))$ is

$$\mathrm{vlft}(\mathbb{G}^\# \circ dV)\Big|_{TU}(v_q) = G^{ij}(q) \frac{\partial V}{\partial q^j}(q) \frac{\partial}{\partial v^i}(v_q).$$

The tangent lift $(\mathrm{vlft}(\mathbb{G}^{\#}\circ dV))^T : TTQ \to TTTQ$ in coordinates $(TTU, TT\phi = (q^1,\ldots,q^n, v^1,\ldots,v^n,\dot{q}^1,\ldots,\dot{q}^n,\dot{v}^1,\ldots,\dot{v}^n))$ is

$$(\mathrm{vlft}(\mathbb{G}^{\#}\circ dV))^T\Big|_{TTU}(w_{v_q}) = G^{ij}(q)\frac{\partial V}{\partial q^j}(q)\frac{\partial}{\partial v^i}(w_{v_q}) + \frac{\partial X^i}{\partial q^j}(v_q)\dot{q}^j\frac{\partial}{\partial \dot{q}^i}(w_{v_q}) + \frac{\partial X^i}{\partial v^j}(v_q)\dot{v}^j\frac{\partial}{\partial \dot{q}^i}(w_{v_q}) + \frac{\partial Y^i}{\partial q^j}(v_q)\dot{q}^j\frac{\partial}{\partial \dot{v}^i}(w_{v_q}) + \frac{\partial Y^i}{\partial v^j}(v_q)\dot{v}^j\frac{\partial}{\partial \dot{v}^i}(w_{v_q})$$

where $X^i(v_q) = 0$ and $Y^i(v_q) = G^{ij}(q)\frac{\partial V}{\partial q^j}(q)$. Hence, given $v_1\oplus v_2 \in T_{q_0}Q\oplus T_{q_0}Q \cong T_{0_{q_0}}TQ$,

$$\begin{aligned}
&(\mathrm{vlft}(\mathbb{G}^{\#}\circ dV))^T(v_1\oplus v_2)\\
&= (\mathrm{vlft}(\mathbb{G}^{\#}\circ dV))^T\Big|_{TTU}(v_1\oplus v_2)\\
&= 0+0+0+\left(\frac{\partial G^{ij}}{\partial q^k}(q_0)\underbrace{\frac{\partial V}{\partial q^j}(q_0)}_{0} + G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\right)v_1^k\frac{\partial}{\partial \dot{v}^i}(v_1\oplus v_2)+0\\
&= G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)v_1^k\frac{\partial}{\partial \dot{v}^i}(v_1\oplus v_2)\\
&= (0\oplus 0\oplus 0\oplus(\mathbb{G}(q_0)^{\#}\circ \mathrm{Hess}V(q_0)^{\flat}(v_1)))_{v_1\oplus v_2},
\end{aligned}$$

where $v_1 = v_1^i\frac{\partial}{\partial q^i}(q_0)$ and $v_2 = v_2^i\frac{\partial}{\partial q^i}(q_0)$. □

Lemma 2.11 (Tangent Lift of the External Force)**.** *Let $\Sigma = (Q,\mathbb{G},V,F)$ be a C^∞-forced simple mechanical system and consider an equilibrium configuration $q_0\in Q$ for Σ. If $F(Z(TQ)) = Z(T^*Q)$, then, for all $v_1\oplus v_2 \in T_{q_0}Q\oplus T_{q_0}Q\cong T_{0_{q_0}}TQ$, we have*

$$(\mathit{vlft}(\mathbb{G}^{\#}\circ F))^T(v_1\oplus v_2) = (0\oplus 0\oplus 0\oplus(\mathbb{G}(q_0)^{\#}\circ d_1F(0_{q_0})(v_1)+\mathbb{G}(q_0)^{\#}\circ d_2F(0_{q_0})(v_2)))_{v_1\oplus v_2},$$

where we have defined $d_1F(0_{q_0}), d_2F(0_{q_0}) \in L(T_{q_0}Q;T^*_{q_0}Q)$ such that given a local chart $(U,\phi=(q^1,\ldots,q^n))$ of Q and corresponding local chart $(TU, T\phi = (q^1,\ldots,q^n,v^1,\ldots,v^n))$ of TQ,

$$[d_1F(0_{q_0})] = \left[\frac{\partial F_i}{\partial q^j}(0_{q_0})\right],\quad [d_2F(0_{q_0})] = \left[\frac{\partial F_i}{\partial v^j}(0_{q_0})\right].$$

Proof. Consider $(U,\phi=(q^1,\ldots,q^n))$, a local chart of Q around q_0. The external force in coordinates $(U,\phi=(q^1,\ldots,q^n))$ is

$$F\Big|_{TU}(v_q) = F_i(v_q)dq^i(q),$$

and so

$$\mathbb{G}^{\#}\circ F\Big|_{TU}(v_q) = G^{ij}(q)F_j(v_q)\frac{\partial}{\partial q^i}(q).$$

The vertical lift

$$\begin{aligned}\mathrm{vlft}(\mathbb{G}^{\#}\circ F):\ & TQ \to TTQ\\ & v_q \mapsto \mathrm{vlft}_{v_q}(\mathbb{G}^{\#}\circ F(v_q))\end{aligned}$$

in coordinates $(TU, T\phi = (q^1, \dots, q^n, v^1, \dots, v^n))$ is

$$\mathrm{vlft}(\mathbb{G}^\# \circ F)\Big|_{TU}(v_q) = G^{ij}(q)F_j(v_q)\frac{\partial}{\partial v^i}(v_q).$$

The tangent lift $(\mathrm{vlft}(\mathbb{G}^\# \circ F))^T : TTQ \to TTTQ$ in coordinates $(TTU, TT\phi = (q^1, \dots, q^n, v^1, \dots, v^n, \dot{q}^1, \dots, \dot{q}^n, \dot{v}^1, \dots, \dot{v}^n))$ is

$$\begin{aligned}(\mathrm{vlft}(\mathbb{G}^\# \circ F))^T\Big|_{TTU}(w_{v_q}) &= G^{ij}(q)F_j(v_q)\frac{\partial}{\partial v^i}(w_{v_q}) + \frac{\partial X^i}{\partial q^j}(v_q)\dot{q}^j\frac{\partial}{\partial \dot{q}^i}(w_{v_q}) \\ &\quad + \frac{\partial X^i}{\partial v^j}(v_q)\dot{v}^j\frac{\partial}{\partial \dot{q}^i}(w_{v_q}) + \frac{\partial Y^i}{\partial q^j}(v_q)\dot{q}^j\frac{\partial}{\partial \dot{v}^i}(w_{v_q}) + \frac{\partial Y^i}{\partial v^j}(v_q)\dot{v}^j\frac{\partial}{\partial \dot{v}^i}(w_{v_q})\end{aligned}$$

where $X^i(v_q) = 0$ and $Y^i(v_q) = G^{ij}(q)F_j(v_q)$. Hence, given $v_1 \oplus v_2 \in T_{q_0}Q \oplus T_{q_0}Q \cong T_{0_{q_0}}TQ$,

$$\begin{aligned}(\mathrm{vlft}(\mathbb{G}^\# \circ F))^T(v_1 \oplus v_2) &= (\mathrm{vlft}(\mathbb{G}^\# \circ F))^T\Big|_{TTU}(v_1 \oplus v_2) \\ &= 0 + 0 + 0 + \left(\frac{\partial G^{ij}}{\partial q^k}(q_0)\underbrace{F_j(0_{q_0})}_{0} + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\right)v_1^k\frac{\partial}{\partial \dot{v}^i}(v_1 \oplus v_2) \\ &\quad + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})v_2^k\frac{\partial}{\partial \dot{v}^i}(v_1 \oplus v_2) \\ &= G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})v_1^k\frac{\partial}{\partial \dot{v}^i}(v_1 \oplus v_2) + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})v_2^k\frac{\partial}{\partial \dot{v}^i}(v_1 \oplus v_2) \\ &= (0 \oplus 0 \oplus 0 \oplus (\mathbb{G}(q_0)^\# \circ d_1F(0_{q_0})(v_1) + \mathbb{G}(q_0)^\# \circ d_2F(0_{q_0})(v_2)))_{v_1 \oplus v_2},\end{aligned}$$

where $v_1 = v_1^i\frac{\partial}{\partial q^i}(q_0)$ and $v_2 = v_2^i\frac{\partial}{\partial q^i}(q_0)$. □

Now, putting everything together, we get the following.

Proposition 2.12 (Linearization of a Forced Simple Mechanical System)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F)$ be a C^∞-forced simple mechanical system such that $F(Z(TQ)) = Z(T^*Q)$ and consider an equilibrium configuration $q_0 \in Q$ for Σ (i.e. we have $dV(q_0) = 0$). Then the linearization of (2.1.2) at 0_{q_0} in the decomposition $T_{0_{q_0}}TQ \cong T_{q_0}Q \oplus T_{q_0}Q$ is given by*

$$A_\Sigma(q_0) = \begin{pmatrix} 0 & id_{T_{q_0}Q} \\ -\mathbb{G}^\# \circ \mathit{Hess}V(q_0)^\flat + \mathbb{G}^\# \circ d_1F(0_{q_0}) & \mathbb{G}^\# \circ d_2F(0_{q_0}) \end{pmatrix},$$

and the linearized equations of motion are

$$\begin{pmatrix} \dot{q}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} 0 & id_{T_{q_0}Q} \\ -\mathbb{G}^\# \circ \mathit{Hess}V(q_0)^\flat + \mathbb{G}^\# \circ d_1F(0_{q_0}) & \mathbb{G}^\# \circ d_2F(0_{q_0}) \end{pmatrix}\begin{pmatrix} q(t) \\ v(t) \end{pmatrix}.$$

*In other words, the linearization is given by the forced linear mechanical system $(T_{q_0}Q, \mathbb{G}(q_0),$
$\mathit{Hess}V(q_0), (d_1F(q_0), d_2F(q_0)))$.*

Proof. This follows from combining the results of Lemmas 2.8, 2.9, 2.10 and 2.11. □

2.3. Stability of Equilibria

In this section, we look at the stability of equilibria. For a nonlinear system, the stability of an equilibrium point is typically analyzed in one of two ways: (1) stability of its linearization, or (2) Lyapunov stability. For unconstrained mechanical systems, the same two approaches can be applied for an equilibrium configuration q_0 by considering the associated vector field and its equilibrium point 0_{q_0}. Due to the nature of mechanical systems and the structure of its linearizations, we can obtain specialized results. In particular, for the unconstrained mechanical systems that we study here, i.e. those subject to dissipative forces, we can examine more closely the spectrum of the linearization, and for Lyapunov stability, use the energy function as a Lyapunov function.

Definition 2.13 (Dissipative Force)**.** Consider a time-independent force $F : TQ \to T^*Q$.
(i) F is **dissipative** if, for all $v_q \in TQ$, we have $\langle F(v_q); v_q\rangle \leq 0$.
(ii) F is **strictly dissipative** if F is dissipative and $\langle F(v_q); v_q\rangle = 0$ only when $v_q \in Z(TQ)$.

Definition 2.14 (Rayleigh Dissipation)**.** A $(0, 2)$-tensor field R_{diss} on Q is a **Rayleigh dissipation function** if R_{diss} is C^∞, symmetric, and positive-semidefinite. If R_{diss} is positive-definite, then it is **strict**. The dissipative force associated with R_{diss} is $-R^\flat_{\text{diss}} : TQ \to T^*Q$.

Let us compile some basic results regarding the stability of equilibria via linearization. We refer to [3] for the proofs and additional details.

Theorem 2.15 (Normal Form for a Symmetric Bilinear Map)**.** *Consider a finite-dimensional $\mathbb{R}$-vector space V. Given $B \in \Sigma_2(V)$, there exists a basis $\{e_1, \ldots, e_n\}$ for V such that*

$$[B] = \begin{pmatrix} I_{p\times p} & 0 & 0 \\ 0 & -I_{q\times q} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Definition 2.16 (Rank, Index and Signature)**.** Referring to Theorem 2.15, we define the following.
(i) The **rank of** B is the number of nonzero elements on the diagonal, i.e. $\text{rank}(B) = p + q$.
(ii) The **index of** B is the number of -1's on the diagonal, i.e. $\text{ind}(B) = q$.
(iii) The **signature of** B is the number of 1's on the diagonal minus the number of -1's on the diagonal, i.e. $\text{sig}(B) = p - q$.

Proposition 2.17 (Eigenvalues of an Unforced Linear Mechanical System)**.** *Let* $(V, M, K, (0, 0))$ *be an unforced linear mechanical system and consider* $A = \begin{pmatrix} 0 & id_V \\ -M^\sharp \circ K^\flat & 0 \end{pmatrix}$.
(i) The eigenvalues of $M^\sharp \circ K^\flat$ are real.
(ii) $l \in \mathbb{R}$ is an eigenvalue of $M^\# \circ K^\flat$ if and only if $\lambda = \pm\sqrt{-l}$ are eigenvalues of A.

Proposition 2.18 (Stability of an Unforced Linear Mechanical System)**.** *Let* $(V, M, K, (0, 0))$ *be an unforced linear mechanical system.*

(i) The equilibrium point 0 *is spectrally stable (see Definition A.9) if and only if* $ind(K) = 0$.
(ii) The equilibrium point 0 *is stable if and only if* $rank(K) = n$ *and* $ind(K) = 0$.
(iii) If $ind(K) > 0$*, then the equilibrium point* 0 *is unstable.*

Theorem 2.19 (Stability of a Linear Mechanical System with Linear Dissipation)**.** *Let* $\Sigma = (V, M, K, (0, -R^\flat))$ *be a forced linear mechanical system, where* $R \in \Sigma_2(V)$ *and* R *is positive-semidefinite. (We call* R *a* ***linear Rayleigh dissipation function*** *and we say that* Σ *is a linear mechanical system subject to a* ***linear dissipative force****.)*
(i) If K *is not positive-semidefinite, then the equilibrium point* 0 *is unstable.*
(ii) The equilibrium point 0 *is stable if and only if* K *is positive-semidefinite and* $Ker(K^\flat) \cap Ker(R^\flat) = \{0\}$.
(iii) The equilibrium point 0 *is asymptotically stable if and only if* K *is positive-definite and* $\langle M^\# \circ K^\flat, Im(M^\# \circ R^\flat)\rangle = V$.

Note that for an $\mathbb{F}$-vector space V, a subset $\mathcal{L} \subseteq L(V;V)$, and a subspace $U \subseteq V$, $\langle \mathcal{L}, U\rangle$ is defined as the smallest subspace of V containing U that is also an invariant subspace for each $A \in \mathcal{L}$.

Theorem 2.20 (Stability of a Rayleigh Dissipative Unconstrained Mechanical System via Linearization)**.** *Let* $\Sigma = (Q, \mathbb{G}, V, -R^\flat_{diss})$ *be a* C^∞*-forced simple mechanical system subject to the force associated with a Rayleigh dissipation function* R_{diss} *and suppose* $q_0 \in Q$ *is an equilibrium configuration for* Σ.
(i) If $HessV(q_0)$ *is not positive-semidefinite, then* q_0 *is unstable.*
(ii) If $HessV(q_0)$ *is positive-definite and* $\langle \mathbb{G}(q_0)^\# \circ HessV(q_0)^\flat, Im(\mathbb{G}(q_0)^\# \circ R_{diss}(q_0)^\flat)\rangle = T_{q_0}Q$*, then* q_0 *is locally asymptotically stable.*

Now, let us examine the stability of equilibria via Lyapunov methods.

Proposition 2.21 (Time Derivative of Energy for Unconstrained Mechanical Systems)**.** *Let* $\Sigma = (Q, \mathbb{G}, V, F)$ *be a* C^∞*-forced simple mechanical system, and consider a curve* $\gamma : I \to Q$ *satisfying the equations of motion (2.1.1). Then,*

$$\frac{dE(\gamma'(t))}{dt} = \langle F(\gamma'(t)); \gamma'(t)\rangle.$$

Equivalently, let X *be the associated vector field for* Σ*. Then,*

$$\mathcal{L}_X E(v_q) = \langle F(v_q); v_q\rangle.$$

Proof. We know that the total energy is

$$E(v_q) = \frac{1}{2}\mathbb{G}(q)(v_q, v_q) + V(q),$$

hence

$$E(\gamma'(t)) = \frac{1}{2}\mathbb{G}(\gamma(t))(\gamma'(t), \gamma'(t)) + V(\gamma(t)),$$

and so taking the time derivative we get

$$\begin{aligned}
&\frac{dE(\gamma'(t))}{dt}\\
&= \frac{d}{dt}\left(\frac{1}{2}\mathbb{G}(\gamma(t))(\gamma'(t),\gamma'(t)) + V(\gamma(t))\right)\\
&= \frac{1}{2}(\underbrace{\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\mathbb{G}}_{0})(\gamma(t))(\gamma'(t),\gamma'(t)) + \frac{1}{2}\mathbb{G}(\gamma(t))(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t),\gamma'(t))\\
&\quad + \frac{1}{2}\underbrace{\mathbb{G}(\gamma(t))(\gamma'(t),\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t))}_{\mathbb{G}(\gamma(t))(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t),\gamma'(t))} + (\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}V)(\gamma(t))\\
&= \mathbb{G}(\gamma(t))(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle\\
&= \mathbb{G}(\gamma(t))(-\mathbb{G}^{\#}\circ dV(\gamma(t)) + \mathbb{G}^{\#}\circ F(\gamma'(t)),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle\\
&= -\mathbb{G}(\gamma(t))(\mathbb{G}^{\#}\circ dV(\gamma(t)),\gamma'(t)) + \mathbb{G}(\gamma(t))(\mathbb{G}^{\#}\circ F(\gamma'(t)),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle\\
&= -\langle \mathbb{G}^{\flat}(\mathbb{G}^{\#}\circ dV(\gamma(t)));\gamma'(t)\rangle + \langle \mathbb{G}^{\flat}(\mathbb{G}^{\#}\circ F(\gamma'(t)));\gamma'(t)\rangle + \langle dV(\gamma(t));\gamma'(t)\rangle\\
&= -\langle dV(\gamma(t));\gamma'(t)\rangle + \langle F(\gamma'(t));\gamma'(t)\rangle + \langle dV(\gamma(t));\gamma'(t)\rangle\\
&= \langle F(\gamma'(t));\gamma'(t)\rangle.
\end{aligned}$$

The equivalence of the time derivative and Lie derivative expressions is Proposition A.11. □

Proposition 2.22 (Dissipative Forces Decrease Energy)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F_{diss})$ be a C^∞-forced simple mechanical system where F_{diss} is dissipative, and consider a curve $\gamma : I \to Q$ satisfying the equations of motion (2.1.1). Then, the energy $E(\gamma'(t))$ is non-increasing. Equivalently, considering the associated vector field X for Σ, we have that $\mathcal{L}_X E(v_q) \leq 0$ for all $v_q \in TQ$.*

Proof. This follows directly from Proposition 2.21. □

Theorem 2.23 (Stability of a Dissipative Unconstrained Mechanical System via Lyapunov Methods)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F_{diss})$ be a C^∞-forced simple mechanical system where F_{diss} is dissipative, and consider an equilibrium configuration $q_0 \in Q$ for Σ, i.e. we have $dV(q_0) = 0$ since $F(Z(TQ)) = Z(T^*Q)$.*

(i) If V is locally positive-definite about q_0, then q_0 is stable.
(ii) If q_0 is an isolated local minimum for V and F_{diss} is strictly dissipative, then q_0 is locally asymptotically stable.

Note that q_0 being an isolated local minimum for V implies that V is locally positive-definite about q_0. The converse is only true for analytic functions.

Proof. (i) Consider the energy function $E : TQ \to \mathbb{R}$, $E(v_q) = \frac{1}{2}\mathbb{G}(q)(v_q, v_q) + V(q)$. Without loss of generality, assume $V(q_0) = 0$.
Claim: E is a Lyapunov function for the associated vector field X for Σ at the point 0_{q_0}.
Proof of claim:

1. We have $E(0_{q_0}) = 0$.
2. $E(\gamma'(t))$ is non-increasing by Proposition 2.22. In other words, $\mathcal{L}_X E$ is negative-semidefinite about 0_{q_0}.
3. We need to show that there exists a neighbourhood $TU \subseteq TQ$ of 0_{q_0} such that for all $v_q \in TU \setminus \{0_{q_0}\}$, we have $E(v_q) > 0$. Since V is locally positive-definite about q_0, there exists a neighbourhood $U \subseteq Q$ of q_0 such that for all $q \in U \setminus \{q_0\}$, we have $V(q) > 0$. Now, for all $v_q \in TU \setminus Z(TU)$, we have $\mathbb{G}(q)(v_q, v_q) > 0$. Hence, for all $v_q \in TU \setminus \{0_{q_0}\}$, we have $E(v_q) > 0$. Hence E is a Lyapunov function for the associated vector field at the point 0_{q_0} as claimed.

Hence, by the Lyapunov stability criteria (Theorem A.16), q_0 is stable.

(ii) Since q_0 is an isolated local minimum for V, V is locally positive-definite about q_0 and so (i) applies. E is a Lyapunov function. Also, since q_0 is an isolated local minimum for V, there exists a neighbourhood $W \subseteq Q$ of q_0 such that for all $q \in W \setminus \{q_0\}$, we have $V(q) > V(q_0)$, and W contains no critical points other than q_0. We have that E is positive-definite on TW and $E(\gamma'(t))$ is non-increasing. Let $A = \{v_q \in TW \mid \langle F_{\text{diss}}(v_q); v_q\rangle = 0\} = Z(TW)$.

Claim: $\{0_{q_0}\}$ is the only positively invariant set in A.

Proof of claim: Suppose $0_{q_1} \neq 0_{q_0}$ is another point. Then $q_1 \neq q_0$ implies that $\text{grad}V(q_1) \neq 0$. Hence the solution will leave A, proving the claim.

Hence the corollary to LaSalle Invariance Principle (Corollary A.22) implies that q_0 is locally asymptotically stable.

□

Chapter 3

Nonholonomic Mechanical Systems: Equations of Motion and Equilibria

In this chapter, we begin to study constrained mechanical systems, i.e. simple mechanical systems subject to velocity constraints. Velocity constraints are modeled as a distribution on the configuration manifold, and can either be holonomic or nonholonomic. Note that we only consider linear velocity constraints which are regular C^∞-distributions. Afterwards, we look at how the equations of motion and equilibria change when these constraints are added.

3.1. Velocity Constraints

Definition 3.1 (Annihilator)**.** Let V be a finite-dimensional $\mathbb{R}$-vector space and $S \subseteq V$. We define the **annihilator of** S by

$$\operatorname{ann}(S) = \{\alpha \in V^* \mid \alpha(v) = 0 \ \forall v \in S\}.$$

Now, given a distribution D on a C^∞-manifold M, we define the **annihilator of** D, $\operatorname{ann}(D)$, such that

$$\operatorname{ann}(D)_q = \operatorname{ann}(D_q).$$

Note that since $\operatorname{ann}(S)$ is a subspace of V^*, $\operatorname{ann}(D)$ is a codistribution on M.

Definition 3.2 (Linear Velocity Constraint)**.** Let Q be a C^∞-manifold. A distribution D on Q is a C^∞**-linear velocity constraint** if $\operatorname{ann}(D)$ is a C^∞-codistribution. A C^∞-linear velocity constraint D is **regular** if D is a regular distribution. Alternatively, a distribution D on Q is a **regular** C^∞**-linear velocity constraint** if D is a regular C^∞-distribution.

Remark 3.3. We will only work with regular C^∞-linear velocity constraints and so the technicalities above involving regularity and smoothness can be ignored.

Often we will say velocity constraints, constraint distribution, regular constraints, or simply constraints in reference to a regular C^∞-linear velocity constraint. Constraints can either be holonomic or nonholonomic. These are defined as follows.

Definition 3.4 (Holonomic and Nonholonomic Constraints)**.** Let D be a regular C^∞-linear velocity constraint on a C^∞-manifold Q. Then,
(i) D is **holonomic** if D is integrable, and
(ii) D is **nonholonomic** if D is not holonomic.

Note that we are not interested in mechanical systems subject to holonomic constraints. As discussed in [3], if the distribution is integrable, this means that all curves through a point $q \in Q$ satisfying the constraint must lie on the maximal integral manifold for D through q. In this case, we can always restrict our analysis to the maximal integral manifold of the distribution, and then the results for unconstrained mechanical systems would apply. It is the nonholonomicity of constraints which introduces additional complexities into the equations of motion and stability analysis.

Definition 3.5 (Totally Nonholonomic)**.** Let D be a nonholonomic constraint on a C^∞-manifold Q. Denote by $\Gamma^\infty(D)$ the set of C^∞-vector fields on Q taking values in D and define $\mathrm{Lie}^\infty(D)$ as the smallest subspace of $\Gamma^\infty(TQ)$ such that $\Gamma^\infty(D) \subseteq \mathrm{Lie}^\infty(D)$ and, for all $X, Y \in \mathrm{Lie}^\infty(D)$, we have $[X, Y] \in \mathrm{Lie}^\infty(D)$. We say that D is **totally nonholonomic** if, for all $q \in Q$, we have $\mathrm{Lie}^\infty(D)_q \triangleq \{X(q) \mid X \in \mathrm{Lie}^\infty(D)\} = T_qQ$.

Essentially, if a constraint is totally nonholonomic, for every point in the configuration space, although our instantaneous motion is restricted to a subspace of directions, it is possible to achieve net movement in any direction.

3.2. Equations of Motion

Here, we study the equations of motion for nonholonomic mechanical systems. Although the results obviously apply for holonomic mechanical systems as well, they do not provide any additional information of interest to us.

Definition 3.6 (Forced Simple Mechanical System with Constraints)**.** A **C^∞-forced simple mechanical system with constraints** is a 5-tuple $(Q, \mathbb{G}, V, F, D)$, where
(i) $(Q, \mathbb{G}, V, F)$ is a C^∞-forced simple mechanical system, and
(ii) D is a C^∞-linear velocity constraint.

Note that since we are only interested in constraints that are regular, we will talk about C^∞-forced simple mechanical systems with regular constraints. This is our general notion of a constrained mechanical system.

Definition 3.7 (Curve Satisfying a Constraint)**.** Let D be a C^∞-linear velocity constraint on a C^∞-manifold Q. A C^∞-curve $\gamma : I \to Q$ **satisfies the constraint** D if $\gamma'(t) \in D_{\gamma(t)}$ for all $t \in I$.

Definition 3.8 (Constraint Force)**.** Let D be a C^∞-linear velocity constraint on a C^∞-manifold Q. A **constraint force** is a force taking values in $\mathrm{ann}(D)$. Given a continuous curve $\gamma : I \to Q$, a covector field $\alpha : I \to T^*Q$ along γ is called a **constraint force along** γ if, for all $t \in I$, we have $\alpha(t) \in \mathrm{ann}(D)_{\gamma(t)}$.

A physical motivation for the definition of constraint forces is that they do no work on curves satisfying the constraint. The converse is in general not true, though.

Definition 3.9 (Orthogonal Projections)**.** Let D be a regular C^∞-linear velocity constraint on a C^∞-manifold Q. We define two vector bundle maps over id_Q, $P_D : TQ \to TQ$ and $P_D^\perp : TQ \to TQ$, such that, for all $v_q \in TQ$,
(i) $v_q = P_D(v_q) \oplus P_D^\perp(v_q)$, and
(ii) $P_D(v_q) \in D_q$ and $P_D^\perp(v_q) \in D_q^\perp$.

Note that P_D and $P_D^\perp$ can be treated as $(1,1)$-tensor fields.

Proposition 3.10 (Equations of Motion for Constrained Mechanical Systems)**.** *Let* $(Q, \mathbb{G}, V,$
$F, D)$ *be a* C^∞*-forced simple mechanical system with regular constraints. A curve* $\gamma : I \to Q$ *satisfies the constraint* D *and the Lagrange-d'Alembert Principle for the force* $F + \alpha$ *and Lagrangian* $L(v_q) = \frac{1}{2}\mathbb{G}(v_q, v_q) - V(q)$ *where* α *is a constraint force along* γ*, i.e.*

$$\begin{cases} \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -gradV(\gamma(t)) + \mathbb{G}^\# \circ F(\gamma'(t)) + \mathbb{G}^\# \circ \alpha(t) \\ \gamma'(t) \in D_{\gamma(t)} \end{cases}$$

if and only if there exists a vector field λ *along* γ *such that* $\lambda(t) \in D^\perp_{\gamma(t)}$*, and*

$$\begin{cases} \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -gradV(\gamma(t)) + \mathbb{G}^\# \circ F(\gamma'(t)) + \lambda(t) \\ P_D^\perp(\gamma'(t)) = 0. \end{cases} \tag{3.2.1}$$

Proof. This follows since
1. $P_D^\perp(\gamma'(t)) = 0 \iff \gamma'(t) \in D_{\gamma(t)}$ for all $t \in I$, and
2. points in $\mathrm{ann}(D)$ are mapped bijectively onto $D^\perp$ by $\mathbb{G}^\#$, i.e. $\lambda(t) = \mathbb{G}^\# \circ \alpha(t)$.

□

Proposition 3.10 gives us the equations of motion for constrained mechanical systems, but we still need to solve for the unknown λ. Note that λ is in fact a "Lagrange multiplier". We introduce the constrained affine connection here, which will help us simplify the equations of motion after λ is solved for.

Definition 3.11 (Constrained Affine Connection)**.** Let Q be a C^∞-manifold, $\mathbb{G}$ a C^∞-Riemannian metric on Q, and D a regular C^∞-linear velocity constraint on Q. The **constrained affine connection**, $\overset{D}{\nabla}$, is

$$\overset{D}{\nabla}_X Y \triangleq \overset{\mathbb{G}}{\nabla}_X Y + (\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y),$$

where $P_D^\perp : TQ \to TQ$ is thought of as a $(1,1)$-tensor field. We denote the geodesic spray associated with $\overset{D}{\nabla}$ by $\overset{D}{S}$.

Proposition 3.12 (Properties of the Constrained Connection)**.** *Let* Q *be a* C^∞*-manifold,* $\mathbb{G}$ *a* C^∞*-Riemannian metric on* Q*, and* D *a regular* C^∞*-linear velocity constraint on* Q*. Consider the Levi-Civita affine connection* $\overset{\mathbb{G}}{\nabla}$ *and the constrained connection* $\overset{D}{\nabla}$*. Given* $X \in \Gamma^\infty(TQ)$ *and* $Y \in \Gamma^\infty(D)$*, we have* *(i)* $\overset{D}{\nabla}_X Y = P_D(\overset{\mathbb{G}}{\nabla}_X Y)$ *and therefore* $\overset{D}{\nabla}_X Y \in \Gamma^\infty(D)$, *and*

(ii) $(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) = -P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y)$ *and therefore* $(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) \in \Gamma^\infty(D^\perp)$.

Proof. Let $X \in \Gamma^\infty(TQ)$ and $Y \in \Gamma^\infty(D)$. We have that $P_D^\perp(Y) = 0$, hence taking the covariant derivative we get

$$\overset{\mathbb{G}}{\nabla}_X(P_D^\perp(Y)) = 0,$$

and by the product rule for the covariant derivative we get

$$(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) + P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y) = 0.$$

Rearranging, we get that $(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) = -P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y) \in \Gamma^\infty(D^\perp)$. This proves the second statement. Now,

$$\begin{aligned}\overset{D}{\nabla}_X Y &= \overset{\mathbb{G}}{\nabla}_X Y + (\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y)\\ &= \overset{\mathbb{G}}{\nabla}_X Y - P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y)\\ &= P_D(\overset{\mathbb{G}}{\nabla}_X Y) \in \Gamma^\infty(D),\end{aligned}$$

which proves the first statement. □

Definition 3.13. Let D be a regular C^∞-distribution on a C^∞-manifold Q. A C^∞-affine connection ∇ **restricts** to D if, for all $X \in \Gamma^\infty(TQ)$ and for all $Y \in \Gamma^\infty(D)$, we have $\nabla_X Y \in \Gamma^\infty(D)$.

Hence we can say that $\overset{D}{\nabla}$ restricts to D.

Theorem 3.14 (Equations of Motion for Constrained Mechanical Systems using the Constrained Connection)**.** *Let $(Q, \mathbb{G}, V, F, D)$ be C^∞-forced simple mechanical system with regular constraints. A curve $\gamma : I \to Q$ satisfies the constraint D and the Lagrange-d'Alembert Principle for the force $F + \alpha$ and Lagrangian $L(v_q) = \frac{1}{2}\mathbb{G}(v_q, v_q) - V(q)$ where α is a constraint force along γ, i.e.*

$$\begin{cases}\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -gradV(\gamma(t)) + \mathbb{G}^\# \circ F(\gamma'(t)) + \mathbb{G}^\# \circ \alpha(t)\\ \gamma'(t) \in D_{\gamma(t)}\end{cases}$$

if and only if

$$\overset{D}{\nabla}_{\gamma'(t)}\gamma'(t) = -P_D \circ gradV(\gamma(t)) + P_D \circ \mathbb{G}^\# \circ F(\gamma'(t)) \tag{3.2.2}$$

and $\gamma'(t_0) \in D_{\gamma(t_0)}$ for some $t_0 \in I$.

Proof. We only consider the forward direction of the proof here. We refer to [3] for a complete proof.

Applying Proposition 3.10, it is clear that since $P_D^\perp(\gamma'(t)) = 0$, there exists $t_0 \in I$ such that $\gamma'(t_0) \in D_{\gamma(t_0)}$. By Proposition 3.12, we have that $P_D^\perp(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)) =$

$-(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} P_D^{\perp})(\gamma'(t))$. Now, consider $\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -\mathrm{grad}V(\gamma(t)) + \mathbb{G}^{\#} \circ F(\gamma'(t)) + \lambda(t)$. We solve for $\lambda(t)$:

$$\begin{aligned}\lambda(t) &= \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) + \mathrm{grad}V(\gamma(t)) - \mathbb{G}^{\#} \circ F(\gamma'(t)) \\ &= P_D^{\perp}(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)) + P_D^{\perp}(\mathrm{grad}V(\gamma(t))) - P_D^{\perp}(\mathbb{G}^{\#} \circ F(\gamma'(t))) \\ &= -(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} P_D^{\perp})(\gamma'(t)) + P_D^{\perp}(\mathrm{grad}V(\gamma(t))) - P_D^{\perp}(\mathbb{G}^{\#} \circ F(\gamma'(t))).\end{aligned}$$

Hence, substituting this expression for $\lambda(t)$, we get

$$\begin{aligned}\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = &-(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} P_D^{\perp})(\gamma'(t)) - P_D^{\perp}(\mathbb{G}^{\#} \circ F(\gamma'(t))) + P_D^{\perp}(\mathrm{grad}V(\gamma(t))) \\ &+ \mathbb{G}^{\#} \circ F(\gamma'(t)) - \mathrm{grad}V(\gamma(t)),\end{aligned}$$

and so

$$\overset{D}{\nabla}_{\gamma'(t)}\gamma'(t) = P_D(\mathbb{G}^{\#} \circ F(\gamma'(t))) - P_D(\mathrm{grad}V(\gamma(t))),$$

using the fact that $P_D = \mathrm{id}_{TQ} - P_D^{\perp}$. □

3.3. Equilibria

In contrast to dissipative mechanical systems without velocity constraints, the critical points of the potential function do not fully characterize the equilibria of a nonholonomic mechanical system. While any critical point of the potential function is still an equilibrium configuration for the nonholonomic mechanical system, the velocity constraints introduce new equilibria, which correspond to points in which the gradient of the potential function is orthogonal to the allowable velocities. In fact, the equilibria tend to no longer be isolated, and tend to form a manifold of equilibria. However, this is not always the case, contrary to the claims in [11] and assumptions made by many researchers. Equilibria can still be isolated, and the equilibria do not necessarily form a submanifold of the configuration space. We borrow examples from [6] and rework them here.

Definition 3.15 (Equilibrium Configuration)**.** Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints. A point $q_0 \in Q$ is an **equilibrium configuration for** Σ if the trivial curve $\gamma(t) = q_0$ satisfies the equations of motion (3.2.2).

Let us characterize the equilibrium configurations of a nonholonomic mechanical system.

Lemma 3.16 (Characterization of Equilibrium Configurations)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$. A point $q_0 \in Q$ is an equilibrium configuration for Σ if and only if $dV(q_0) \in ann(D_{q_0})$ (equivalently, $P_D(\mathbb{G}^{\#} \circ dV(q_0)) = 0$).*

Proof. ($\Rightarrow$) Since q_0 is an equilibrium configuration we have that $\gamma(t) = q_0$, $\gamma'(t) = 0_{q_0}$, and $\overset{D}{\nabla}_{\gamma'(t)}\gamma'(t) = 0$. So $-P_D(\mathbb{G}^{\#} \circ dV(\gamma(t))) + P_D(\mathbb{G}^{\#} \circ F(\gamma'(t))) = 0$. Since $F(\gamma'(t)) = 0$, we get $-P_D(\mathbb{G}^{\#} \circ dV(q_0)) = 0$, or $P_D(\mathbb{G}^{\#} \circ dV(q_0)) = 0$.

($\Leftarrow$) Suppose $P_D(\mathbb{G}^\# \circ dV(q_0)) = 0$. Consider $\gamma(t) = q_0$. Then $\gamma'(t) = 0_{q_0}$ and

$$\overset{D}{\nabla}_{\gamma'(t)}\gamma'(t) = -P_D(\mathbb{G}^\# \circ dV(\gamma(t))) + P_D(\mathbb{G}^\# \circ F(\gamma'(t)))$$

is satisfied. Hence q_0 is an equilibrium configuration. $\square$

The following is an example of a nonholonomic mechanical system in which the set of equilibria do not form a manifold. The original example is found in [6]. For this example (and subsequent examples), we use *Mathematica* to aid us in heavy computations and symbolic manipulations.

Example 3.17. Consider $\Sigma = (Q, \mathbb{G}, V, F, D)$, a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, where

(i) $Q = \mathbb{R}^3$ with global coordinate charts $(\mathbb{R}^3, \phi = (x, y, z))$ and $(T\mathbb{R}^3, T\phi = (x, y, z, u, v, w))$,
(ii) $\mathbb{G} = dx \otimes dx + dy \otimes dy + dz \otimes dz$,
(iii) $V(q) = \frac{1}{2}(x^2 + y^2 + z^2)$,
(iv) $F(v_q) = -udx(q) - vdy(q) - wdz(q)$, and
(v) D is defined such that $D^\perp = \operatorname{span}\left\{x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + (1 + x - y)\frac{\partial}{\partial z}\right\}$.

Observe that we can write $D = \operatorname{span}\{X_1, X_2\}$, where

$$X_1 = (1 + x)\frac{\partial}{\partial x} + (-x)\frac{\partial}{\partial y} + (-x)\frac{\partial}{\partial z},$$

$$X_2 = (y)\frac{\partial}{\partial x} + (1 - y)\frac{\partial}{\partial y} + (-y)\frac{\partial}{\partial z}.$$

Indeed, we have that $x \cdot (1 + x) + y \cdot (-x) + (1 + x - y) \cdot (-x) = 0$ and $x \cdot y + y \cdot (1 - y) + (1 + x - y) \cdot (-y) = 0$, and

$$\det\begin{bmatrix} x & y & 1 + x - y \\ 1 + x & -x & -x \\ y & 1 - y & -y \end{bmatrix} = 2x^2 - 2xy + 2x + 2y^2 - 2y + 1,$$

which can be verified is never zero.

Now, let us verify that the constraints are nonholonomic (in fact, totally nonholonomic). We compute

$$[X_1, X_2] = (-x - y)\frac{\partial}{\partial x} + (x + y)\frac{\partial}{\partial y} + (x + y)\frac{\partial}{\partial z},$$

and get

$$\det\begin{bmatrix} 1 + x & -x & -x \\ y & 1 - y & -y \\ -x - y & x + y & x + y \end{bmatrix} = x + y,$$

so $\{X_1, X_2, [X_1, X_2]\}$ is linearly independent everywhere except on the line $x + y = 0$. Hence, one bracket is not enough to show that D is totally nonholonomic. We compute another bracket

$$[X_1, [X_1, X_2]] = (-1 + x + y)\frac{\partial}{\partial x} + (1 - x - y)\frac{\partial}{\partial y} + (1 - x - y)\frac{\partial}{\partial z}.$$

Here, for $\{X_1, X_2, [X_1, [X_1, X_2]]\}$, we have

$$\det \begin{bmatrix} 1+x & -x & -x \\ y & 1-y & -y \\ -1+x+y & 1-x-y & 1-x-y \end{bmatrix} = -x-y+1,$$

so we have linear independence everywhere except on the line $x+y=1$. Now, since the solutions to $x+y=1$ do not coincide with any solutions of $x+y=0$, we can conclude that D is totally nonholonomic.

The equilibria of Σ are all $q \in Q$ such that $\mathbb{G}^{\#} \circ dV(q) \in D^{\perp}$. We solve

$$\begin{aligned} \mathbb{G}^{\#} \circ dV(q) &= \lambda \left(x\frac{\partial}{\partial x}(q) + y\frac{\partial}{\partial y}(q) + (1+x-y)\frac{\partial}{\partial z}(q) \right) \\ x\frac{\partial}{\partial x}(q) + y\frac{\partial}{\partial y}(q) + z\frac{\partial}{\partial z}(q) &= \lambda \left(x\frac{\partial}{\partial x}(q) + y\frac{\partial}{\partial y}(q) + (1+x-y)\frac{\partial}{\partial z}(q) \right) \\ &\Rightarrow \begin{cases} x = x\lambda \\ y = y\lambda \\ z = (1+x-y)\lambda \end{cases} . \end{aligned}$$

If $x=0$ and $y=0$, then z is arbitrary. If $x \neq 0$ or $y \neq 0$, then $\lambda = 1$ and so $z = 1+x-y$ for all x and y. Hence the set of equilibria is $\{(0,0,z) \mid z \in \mathbb{R}\} \cup \{(x,y,z) \mid x-y-z=-1\}$, shown in Figure 3.1. It is not a submanifold of Q.

The following is an example of a nonholonomic mechanical system which has an isolated equilibrium configuration. The original example is found in [6].

Example 3.18. Consider $\Sigma = (Q, \mathbb{G}, V, F, D)$, a C^{∞}-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, where

(i) $Q = \mathbb{R}^3$ with global coordinate charts $(\mathbb{R}^3, \phi = (x,y,z))$ and $(T\mathbb{R}^3, T\phi = (x,y,z,u,v,w))$,
(ii) $\mathbb{G} = dx \otimes dx + dy \otimes dy + dz \otimes dz$,
(iii) $V(q) = \frac{1}{2}(x-1)^2 + \frac{1}{2}(y-1)^2$,
(iv) $F(v_q) = -udx(q) - vdy(q) - wdz(q)$, and
(v) D is defined such that $D^{\perp} = \operatorname{span}\left\{ -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} + ((y-1)^2+z^2)\frac{\partial}{\partial z} \right\}$.

Observe that we can write $D = \operatorname{span}\{X_1, X_2\}$, where

$$X_1 = (x)\frac{\partial}{\partial x} + ((y-1)^2 + z^2 + y)\frac{\partial}{\partial y} + (-x)\frac{\partial}{\partial z},$$

$$X_2 = ((y-1)^2 + z^2)\frac{\partial}{\partial x} + (0)\frac{\partial}{\partial y} + (y)\frac{\partial}{\partial z}.$$

Indeed, we have that $(-y) \cdot x + x \cdot ((y-1)^2 + z^2 + y) + ((y-1)^2 + z^2) \cdot (-x) = 0$ and $(-y) \cdot ((y-1)^2 + z^2) + x \cdot 0 + ((y-1)^2 + z^2) \cdot y = 0$, and

$$\det \begin{bmatrix} -y & x & (y-1)^2+z^2 \\ x & (y-1)^2+z^2+y & -x \\ (y-1)^2+z^2 & 0 & y \end{bmatrix}$$
$$= y(-x^2 - y^3 + y^2 - yz^2 - y) + ((y-1)^2+z^2)(-x^2 - ((y-1)^2+z^2)((y-1)^2+y+z^2)),$$

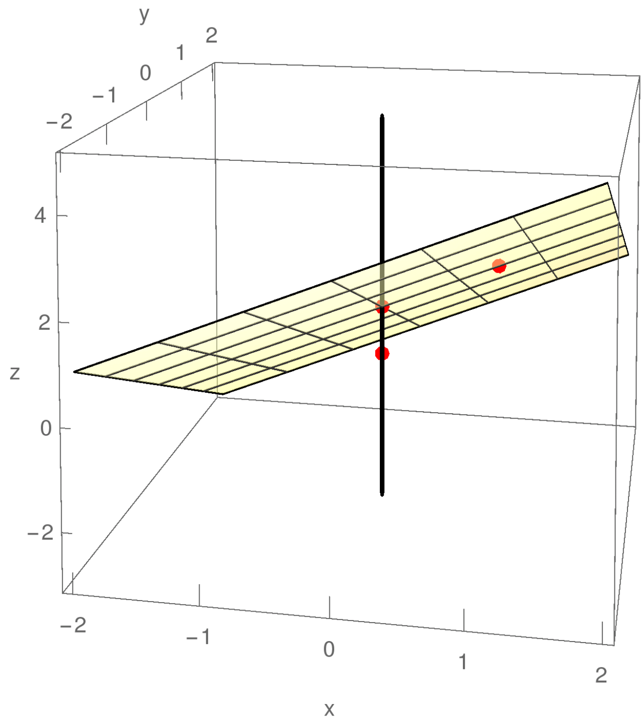

Figure 3.1: The set of equilibria for Σ in Example 3.17. Points $(0,0,0)$, $(0,0,1)$, and $(1,0,2)$ are indicated in red.

which can be verified is never zero.

Now, let us verify that the constraints are nonholonomic (in fact, totally nonholonomic). We compute

$$[X_1, X_2] = p_1(x,y,z)\frac{\partial}{\partial x} + p_2(x,y,z)\frac{\partial}{\partial y} + p_3(x,y,z)\frac{\partial}{\partial z},$$

where

$$\begin{aligned}
p_1(x,y,z) &= -(y-1)^2 - 2xz - z^2 + 2(y-1)\left((y-1)^2 + y + z^2\right),\\
p_2(x,y,z) &= -2yz,\\
p_3(x,y,z) &= 2(y-1)^2 + y + 2z^2,
\end{aligned}$$

and get

$$\begin{aligned}
&\det\begin{bmatrix} x & (y-1)^2 + z^2 + y & -x \\ (y-1)^2 + z^2 & 0 & y \\ p_1(x,y,z) & p_2(x,y,z) & p_3(x,y,z) \end{bmatrix}\\
&= 2\left(-1 + y - z^2\right)\left(1 - y + y^2 + z^2\right)^2,
\end{aligned}$$

and so it can be determined that $\{X_1, X_2, [X_1, X_2]\}$ is linearly independent everywhere except on $U_1 = \{(x,y,z) \in \mathbb{R}^3 \mid y = z^2 + 1\}$. Hence, one bracket is not enough to show that

D is totally nonholonomic. We compute another bracket

$$[X_1,[X_1,X_2]] = p_4(x,y,z)\frac{\partial}{\partial x} + p_5(x,y,z)\frac{\partial}{\partial y} + p_6(x,y,z)\frac{\partial}{\partial z},$$

where

$$\begin{aligned} p_4(x,y,z) &= (-1+y)^2 + z^2 - x(-2x - 2z + 4(-1+y)z) \\ &\quad - 2(-1+y)\left((-1+y)^2 + y + z^2\right) + \left((-1+y)^2 + y + z^2\right) \\ &\quad \cdot \left(-2(-1+y) + 2(1+2(-1+y))(-1+y) + 2\left((-1+y)^2 + y + z^2\right)\right), \\ p_5(x,y,z) &= 2xy + 2(1+2(-1+y))yz - 2z\left((-1+y)^2 + y + z^2\right) \\ &\quad - 2z\left(2(-1+y)^2 + y + 2z^2\right), \\ p_6(x,y,z) &= -(-1+y)^2 - 6xz - z^2 + (1+4(-1+y))\left((-1+y)^2 + y + z^2\right) \\ &\quad + 2(-1+y)\left((-1+y)^2 + y + z^2\right). \end{aligned}$$

Here, for $\{X_1, X_2, [X_1,[X_1,X_2]]\}$, we have

$$\begin{aligned} &\det\begin{bmatrix} x & (y-1)^2 + z^2 + y & -x \\ (y-1)^2 + z^2 & 0 & y \\ p_4(x,y,z) & p_5(x,y,z) & p_6(x,y,z) \end{bmatrix} \\ &= 2\left(1 - y + y^2 + z^2\right)(3 - 8y + 11y^2 - 8y^3 + 3y^4 + 6xz - 6xyz + 2xy^2z + 6z^2 \\ &\quad - 10yz^2 + 8y^2z^2 - 2y^3z^2 + 6xz^3 + 3z^4 - 2yz^4), \end{aligned}$$

and so it can be determined that we have linear independence everywhere except on

$$\begin{aligned} U_2 = \{(x,y,z) \in \mathbb{R}^3 \mid x = (&-3 + 11y - 22y^2 + 27y^3 - 22y^4 + 11y^5 - 3y^6 - 9z^2 + 24yz^2 \\ &- 35y^2z^2 + 28y^3z^2 - 13y^4z^2 + 2y^5z^2 - 9z^4 + 15yz^4 - 13y^2z^4 + 4y^3z^4 - 3z^6 + 2yz^6) \\ &/(6z - 12yz + 14y^2z - 8y^3z + 2y^4z + 12z^3 - 12yz^3 + 8y^2z^3 + 6z^5), z \neq 0\}. \end{aligned}$$

This region is shown in Figure 3.2. Let us compute yet another bracket

$$[X_2,[X_1,X_2]] = p_7(x,y,z)\frac{\partial}{\partial x} + p_8(x,y,z)\frac{\partial}{\partial y} + p_9(x,y,z)\frac{\partial}{\partial z},$$

where

$$\begin{aligned} p_7(x,y,z) &= 4(-1+y)yz + y(-2x - 2z + 4(-1+y)z) \\ &\quad - 2z\left((-1+y)^2 + z^2\right) - 2z\left(2(-1+y)^2 + y + 2z^2\right), \\ p_8(x,y,z) &= -2y^2 \\ p_9(x,y,z) &= 6yz. \end{aligned}$$

Here, for $\{X_1, X_2, [X_2,[X_1,X_2]]\}$, we have

$$\begin{aligned} &\det\begin{bmatrix} x & (y-1)^2 + z^2 + y & -x \\ (y-1)^2 + z^2 & 0 & y \\ p_7(x,y,z) & p_8(x,y,z) & p_9(x,y,z) \end{bmatrix} \\ &= -4yz\left(1 - y + y^2 + z^2\right)\left(3 - 3y + y^2 + 3z^2\right), \end{aligned}$$

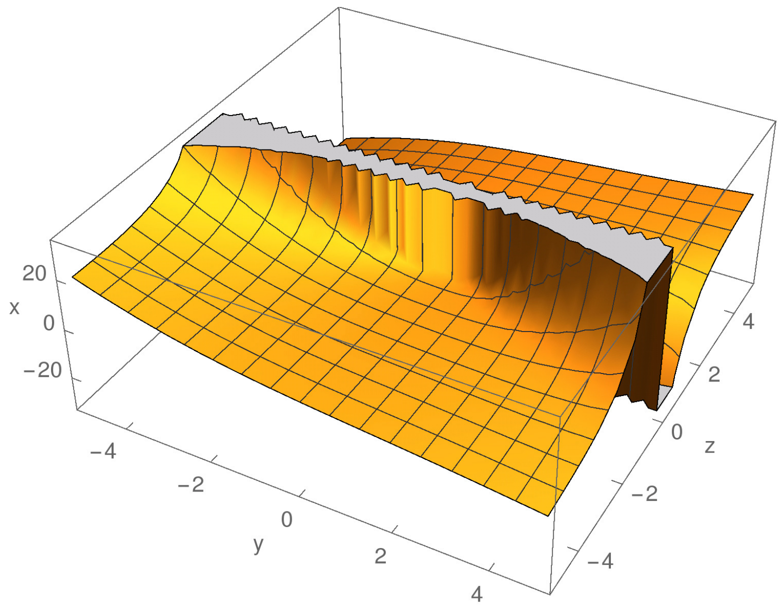

Figure 3.2: The region U_2.

so we have linear independence everywhere except on

$$U_3 = \{(x, y, z) \in \mathbb{R}^3 \mid y = 0\} \cup \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}.$$

Now, it can be verified that $U_1 \cap U_2 \cap U_3 = \emptyset$, so we conclude that D is totally nonholonomic.

The equilibria of Σ are all $q \in Q$ such that $\mathbb{G}^\# \circ dV(q) \in D^\perp$. We solve

$$\mathbb{G}^\# \circ dV(q) = \lambda \left(-y\frac{\partial}{\partial x}(q) + x\frac{\partial}{\partial y}(q) + ((y-1)^2 + z^2)\frac{\partial}{\partial z}(q) \right)$$

$$(x-1)\frac{\partial}{\partial x}(q) + (y-1)\frac{\partial}{\partial y}(q) + (0)\frac{\partial}{\partial z}(q) = \lambda \left(-y\frac{\partial}{\partial x}(q) + x\frac{\partial}{\partial y}(q) + ((y-1)^2 + z^2)\frac{\partial}{\partial z}(q) \right)$$

$$\Rightarrow \begin{cases} x - 1 = -y\lambda \\ y - 1 = x\lambda \\ 0 = ((y-1)^2 + z^2)\lambda \end{cases}.$$

Note that $\lambda = 0$ or $(y-1)^2 + z^2 = 0$. If $\lambda = 0$, then $x = 1$, $y = 1$ and z is arbitrary. If $(y-1)^2 + z^2 = 0$, then $(y-1)^2 = -z^2 \Rightarrow z = 0$, $y = 1$ and $x = 0$. Hence the set of equilibria is $\{(1, 1, z) \mid z \in \mathbb{R}\} \cup \{(0, 1, 0)\}$, shown in Figure 3.3. We have an isolated equilibrium at $(0, 1, 0)$.

We now make use of the well-known implicit function theorem. We state a version of it along the lines of [10].

Theorem 3.19 (Implicit Function Theorem)**.** *Let $F : U \to \mathbb{R}^k$ be a C^∞-function where $U \subseteq \mathbb{R}^n \times \mathbb{R}^k$ is open and consider the standard coordinates $(x^1, \ldots, x^n, y^1, \ldots, y^k)$ on U. Let $(x_0, y_0) \in U$ such that $F(x_0, y_0) = 0$. If det $\left(\left[\frac{\partial F^i}{\partial y^j} \right] \Big|_{(x_0, y_0)} \right) \neq 0$, then there exists*

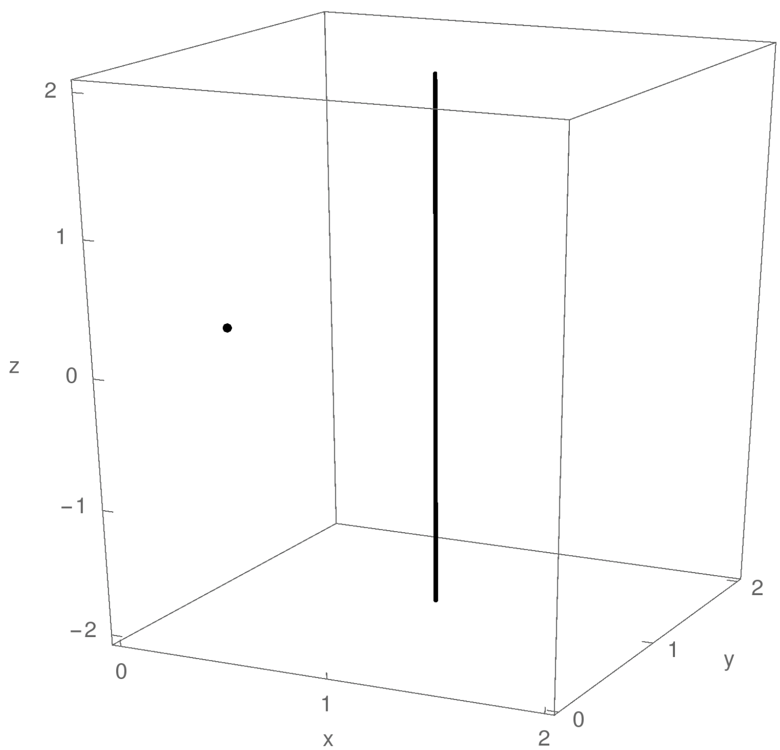

Figure 3.3: The set of equilibria for Σ in Example 3.18.

neighbourhoods $V \subseteq \mathbb{R}^n$ of x_0 and $W \subseteq \mathbb{R}^k$ of y_0 and a C^∞-function $\Phi : V \to W$ such that $F^{-1}(0) \cap (V \times W)$ is the graph of Φ, i.e. $F(x, y) = 0$ for $(x, y) \in V \times W$ if and only if $y = \Phi(x)$.

Proposition 3.20 (Understanding the Nature of Equilibria via the Implicit Function Theorem)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, and consider an equilibrium configuration q_0 for Σ. If the map $P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) : D(q_0) \to D(q_0)$ is invertible, then q_0 is not isolated and, in a neighbourhood of q_0, the set of equilibria forms an $(n-k)$-dimensional submanifold of Q, where n is the dimension of Q and k is the rank of D.*

Proof. Since D is a regular C^∞-distribution on Q, there exists a local chart (U, ϕ) of Q around q_0 and an admissible vector bundle chart (TU, ψ) for TQ such that

1. $\psi : TU \to \phi(U) \times \mathbb{R}^k \times \mathbb{R}^{n-k}$,
2. $\phi \circ \pi_{TQ} \circ \psi^{-1}(x, u, v) = x$, and
3. $\psi(TU \cap D) = \phi(U) \times \mathbb{R}^k \times \{0\}^{n-k}$.

Now consider the map

$$\begin{array}{rccl} \psi \circ P_D \circ \mathbb{G}^\# \circ dV \circ \phi^{-1} : & \phi(U) \subseteq \mathbb{R}^{n-k} \times \mathbb{R}^k & \to & \phi(U) \times \mathbb{R}^k \times \{0\}^{n-k} \\ & (x, y) & \mapsto & (x, y, F(x, y), 0) \end{array}$$

where $F : \phi(U) \subseteq \mathbb{R}^{n-k} \times \mathbb{R}^k \to \mathbb{R}^k$. We have that $P_D \circ \mathbb{G}^\# \circ dV(q_0) = 0$, so letting

$\phi(q_0) = (x_0, y_0)$, we have $F(x_0, y_0) = 0$. If $\det\left(\left[\frac{\partial F^i}{\partial y^j}\right]\Big|_{(x_0,y_0)}\right) \neq 0$ then, by the implicit function theorem, there exists neighbourhoods $V \subseteq \mathbb{R}^{n-k}$ of x_0 and $W \subseteq \mathbb{R}^k$ of y_0 and a C^∞-function $\Phi : V \to W$ such that $F^{-1}(0) \cap (V \times W)$ is the graph of Φ. In other words, the submanifold $Q_0 = \phi^{-1}(\{(x, \Phi(x)) \mid x \in V\})$ of Q satisfies $P_D \circ \mathbb{G}^\# \circ dV(q) = 0$ for all $q \in Q_0$. Hence q_0 is not isolated, and in a neighbourhood of q_0, the set of equilibria forms an $(n-k)$-dimensional submanifold of Q.

So it remains to show that $\det\left(\left[\frac{\partial F^i}{\partial y^j}\right]\Big|_{(x_0,y_0)}\right) \neq 0$. Without loss of generality, we can choose the chart $(U, \phi = (q^1, \dots, q^n))$ such that $\{\frac{\partial}{\partial q^1}(q_0), \dots, \frac{\partial}{\partial q^n}(q_0)\}$ forms an orthogonal basis for $T_{q_0}Q$ and $D_{q_0} = \operatorname{span}\{\frac{\partial}{\partial q^1}(q_0), \dots, \frac{\partial}{\partial q^k}(q_0)\}$. Now consider the map $P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0)$. Let us write this out in coordinates (U, ϕ):

$$\begin{aligned}
&P_D \circ \overset{\mathbb{G}}{\nabla}_{v_{q_0}}(P_D \circ \mathbb{G}^\# \circ dV)(q_0)\\
&\quad = P_D \circ \left(\frac{\partial (P_D \circ \mathbb{G}^\# \circ dV)^i}{\partial q^j}(q_0) v^j_{q_0} + \overset{\mathbb{G}}{\Gamma}{}^i_{js}(q_0) v^j_{q_0} (P_D \circ \mathbb{G}^\# \circ dV)^s(q_0)\right) \frac{\partial}{\partial q^i}(q_0),
\end{aligned}$$

or

$$\begin{aligned}
&P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0)\\
&\quad = P_D \circ \left(\frac{\partial (P_D \circ \mathbb{G}^\# \circ dV)^i}{\partial q^j}(q_0) + \overset{\mathbb{G}}{\Gamma}{}^i_{js}(q_0) (P_D \circ \mathbb{G}^\# \circ dV)^s(q_0)\right) \frac{\partial}{\partial q^i}(q_0) \otimes dq^j(q_0).
\end{aligned}$$

Note that since q_0 is an equilibrium configuration, $P_D \circ \mathbb{G}^\# \circ dV(q_0) = 0$ and so the term containing the Christoffel symbols is zero. Also note that we are only interested in the map $P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) : D(q_0) \to D(q_0)$. More precisely, this is $\pi_D(q_0) \circ P_D(q_0) \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) \circ i_D(q_0)$, where π_D is the orthogonal projection onto D and i_D is the canonical inclusion map. In coordinates, we have

$$\begin{aligned}
&\pi_D \circ P_D \circ \left(\frac{\partial (P_D \circ \mathbb{G}^\# \circ dV)^i}{\partial q^j}(q_0) + \overset{\mathbb{G}}{\Gamma}{}^i_{js}(q_0) \underbrace{(P_D \circ \mathbb{G}^\# \circ dV)^s(q_0)}_{0}\right) \frac{\partial}{\partial q^i}(q_0) \otimes dq^j(q_0), \quad j \in \{1, \dots, k\}\\
&= \frac{\partial (P_D \circ \mathbb{G}^\# \circ dV)^i}{\partial q^j}(q_0) \frac{\partial}{\partial q^i}(q_0) \otimes dq^j(q_0), \quad i, j \in \{1, \dots, k\}.
\end{aligned}$$

Since we have that $P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) : D(q_0) \to D(q_0)$ is invertible, equivalently we have

$$\det\left(\left[\frac{\partial (P_D \circ \mathbb{G}^\# \circ dV)^i}{\partial q^j}\right]\Big|_{q_0}\right) \neq 0.$$

Choosing ψ such that $\psi\Big|_{q_0} = T\phi\Big|_{q_0}$, the invertibility of the map $\left[\frac{\partial (P_D \circ \mathbb{G}^\# \circ dV)^i}{\partial q^j}\right]\Big|_{q_0}$ at q_0 is equivalent to the invertibility of $\left[\frac{\partial F^i}{\partial y^j}\right]\Big|_{(x_0,y_0)}$ at (x_0, y_0). Hence $\det\left(\left[\frac{\partial F^i}{\partial y^j}\right]\Big|_{(x_0,y_0)}\right) \neq 0$, and so the result follows. $\square$

Example 3.21. Consider again Example 3.17. We have

$$\left[\mathbb{G}^{\#}\right] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } [dV] = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

At (x, y, z), the projection of a tangent vector (u, v, w) onto $D(x, y, z)$ is

$$\begin{aligned}
&\begin{pmatrix} u \\ v \\ w \end{pmatrix} - \frac{(u, v, w) \cdot (x, y, 1 + x - y)}{\|(x, y, 1 + x - y)\|^2} \begin{pmatrix} x \\ y \\ 1 + x - y \end{pmatrix} \\
&= \begin{pmatrix} u \\ v \\ w \end{pmatrix} - \frac{xu + yv + (1 + x - y)w}{x^2 + y^2 + (1 + x - y)^2} \begin{pmatrix} x \\ y \\ 1 + x - y \end{pmatrix} \\
&= \frac{\begin{pmatrix} y^2 + (1 + x - y)^2 & -xy & -x(1 + x - y) \\ -xy & x^2 + (1 + x - y)^2 & -y(1 + x - y) \\ -x(1 + x - y) & -y(1 + x - y) & x^2 + y^2 \end{pmatrix}}{x^2 + y^2 + (1 + x - y)^2} \begin{pmatrix} u \\ v \\ w \end{pmatrix},
\end{aligned}$$

so we have

$$[P_D(x, y, z)] = \frac{1}{x^2 + y^2 + (1 + x - y)^2} \begin{pmatrix} y^2 + (1 + x - y)^2 & -xy & -x(1 + x - y) \\ -xy & x^2 + (1 + x - y)^2 & -y(1 + x - y) \\ -x(1 + x - y) & -y(1 + x - y) & x^2 + y^2 \end{pmatrix}.$$

Hence,

$$\left[P_D \circ \mathbb{G}^{\#} \circ dV\right] = \begin{pmatrix} p_1(x, y, z) \\ p_2(x, y, z) \\ p_3(x, y, z) \end{pmatrix},$$

where

$$\begin{aligned}
p_1(x, y, z) &= \frac{-xy^2 + x\left((1 + x - y)^2 + y^2\right) - x(1 + x - y)z}{x^2 + (1 + x - y)^2 + y^2}, \\
p_2(x, y, z) &= \frac{-x^2 y + \left(x^2 + (1 + x - y)^2\right) y - (1 + x - y)yz}{x^2 + (1 + x - y)^2 + y^2}, \\
p_3(x, y, z) &= \frac{-x^2(1 + x - y) - (1 + x - y)y^2 + \left(x^2 + y^2\right) z}{x^2 + (1 + x - y)^2 + y^2}.
\end{aligned}$$

Taking partial derivatives, we get

$$\begin{aligned}
\frac{\partial p_1}{\partial x} &= \frac{2x(1 + x - y) + (1 + x - y)^2 - xz - (1 + x - y)z}{x^2 + (1 + x - y)^2 + y^2} \\
&\quad - \frac{(2x + 2(1 + x - y))\left(-xy^2 + x\left((1 + x - y)^2 + y^2\right) - x(1 + x - y)z\right)}{\left(x^2 + (1 + x - y)^2 + y^2\right)^2}, \\
\frac{\partial p_1}{\partial y} &= \frac{-2xy + x(-2(1 + x - y) + 2y) + xz}{x^2 + (1 + x - y)^2 + y^2} \\
&\quad - \frac{(-2(1 + x - y) + 2y)\left(-xy^2 + x\left((1 + x - y)^2 + y^2\right) - x(1 + x - y)z\right)}{\left(x^2 + (1 + x - y)^2 + y^2\right)^2}, \\
\frac{\partial p_1}{\partial z} &= -\frac{x(1 + x - y)}{x^2 + (1 + x - y)^2 + y^2},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial p_2}{\partial x} &= \frac{-2xy + (2x + 2(1+x-y))y - yz}{x^2 + (1+x-y)^2 + y^2} \\
&\quad - \frac{(2x + 2(1+x-y))\left(-x^2 y + \left(x^2 + (1+x-y)^2\right) y - (1+x-y)yz\right)}{\left(x^2 + (1+x-y)^2 + y^2\right)^2}, \\
\frac{\partial p_2}{\partial y} &= \frac{(1+x-y)^2 - 2(1+x-y)y - (1+x-y)z + yz}{x^2 + (1+x-y)^2 + y^2} \\
&\quad - \frac{(-2(1+x-y) + 2y)\left(-x^2 y + \left(x^2 + (1+x-y)^2\right) y - (1+x-y)yz\right)}{\left(x^2 + (1+x-y)^2 + y^2\right)^2}, \\
\frac{\partial p_2}{\partial z} &= -\frac{(1+x-y)y}{x^2 + (1+x-y)^2 + y^2}, \\
\\
\frac{\partial p_3}{\partial x} &= \frac{-x^2 - 2x(1+x-y) - y^2 + 2xz}{x^2 + (1+x-y)^2 + y^2} \\
&\quad - \frac{(2x + 2(1+x-y))\left(-x^2(1+x-y) - (1+x-y)y^2 + \left(x^2 + y^2\right) z\right)}{\left(x^2 + (1+x-y)^2 + y^2\right)^2}, \\
\frac{\partial p_3}{\partial y} &= \frac{x^2 - 2(1+x-y)y + y^2 + 2yz}{x^2 + (1+x-y)^2 + y^2} \\
&\quad - \frac{(-2(1+x-y) + 2y)\left(-x^2(1+x-y) - (1+x-y)y^2 + \left(x^2 + y^2\right) z\right)}{\left(x^2 + (1+x-y)^2 + y^2\right)^2}, \\
\frac{\partial p_3}{\partial z} &= \frac{x^2 + y^2}{x^2 + (1+x-y)^2 + y^2}.
\end{aligned}$$

Let us now compute $\pi_D(q_0) \circ P_D(q_0) \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) \circ i_D(q_0)$ at various equilibria q_0 (shown in red in Figure 3.1). Note that

$$\left[\overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0)\right] = \begin{pmatrix} \frac{\partial p_1}{\partial x}(q_0) & \frac{\partial p_1}{\partial y}(q_0) & \frac{\partial p_1}{\partial z}(q_0) \\ \frac{\partial p_2}{\partial x}(q_0) & \frac{\partial p_2}{\partial y}(q_0) & \frac{\partial p_2}{\partial z}(q_0) \\ \frac{\partial p_3}{\partial x}(q_0) & \frac{\partial p_3}{\partial y}(q_0) & \frac{\partial p_3}{\partial z}(q_0) \end{pmatrix}.$$

1. At $q_0 = (0,0,1)$, we have $D(q_0) = \operatorname{span}\left\{\frac{\partial}{\partial x}(q_0), \frac{\partial}{\partial y}(q_0)\right\}$ and $D^\perp(q_0) = \operatorname{span}\left\{\frac{\partial}{\partial z}(q_0)\right\}$. We can write

$$[i_D(q_0)] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad [P_D(q_0)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad [\pi_D(q_0)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

so

$$\left[\pi_D(q_0) \circ P_D(q_0) \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) \circ i_D(q_0)\right] = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This is not invertible. Hence this is consistent with the fact that there is no manifold structure at $(0,0,1)$.

2. At $q_0 = (0,0,0)$, we have $D(q_0) = \operatorname{span}\left\{\frac{\partial}{\partial x}(q_0), \frac{\partial}{\partial y}(q_0)\right\}$ and $D^\perp(q_0) = \operatorname{span}\left\{\frac{\partial}{\partial z}(q_0)\right\}$. Again, we can write

$$[i_D(q_0)] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad [P_D(q_0)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad [\pi_D(q_0)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

so

$$\left[\pi_D(q_0) \circ P_D(q_0) \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) \circ i_D(q_0)\right] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This is invertible. Hence, in a neighbourhood of $(0, 0, 0)$, the set of equilibria forms a one-dimensional submanifold of Q.

3. At $q_0 = (1, 0, 2)$, we have $D(q_0) = \operatorname{span}\left\{2\frac{\partial}{\partial x}(q_0) - \frac{\partial}{\partial y}(q_0) - \frac{\partial}{\partial z}(q_0), \frac{\partial}{\partial y}(q_0)\right\}$ and $D^\perp(q_0) = \operatorname{span}\left\{\frac{\partial}{\partial x}(q_0) + 2\frac{\partial}{\partial z}(q_0)\right\}$. In other words, $e_1 = (2, -1, -1)$ and $e_2 = (0, 1, 0)$ form a basis for $D(q_0)$ and $e_3 = (1, 0, 2)$ forms a basis for $D^\perp(q_0)$. In this basis, we can write

$$[i_D(q_0)]_{\{e_1,e_2,e_3\}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad [P_D(q_0)]_{\{e_1,e_2,e_3\}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$[\pi_D(q_0)]_{\{e_1,e_2,e_3\}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The transformation matrix from the basis $\{e_1, e_2, e_3\}$ to standard coordinates $\{u, v, w\}$ is

$$T = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix},$$

and its inverse is

$$T^{-1} = \frac{1}{5}\begin{pmatrix} 2 & 0 & -1 \\ 2 & 5 & -1 \\ 1 & 0 & 2 \end{pmatrix}.$$

Hence we have

$$\left[\pi_D(q_0) \circ P_D(q_0) \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathbb{G}^\# \circ dV)(q_0) \circ i_D(q_0)\right]_{\{e_1,e_2,e_3\}} = \frac{1}{5}\begin{pmatrix} 4 & -1 \\ 4 & -1 \end{pmatrix}.$$

This is not invertible. Hence, in a neighbourhood of $(1, 0, 2)$, we cannot conclude whether the set of equilibria has a manifold structure. Note that from Example 3.17 and referring to Figure 3.1, however, we know the set of equilibria around $(1, 0, 2)$ does form a two-dimensional submanifold of Q.

3.4. Forced Affine Connection Systems

We now consider forced affine connection systems, a generalization of forced simple mechanical systems with constraints. In the next chapter, we will look at the linearization of forced affine connection systems, and then apply these results to obtain the linearization for forced simple mechanical systems with constraints.

Definition 3.22 (Vector Force)**.** Let Q be a C^∞-manifold and consider $Y : TQ \to TQ$. Y is a C^∞**-vector force** if Y is a C^∞-vector bundle map over id_Q.

Note that our notion of a vector force is assumed to be time-independent.

Definition 3.23 (Forced Affine Connection System)**.** A C^∞**-forced affine connection system** is a 4-tuple (Q, ∇, Y, D), where
 (i) Q is a C^∞-manifold,
 (ii) ∇ is a C^∞-affine connection on Q,
 (iii) Y is a C^∞-vector force, and
 (iv) D is a regular C^∞-distribution to which ∇ restricts, and for which Y is D-valued.

Definition 3.24 (Equations of Motion)**.** Let $\Sigma = (Q, \nabla, Y, D)$ be a C^∞-forced affine connection system. The **equations of motion** for Σ are

$$\nabla_{\gamma'(t)}\gamma'(t) = Y(\gamma'(t)), \tag{3.4.1}$$

where $\gamma : I \to Q$ and $\gamma'(t_0) \in D_{\gamma(t_0)}$ for some (and therefore, all) $t_0 \in I$.

Remark 3.25 (Reinterpreting Constrained Mechanical Systems). For a forced simple mechanical system with regular constraints $(Q, \mathbb{G}, V, F, D)$, we have the forced affine connection system $(Q, \overset{D}{\nabla}, P_D(\mathbb{G}^\# \circ F - \mathrm{grad}V), D)$. Note that the equations of motion for the forced simple mechanical system with regular constraints coincide with the equations of motion for the forced affine connection system.

Definition 3.26 (Equilibrium Configuration)**.** Let $\Sigma = (Q, \nabla, Y, D)$ be a C^∞-forced affine connection system. A point $q_0 \in Q$ is an **equilibrium configuration for** Σ if the trivial curve $\gamma(t) = q_0$ satisfies the equations of motion (3.4.1).

Similarly to forced simple mechanical systems, we can consider the associated vector field for forced affine connection systems. Let $\Sigma = (Q, \nabla, Y, D)$ be a C^∞-forced affine connection system and consider a local chart $(U, \phi = (q^1, \ldots, q^n))$ of Q. For a curve $\gamma : I \to Q$ satisfying the equations of motion (3.4.1), let $\phi \circ \gamma(t) = q(t) = (q^1(t), \ldots, q^n(t))$. Write Y as $Y(v_q) = Y^i(v_q)\frac{\partial}{\partial q^i}(q)$ and let Γ^i_{jk} be the Christoffel symbols for ∇ in the chart (U, ϕ). The equations of motion in coordinates are

$$\ddot{q}^i + \Gamma^i_{jk}\dot{q}^j\dot{q}^k = Y^i,$$

or

$$\begin{cases} \dot{q}^i = v^i \\ \dot{v}^i = -\Gamma^i_{jk}v^jv^k + Y^i, \end{cases}$$

where $v(t) = (v^1(t), \ldots, v^n(t)) = \dot{q}(t)$. It can be shown that this gives rise to a well-defined vector field X on D, i.e. $X : D \to TD$ defined by

$$X(v_q) \triangleq S(v_q) + \mathrm{vlft}(Y)(v_q), \tag{3.4.2}$$

where the integral curves of X, projected onto Q (using the canonical projection), are curves satisfying the equations of motion (3.4.1). We will call X the **associated vector field for** Σ. In coordinates this can be written as

$$X\Big|_{TU}(v_q) = v^i\frac{\partial}{\partial q^i}(v_q) + \left(-\Gamma^i_{jk}(q)v^jv^k + Y^i(v_q)\right)\frac{\partial}{\partial v^i}(v_q),$$

where $(TU, T\phi = (q^1, \dots, q^n, v^1, \dots, v^n))$ is the corresponding local chart of TQ for (U, ϕ).

Exactly as before, an equilibrium configuration $q_0 \in Q$ gives rise to an equilibrium point 0_{q_0} for the vector field X, and stability notions can be made for the equilibrium configuration q_0 by reference to the equilibrium point 0_{q_0}.

Lemma 3.27 (Characterization of Equilibrium Configurations)**.** *Let $\Sigma = (Q, \nabla, Y, D)$ be a C^∞-forced affine connection system. A point $q_0 \in Q$ is an equilibrium configuration for Σ if and only if $Y(0_{q_0}) = 0$.*

Proof. The proof is immediate by considering the equations of motion (3.4.1) and setting $\nabla_{\gamma'(t)}\gamma'(t) = 0$. $\square$

Chapter 4

Nonholonomic Mechanical Systems: Linearization about Equilibria

In this chapter, we obtain the linearization of a forced affine connection system and then apply this to nonholonomic mechanical systems. We then examine the validity of an alternative approach to linearization.

4.1. Linearization of a Forced Affine Connection System

We will require Lemma 2.8 for decomposing the state space of the linearization.

Lemma 4.1 (Tangent Lift of the Geodesic Spray)**.** *Let $\Sigma = (Q, \nabla, W, D)$ be a C^∞-forced affine connection system, and consider an equilibrium configuration $q_0 \in Q$ for Σ. Then, for all $v_1 \oplus v_2 \in T_{q_0}Q \oplus T_{q_0}Q \cong T_{0_{q_0}}TQ$, we have*

$$S^T(v_1 \oplus v_2) = (0 \oplus 0 \oplus v_2 \oplus 0)_{v_1 \oplus v_2}.$$

Proof. The proof is exactly the same as that of Lemma 2.9. □

Lemma 4.2 (Tangent Lift of the Vector Force)**.** *Let $\Sigma = (Q, \nabla, W, D)$ be a C^∞-forced affine connection system, and consider an equilibrium configuration $q_0 \in Q$ for Σ. Then, for all $v_1 \oplus v_2 \in T_{q_0}Q \oplus T_{q_0}Q \cong T_{0_{q_0}}TQ$, we have*

$$\mathit{vlft}(W)^T(v_1 \oplus v_2) = (0 \oplus 0 \oplus 0 \oplus (d_1W(0_{q_0})(v_1) + d_2W(0_{q_0})(v_2)))_{v_1 \oplus v_2},$$

where we have defined $d_1W(0_{q_0}), d_2W(0_{q_0}) \in L(T_{q_0}Q; T_{q_0}Q)$ such that given a local chart $(U, \phi = (q^1, \dots, q^n))$ of Q and corresponding local chart $(TU, T\phi = (q^1, \dots, q^n, v^1, \dots, v^n))$ of TQ,

$$[d_1W(0_{q_0})] = \left[\frac{\partial W^i}{\partial q^j}(0_{q_0})\right],$$
$$[d_2W(0_{q_0})] = \left[\frac{\partial W^i}{\partial v^j}(0_{q_0})\right].$$

Proof. The proof is analogous to that of Lemma 2.11. □

Proposition 4.3 (Linearization of a Forced Affine Connection System)**.** *Let* $\Sigma = (Q, \nabla, W, D)$ *be a* C^∞*-forced affine connection system and consider an equilibrium configuration* $q_0 \in Q$ *for* Σ*. Then the linearization of (3.4.2) at* 0_{q_0} *in the decomposition* $T_{0_{q_0}}TQ \cong T_{q_0}Q \oplus T_{q_0}Q$ *is given by*

$$A_\Sigma(q_0) = \begin{pmatrix} 0 & id_{T_{q_0}Q} \\ d_1W(0_{q_0}) & d_2W(0_{q_0}) \end{pmatrix},$$

and the linearized equations of motion are

$$\begin{pmatrix} \dot{q}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} 0 & id_{T_{q_0}Q} \\ d_1W(0_{q_0}) & d_2W(0_{q_0}) \end{pmatrix} \begin{pmatrix} q(t) \\ v(t) \end{pmatrix}.$$

Proof. This follows from combining the results of Lemmas 2.8, 4.1 and 4.2. □

4.2. Linearization of a Nonholonomic Mechanical System

In contrast to the unconstrained mechanical systems we considered where equilibria corresponded to critical points of the potential function, this is no longer true for constrained mechanical systems. While critical points are still equilibra, they are in general not the only ones. In other words, we must linearize the equations of motion about a point which is not necessarily a critical point of the potential function. Before looking at this more general case, let us first restrict ourselves to equilibria which are critical points.

Proposition 4.4 (Linearization of a Nonholonomic Mechanical System at a Critical Point)**.** *Let* $\Sigma = (Q, \mathbb{G}, V, F, D)$ *be a* C^∞*-forced simple mechanical system with regular constraints such that* $F(Z(TQ)) = Z(T^*Q)$*, and consider an equilibrium configuration* $q_0 \in Q$ *for* Σ *such that* $dV(q_0) = 0$*. Then the linearized equations of motion at* 0_{q_0} *in the decomposition* $T_{0_{q_0}}TQ \cong T_{q_0}Q \oplus T_{q_0}Q$ *are given by*

$$\begin{pmatrix} \dot{q}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} 0 & id_{T_{q_0}Q} \\ -P_D \circ \mathbb{G}^\# \circ HessV^\flat(q_0) + P_D \circ \mathbb{G}^\# \circ d_1F(0_{q_0}) & P_D \circ \mathbb{G}^\# \circ d_2F(0_{q_0}) \end{pmatrix} \begin{pmatrix} q(t) \\ v(t) \end{pmatrix}.$$

Proof. The equations of motion are

$$\overset{D}{\nabla}_{\gamma'(t)}\gamma'(t) = P_D(\mathbb{G}^\# \circ F(\gamma'(t)) - \mathrm{grad}V(\gamma(t))),$$

and $\gamma'(t_0) \in D_{\gamma(t_0)}$ for some $t_0 \in I$. This can be linearized using Proposition 4.3 by considering $\nabla = \overset{D}{\nabla}$ and $W(v_q) = P_D(\mathbb{G}^\# \circ F(v_q) - \mathrm{grad}V(q))$.

Consider $(U, \phi = (q^1, \dots, q^n))$, a local chart of Q around q_0. We write out

$$\begin{aligned}
W\Big|_{TU}(v_q) &= W^i(v_q)\frac{\partial}{\partial q^i}(q) \\
&= P_D(\mathbb{G}^\# \circ F(v_q) - \underbrace{\mathrm{grad}V}_{\mathbb{G}^\# \circ dV}(q)) \\
&= P_D{}^i_j(q)\frac{\partial}{\partial q^i}(q) \otimes dq^j(q)\left(G^{st}(q)F_t(v_q)\frac{\partial}{\partial q^s}(q) - G^{st}(q)\frac{\partial V}{\partial q^t}(q)\frac{\partial}{\partial q^s}(q)\right) \\
&= P_D{}^i_j\frac{\partial}{\partial q^i} \otimes dq^j\left(G^{st}F_t\frac{\partial}{\partial q^s} - G^{st}\frac{\partial V}{\partial q^t}\frac{\partial}{\partial q^s}\right)\Big|_{v_q} \\
&= P_D{}^i_j G^{jk}F_k\frac{\partial}{\partial q^i} - P_D{}^i_j G^{jk}\frac{\partial V}{\partial q^k}\frac{\partial}{\partial q^i}\Big|_{v_q} \\
&= \left(P_D{}^i_j(q)G^{jk}(q)F_k(v_q) - P_D{}^i_j(q)G^{jk}(q)\frac{\partial V}{\partial q^k}(q)\right)\frac{\partial}{\partial q^i}(q),
\end{aligned}$$

which gives us that $W^i(v_q) = P_D{}^i_j(q)G^{jk}(q)F_k(v_q) - P_D{}^i_j(q)G^{jk}(q)\frac{\partial V}{\partial q^k}(q)$. Now, we compute

$$[d_1W(0_{q_0})] = \left[\frac{\partial W^i}{\partial q^j}(0_{q_0})\right] = \left[\frac{\partial}{\partial q^j}\left(P_D{}^i_s G^{st}F_t - P_D{}^i_s G^{st}\frac{\partial V}{\partial q^t}\right)(0_{q_0})\right]$$

by first computing the derivative

$$\begin{aligned}
\frac{\partial}{\partial q^j}\left(P_D{}^i_s G^{st}F_t - P_D{}^i_s G^{st}\frac{\partial V}{\partial q^t}\right) &= \frac{\partial P_D{}^i_s}{\partial q^j}G^{st}F_t + P_D{}^i_s\frac{\partial G^{st}}{\partial q^j}F_t + P_D{}^i_s G^{st}\frac{\partial F_t}{\partial q^j} \\
&\quad - \frac{\partial P_D{}^i_s}{\partial q^j}G^{st}\frac{\partial V}{\partial q^t} - P_D{}^i_s\frac{\partial G^{st}}{\partial q^j}\frac{\partial V}{\partial q^t} - P_D{}^i_s G^{st}\frac{\partial^2 V}{\partial q^j \partial q^t}
\end{aligned}$$

and now evaluating at 0_{q_0}

$$\begin{aligned}
&\frac{\partial}{\partial q^j}\left(P_D{}^i_s G^{st}F_t - P_D{}^i_s G^{st}\frac{\partial V}{\partial q^t}\right)(0_{q_0}) \\
&= \frac{\partial P_D{}^i_s}{\partial q^j}(q_0)G^{st}(q_0)\underbrace{F_t(0_{q_0})}_{0} + P_D{}^i_s(q_0)\frac{\partial G^{st}}{\partial q^j}(q_0)\underbrace{F_t(0_{q_0})}_{0} + P_D{}^i_s(q_0)G^{st}(q_0)\frac{\partial F_t}{\partial q^j}(0_{q_0}) \\
&\quad - \frac{\partial P_D{}^i_s}{\partial q^j}(q_0)G^{st}(q_0)\underbrace{\frac{\partial V}{\partial q^t}(q_0)}_{0} - P_D{}^i_s(q_0)\frac{\partial G^{st}}{\partial q^j}(q_0)\underbrace{\frac{\partial V}{\partial q^t}(q_0)}_{0} - P_D{}^i_s(q_0)G^{st}(q_0)\frac{\partial^2 V}{\partial q^j \partial q^t}(q_0) \\
&= P_D{}^i_s(q_0)G^{st}(q_0)\frac{\partial F_t}{\partial q^j}(0_{q_0}) - P_D{}^i_s(q_0)G^{st}(q_0)\frac{\partial^2 V}{\partial q^j \partial q^t}(q_0).
\end{aligned} \tag{4.2.1}$$

So we have that

$$d_1W(0_{q_0})(v_1) = P_D \circ \mathbb{G}^\# \circ d_1F(0_{q_0})(v_1) - P_D \circ \mathbb{G}^\# \circ \mathrm{Hess}V^\flat(q_0)(v_1).$$

Next we compute

$$[d_2W(0_{q_0})] = \left[\frac{\partial W^i}{\partial v^j}(0_{q_0})\right] = \left[\frac{\partial}{\partial v^j}\left(P_D{}^i_s G^{st}F_t - P_D{}^i_s G^{st}\frac{\partial V}{\partial q^t}\right)(0_{q_0})\right]$$

again by first computing the derivative

$$\begin{aligned}\frac{\partial}{\partial v^j}\left(P_D{}^i_s G^{st} F_t - P_D{}^i_s G^{st}\frac{\partial V}{\partial q^t}\right) &= \underbrace{\frac{\partial P_D{}^i_s}{\partial v^j}}_{0} G^{st} F_t + P_D{}^i_s \underbrace{\frac{\partial G^{st}}{\partial v^j}}_{0} F_t + P_D{}^i_s G^{st}\frac{\partial F_t}{\partial v^j} \\ &\quad - \underbrace{\frac{\partial P_D{}^i_s}{\partial v^j}}_{0} G^{st}\frac{\partial V}{\partial q^t} - P_D{}^i_s \underbrace{\frac{\partial G^{st}}{\partial v^j}}_{0}\frac{\partial V}{\partial q^t} - P_D{}^i_s G^{st}\underbrace{\frac{\partial^2 V}{\partial v^j \partial q^t}}_{0} \\ &= P_D{}^i_s G^{st}\frac{\partial F_t}{\partial v^j}\end{aligned}$$

and then evaluating at 0_{q_0}

$$\frac{\partial}{\partial v^j}\left(P_D{}^i_s G^{st} F_t - P_D{}^i_s G^{st}\frac{\partial V}{\partial q^t}\right)(0_{q_0}) = P_D{}^i_s(q_0) G^{st}(q_0)\frac{\partial F_t}{\partial v^j}(0_{q_0}).$$

In this case, we get that

$$d_2 W(0_{q_0})(v_2) = P_D \circ \mathbb{G}^{\#} \circ d_2 F(0_{q_0})(v_2).$$

Hence the linearization is

$$\begin{pmatrix}\dot q(t)\\ \dot v(t)\end{pmatrix} = \begin{pmatrix} 0 & \mathrm{id}_{T_{q_0}Q} \\ -P_D \circ \mathbb{G}^{\#} \circ \mathrm{Hess}V^{\flat}(q_0) + P_D \circ \mathbb{G}^{\#} \circ d_1 F(0_{q_0}) & P_D \circ \mathbb{G}^{\#} \circ d_2 F(0_{q_0})\end{pmatrix}\begin{pmatrix} q(t)\\ v(t)\end{pmatrix}.$$

$\square$

Remark 4.5. This is the same form as the linearization for the unconstrained case, except that now we have the projection map P_D in front of each term.

Now, let us consider the general case. First, we prove the following lemma.

Lemma 4.6 (Covariant Derivative of the Projected Potential Force at an Equilibrium Configuration in Coordinates)**.** *Consider the D-valued vector field $P_D \circ \mathrm{grad}V : Q \to D$. Let $v_q \in TQ$ where q is an equilibrium configuration, i.e. we have $P_D \circ \mathrm{grad}V(q) = 0$. Then, for a local chart $(U, \phi = (q^1, \dots, q^n))$ of Q around q and corresponding local chart $(TU, T\phi = (q^1, \dots, q^n, v^1, \dots, v^n))$ of TQ, we have*

$$\begin{aligned}&\overset{\mathbb{G}}{\nabla}_{v_q}\left(P_D \circ \mathrm{grad}V\right)(q) \\ &\quad = \left(\frac{\partial P_D{}^i_j}{\partial q^k}(q) G^{jt}(q)\frac{\partial V}{\partial q^t}(q) + P_D{}^i_j(q)\frac{\partial G^{jt}}{\partial q^k}(q)\frac{\partial V}{\partial q^t}(q) + P_D{}^i_j(q) G^{jt}(q)\frac{\partial^2 V}{\partial q^k \partial q^t}(q)\right) v^k \frac{\partial}{\partial q^i}(q).\end{aligned}$$

Proof. By direct computation, we get

$$
\begin{aligned}
&\overset{\mathbb{G}}{\nabla}_{v_q}\left(P_D\circ \operatorname{grad}V\right)(q)\\
&=\left(\overset{\mathbb{G}}{\nabla}_{v_q}P_D\right)(\underbrace{\operatorname{grad}V}_{\mathbb{G}^\#\circ dV}(q))+P_D\left(\left(\overset{\mathbb{G}}{\nabla}_{v_q}\underbrace{\operatorname{grad}V}_{\mathbb{G}^\#\circ dV}\right)(q)\right)\\
&=\left(\frac{\partial {P_D}^i_j}{\partial q^k}(q)v^k+{P_D}^q_j(q)\overset{\mathbb{G}}{\Gamma}{}^i_{qk}(q)v^k-{P_D}^i_k(q)\overset{\mathbb{G}}{\Gamma}{}^k_{pj}(q)v^p\right)\frac{\partial}{\partial q^i}(q)\otimes dq^j(q)\left(G^{ab}(q)\frac{\partial V}{\partial q^b}(q)\frac{\partial}{\partial q^a}(q)\right)\\
&\quad+{P_D}^i_j(q)\frac{\partial}{\partial q^i}(q)\otimes dq^j(q)\left(\left(\frac{\partial(\mathbb{G}^\#\circ dV)^a}{\partial q^b}(q)v^b+\overset{\mathbb{G}}{\Gamma}{}^a_{bc}(q)v^b(\mathbb{G}^\#\circ dV)^c(q)\right)\frac{\partial}{\partial q^a}(q)\right)\\
&=\left(\frac{\partial {P_D}^i_j}{\partial q^k}v^k+{P_D}^q_j\overset{\mathbb{G}}{\Gamma}{}^i_{qk}v^k-{P_D}^i_k\overset{\mathbb{G}}{\Gamma}{}^k_{pj}v^p\right)\frac{\partial}{\partial q^i}\otimes dq^j\left(G^{ab}\frac{\partial V}{\partial q^b}\frac{\partial}{\partial q^a}\right)\\
&\quad+{P_D}^i_j\frac{\partial}{\partial q^i}\otimes dq^j\left(\left(\frac{\partial(\mathbb{G}^\#\circ dV)^a}{\partial q^b}v^b+\overset{\mathbb{G}}{\Gamma}{}^a_{bc}v^b(\mathbb{G}^\#\circ dV)^c\right)\frac{\partial}{\partial q^a}\right)\Bigg|_{v_q}\\
&=\left(\frac{\partial {P_D}^i_j}{\partial q^k}v^k+{P_D}^q_j\overset{\mathbb{G}}{\Gamma}{}^i_{qk}v^k-{P_D}^i_k\overset{\mathbb{G}}{\Gamma}{}^k_{pj}v^p\right)\left(G^{jb}\frac{\partial V}{\partial q^b}\right)\frac{\partial}{\partial q^i}\\
&\quad+{P_D}^i_j\left(\frac{\partial(\mathbb{G}^\#\circ dV)^j}{\partial q^b}v^b+\overset{\mathbb{G}}{\Gamma}{}^j_{bc}v^b(\mathbb{G}^\#\circ dV)^c\right)\frac{\partial}{\partial q^i}\Bigg|_{v_q}\\
&=\left(\frac{\partial {P_D}^i_j}{\partial q^k}v^kG^{jb}\frac{\partial V}{\partial q^b}+{P_D}^q_j\overset{\mathbb{G}}{\Gamma}{}^i_{qk}v^kG^{jb}\frac{\partial V}{\partial q^b}-{P_D}^i_k\overset{\mathbb{G}}{\Gamma}{}^k_{pj}v^pG^{jb}\frac{\partial V}{\partial q^b}\right)\frac{\partial}{\partial q^i}\\
&\quad+\left({P_D}^i_j\frac{\partial(G^{jt}\frac{\partial V}{\partial q^t})}{\partial q^b}v^b+{P_D}^i_j\overset{\mathbb{G}}{\Gamma}{}^j_{bc}v^bG^{ct}\frac{\partial V}{\partial q^t}\right)\frac{\partial}{\partial q^i}\Bigg|_{v_q}\\
&=\Bigg(\frac{\partial {P_D}^i_j}{\partial q^k}v^kG^{jb}\frac{\partial V}{\partial q^b}+{P_D}^q_j\overset{\mathbb{G}}{\Gamma}{}^i_{qk}v^kG^{jb}\frac{\partial V}{\partial q^b}\underbrace{-{P_D}^i_k\overset{\mathbb{G}}{\Gamma}{}^k_{pj}v^pG^{jb}\frac{\partial V}{\partial q^b}}_{\text{cancels with (*)}}+{P_D}^i_j\frac{\partial G^{jt}}{\partial q^b}\frac{\partial V}{\partial q^t}v^b+{P_D}^i_jG^{jt}\frac{\partial^2 V}{\partial q^b\partial q^t}v^b\\
&\quad+\underbrace{{P_D}^i_j\overset{\mathbb{G}}{\Gamma}{}^j_{bc}v^bG^{ct}\frac{\partial V}{\partial q^t}}_{(*)}\Bigg)\frac{\partial}{\partial q^i}\Bigg|_{v_q}\\
&=\left(\frac{\partial {P_D}^i_j}{\partial q^k}v^kG^{jb}\frac{\partial V}{\partial q^b}+\overset{\mathbb{G}}{\Gamma}{}^i_{qk}v^k\underbrace{{P_D}^q_jG^{jb}\frac{\partial V}{\partial q^b}}_{0}+{P_D}^i_j\frac{\partial G^{jt}}{\partial q^b}\frac{\partial V}{\partial q^t}v^b+{P_D}^i_jG^{jt}\frac{\partial^2 V}{\partial q^b\partial q^t}v^b\right)\frac{\partial}{\partial q^i}\Bigg|_{v_q}\\
&=\left(\frac{\partial {P_D}^i_j}{\partial q^k}(q)G^{jt}(q)\frac{\partial V}{\partial q^t}(q)+{P_D}^i_j(q)\frac{\partial G^{jt}}{\partial q^k}(q)\frac{\partial V}{\partial q^t}(q)+{P_D}^i_j(q)G^{jt}(q)\frac{\partial^2 V}{\partial q^k\partial q^t}(q)\right)v^k\frac{\partial}{\partial q^i}(q).
\end{aligned}
$$

□

Proposition 4.7 (Linearization of a Nonholonomic Mechanical System)**.** *Let $\Sigma = (Q,\mathbb{G},V,F,D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, and consider an equilibrium configuration $q_0 \in Q$ for Σ. Then the linearized equations of motion at 0_{q_0} in the decomposition $T_{0_{q_0}}TQ \cong T_{q_0}Q \oplus T_{q_0}Q$ are given* *by*

$$\begin{pmatrix}\dot q(t)\\ \dot v(t)\end{pmatrix} = \begin{pmatrix} 0 & id_{T_{q_0}Q} \\ -P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ gradV)(q_0) + P_D \circ \mathbb{G}^\# \circ d_1F(0_{q_0}) & P_D \circ \mathbb{G}^\# \circ d_2F(0_{q_0})\end{pmatrix}\begin{pmatrix} q(t)\\ v(t)\end{pmatrix}.$$

Proof. The proof is exactly the same as that of Proposition 4.4 except for the calculation of $d_1W(0_{q_0})$. Note that before, when we evaluated at 0_{q_0} in (4.2.1), we used the fact that q_0 was a critical point to obtain zeroes in two of the terms. This no longer applies here.

Let us compute

$$[d_1W(0_{q_0})] = \left[\frac{\partial W^i}{\partial q^j}(0_{q_0})\right] = \left[\frac{\partial}{\partial q^j}\left({P_D}^i_s G^{st}F_t - {P_D}^i_s G^{st}\frac{\partial V}{\partial q^t}\right)(0_{q_0})\right]$$

by again first computing the derivative, but expanding only the first term to get

$$\frac{\partial}{\partial q^j}\left({P_D}^i_s G^{st}F_t - {P_D}^i_s G^{st}\frac{\partial V}{\partial q^t}\right) = \frac{\partial {P_D}^i_s}{\partial q^j}G^{st}F_t + {P_D}^i_s\frac{\partial G^{st}}{\partial q^j}F_t + {P_D}^i_s G^{st}\frac{\partial F_t}{\partial q^j} - \frac{\partial}{\partial q^j}\left({P_D}^i_s G^{st}\frac{\partial V}{\partial q^t}\right).$$

Now, note that since P_D is idempotent, $P_D \circ P_D \circ \mathbb{G}^\# \circ dV = P_D \circ \mathbb{G}^\# \circ dV$, and so

$$\begin{aligned}
{P_D}^i_s G^{st}\frac{\partial V}{\partial q^t} &= {P_D}^i_s {P_D}^s_t G^{tr}\frac{\partial V}{\partial q^r}\\
\frac{\partial}{\partial q^j}\left({P_D}^i_s G^{st}\frac{\partial V}{\partial q^t}\right) &= \frac{\partial}{\partial q^j}\left({P_D}^i_s {P_D}^s_t G^{tr}\frac{\partial V}{\partial q^r}\right)\\
&= \frac{\partial {P_D}^i_s}{\partial q^j}{P_D}^s_t G^{tr}\frac{\partial V}{\partial q^r} + {P_D}^i_s\left(\frac{\partial {P_D}^s_t}{\partial q^j}G^{tr}\frac{\partial V}{\partial q^r} + {P_D}^s_t\frac{\partial G^{tr}}{\partial q^j}\frac{\partial V}{\partial q^r} + {P_D}^s_t G^{tr}\frac{\partial^2 V}{\partial q^j\partial q^r}\right).
\end{aligned}$$

Evaluating now at 0_{q_0}, we get

$$\begin{aligned}
&\frac{\partial}{\partial q^j}\left({P_D}^i_s G^{st}F_t - {P_D}^i_s G^{st}\frac{\partial V}{\partial q^t}\right)(0_{q_0})\\
&= \frac{\partial {P_D}^i_s}{\partial q^j}(q_0)G^{st}(q_0)\underbrace{F_t(0_{q_0})}_{0} + {P_D}^i_s(q_0)\frac{\partial G^{st}}{\partial q^j}(q_0)\underbrace{F_t(0_{q_0})}_{0}\\
&\quad + {P_D}^i_s(q_0)G^{st}(q_0)\frac{\partial F_t}{\partial q^j}(0_{q_0}) - \frac{\partial {P_D}^i_s}{\partial q^j}(q_0)\underbrace{{P_D}^s_t(q_0)G^{tr}(q_0)\frac{\partial V}{\partial q^r}(q_0)}_{0}\\
&\quad - {P_D}^i_s(q_0)\left(\frac{\partial {P_D}^s_t}{\partial q^j}(q_0)G^{tr}(q_0)\frac{\partial V}{\partial q^r}(q_0) + {P_D}^s_t(q_0)\frac{\partial G^{tr}}{\partial q^j}(q_0)\frac{\partial V}{\partial q^r}(q_0) + {P_D}^s_t(q_0)G^{tr}(q_0)\frac{\partial^2 V}{\partial q^j\partial q^r}(q_0)\right)\\
&= {P_D}^i_s(q_0)G^{st}(q_0)\frac{\partial F_t}{\partial q^j}(0_{q_0})\\
&\quad - {P_D}^i_s(q_0)\left(\frac{\partial {P_D}^s_t}{\partial q^j}(q_0)G^{tr}(q_0)\frac{\partial V}{\partial q^r}(q_0) + {P_D}^s_t(q_0)\frac{\partial G^{tr}}{\partial q^j}(q_0)\frac{\partial V}{\partial q^r}(q_0) + {P_D}^s_t(q_0)G^{tr}(q_0)\frac{\partial^2 V}{\partial q^j\partial q^r}(q_0)\right).
\end{aligned}$$

So, applying Lemma 4.6, we have that

$$d_1W(0_{q_0})(v_1) = P_D \circ \mathbb{G}^\# \circ d_1F(0_{q_0})(v_1) - P_D \circ \overset{\mathbb{G}}{\nabla}_{v_1}(P_D \circ \mathrm{grad}V)(q_0),$$

or

$$d_1W(0_{q_0})(v_1) = P_D \circ \mathbb{G}^\# \circ d_1F(0_{q_0})(v_1) - P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathrm{grad}V)(q_0)(v_1).$$

Hence the linearization is

$$\begin{pmatrix} \dot{q}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{id}_{T_{q_0}Q} \\ -P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathrm{grad}V)(q_0) + P_D \circ \mathbb{G}^\# \circ d_1 F(0_{q_0}) & P_D \circ \mathbb{G}^\# \circ d_2 F(0_{q_0}) \end{pmatrix} \begin{pmatrix} q(t) \\ v(t) \end{pmatrix}.$$

□

Note that when $dV(q_0) = 0$, we have $P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathrm{grad}V)(q_0) = P_D \circ \mathbb{G}^\# \circ \mathrm{Hess}V^\flat(q_0)$, so this result is consistent with Proposition 4.4.

4.3. Observation: Linearizing Before Solving for Lagrange Multipliers

Ever since researchers started studying the linearization of nonholonomic mechanical systems, there has been a debate regarding the correct approach. Whittaker's original procedure in [14] involved linearizing the constraints in addition to the equations of motion. In other words, using our differential geometric language, the equations in (3.2.1) are linearized before the Lagrange multipliers are solved for. As discussed in [12], he and other researchers argued that this approach was valid, and because linearizing the nonholonomic constraints gave rise to holonomic ones, the nonholonomicity of constraints played no role in stability analysis. We show here that this approach is in general not valid; it gives us a linearization with missing information.

Again, let us first consider the case where the equilibrium configuration is a critical point of the potential function. Actually, for this case, we will show that the approach is in fact valid.

4.3.1. Direct Linearization of a Nonholonomic Mechanical System at a Critical Point

Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, and consider an equilibrium configuration $q_0 \in Q$ for Σ such that $dV(q_0) = 0$. The equations of motion are

$$\begin{cases} \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -\underbrace{\mathrm{grad}V}_{\mathbb{G}^\# \circ dV}(\gamma(t)) + \mathbb{G}^\# \circ F(\gamma'(t)) + \lambda(t) \\ P_D^\perp(\gamma'(t)) = 0. \end{cases}$$

We want to linearize these equations about the equilibrium point 0_{q_0} and then use the linearized equations to solve for the linearized constraint force.

Consider a local chart $(U, \phi = (q^1, \dots, q^n))$ of Q around q_0, and write

1. $\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = \left(\ddot{q}^i(t) + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(\gamma(t))\dot{q}^j(t)\dot{q}^k(t)\right)\frac{\partial}{\partial q^i}(\gamma(t))$ where $\overset{\mathbb{G}}{\Gamma}{}^i_{jk} : U \to \mathbb{R}$ and $\phi \circ \gamma(t) = (q^1(t), \dots, q^n(t))$,
2. $dV(q) = \frac{\partial V}{\partial q^i}(q)dq^i(q)$ where $\frac{\partial V}{\partial q^i} : U \to \mathbb{R}$,
3. $F(v_q) = F_i(v_q)dq^i(q)$ where $F_i : TU \to \mathbb{R}$,
4. $\lambda(t) = \lambda^i(t)\frac{\partial}{\partial q^i}(\gamma(t))$ where $\lambda_i : \mathbb{R} \to \mathbb{R}$, and

5. $\mathbb{G}(q) = G_{ij}(q)dq^i(q) \otimes dq^j(q)$ where $G_{ij} : U \to \mathbb{R}$.
Hence we have

$$\ddot{q}^i(t) + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(\gamma(t))\dot{q}^j(t)\dot{q}^k(t) = -G^{ij}(\gamma(t))\frac{\partial V}{\partial q^j}(\gamma(t)) + G^{ij}(\gamma(t))F_j(\gamma'(t)) + \lambda^i(t).$$

Let $q(t) = (q^1(t), \ldots, q^n(t))$ and $\dot{q}(t) = (\dot{q}^1(t), \ldots, \dot{q}^n(t))$. The entire equation in coordinates is now

$$\ddot{q}^i(t) + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q(t))\dot{q}^j(t)\dot{q}^k(t) = -G^{ij}(q(t))\frac{\partial V}{\partial q^j}(q(t)) + G^{ij}(q(t))F_j(q(t), \dot{q}(t)) + \lambda^i(t).$$

Consider the linearization about the trajectory $(q(t) = \phi(q_0), \dot{q}(t) = 0, \ddot{q}(t) = 0, \lambda^i(t) = 0)$ using

$$\begin{cases} q(t) = \phi(q_0) + \delta q(t) \\ \dot{q}(t) = 0 + \delta\dot{q}(t) \\ \ddot{q}(t) = 0 + \delta\ddot{q}(t) \\ \lambda^i(t) = 0 + \delta\lambda^i(t). \end{cases}$$

Linearizing

$$\ddot{q}^i + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q)\dot{q}^j\dot{q}^k = -G^{ij}(q)\frac{\partial V}{\partial q^j}(q) + G^{ij}(q)F_j(q, \dot{q}) + \lambda^i,$$

we get

$$\begin{aligned} &\delta\ddot{q}^i + \underbrace{\frac{\partial(\overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q)\dot{q}^j\dot{q}^k)}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}}_{0}\delta q + \underbrace{\frac{\partial(\overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q)\dot{q}^j\dot{q}^k)}{\partial \dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}}_{0}\delta\dot{q} \\ &= -\frac{\partial(G^{ij}(q)\frac{\partial V}{\partial q^j}(q))}{\partial q}\Big|_{q=\phi(q_0)}\delta q + \frac{\partial(G^{ij}(q)F_j(q,\dot{q}))}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}\delta q + \frac{\partial(G^{ij}(q)F_j(q,\dot{q}))}{\partial \dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q} + \delta\lambda^i. \end{aligned}$$

Hence,

$$\begin{aligned} \delta\ddot{q}^i = &-\left(\frac{\partial G^{ij}(q)}{\partial q}\underbrace{\frac{\partial V}{\partial q^j}(q)\Big|_{q=\phi(q_0)}}_{0} + G^{ij}(q)\frac{\partial\frac{\partial V}{\partial q^j}(q)}{\partial q}\Big|_{q=\phi(q_0)}\right)\delta q \\ &+ \left(\frac{\partial G^{ij}(q)}{\partial q}\underbrace{F_j(q,\dot{q})\Big|_{q=\phi(q_0),\dot{q}=0}}_{0} + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}\right)\delta q + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial\dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q} + \delta\lambda^i \\ &= -G^{ij}(q)\frac{\partial\frac{\partial V}{\partial q^j}(q)}{\partial q}\Big|_{q=\phi(q_0)}\delta q + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}\delta q + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial\dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q} + \delta\lambda^i \\ &= -G^{ij}(q)\frac{\partial^2 V(q)}{\partial q^k\partial q^j}\Big|_{q=\phi(q_0)}\delta q^k + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial q^k}\Big|_{q=\phi(q_0),\dot{q}=0}\delta q^k + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial v^k}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q}^k + \delta\lambda^i \\ &= -G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k + \delta\lambda^i. \end{aligned}$$

Hence the linearized equation is

$$\delta\ddot{q}^i(t) = -G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t) + \delta\lambda^i(t).$$

In first order form, we have

$$\begin{pmatrix} \delta\dot{q}^1(t) \\ \vdots \\ \delta\dot{q}^n(t) \\ \delta\ddot{q}^1(t) \\ \vdots \\ \delta\ddot{q}^n(t) \end{pmatrix} = \begin{pmatrix} 0_{n\times n} & I_{n\times n} \\ \left[-G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\right] & \left[G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\right] \end{pmatrix} \begin{pmatrix} \delta q^1(t) \\ \vdots \\ \delta q^n(t) \\ \delta\dot{q}^1(t) \\ \vdots \\ \delta\dot{q}^n(t) \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta\lambda^1(t) \\ \vdots \\ \delta\lambda^n(t) \end{pmatrix}.$$

It seems like we could have almost pulled this result out from the standard linearization

$$A_\Sigma(q_0) = \begin{pmatrix} 0 & \mathrm{id}_{T_{q_0}Q} \\ -\mathbb{G}(q_0)^\# \circ \mathrm{Hess}V(q_0)^\flat + \mathbb{G}(q_0)^\# \circ d_1F(0_{q_0}) & \mathbb{G}(q_0)^\# \circ d_2F(0_{q_0}) \end{pmatrix}.$$

Now, we have to linearize the constraint equations

$$P_D^\perp(\gamma'(t)) = 0.$$

Let us write $P_D^\perp(q) = P_{D\,j}^{\perp\,i}(q)\frac{\partial}{\partial q^i}(q) \otimes dq^j(q)$ where $P_{D\,j}^{\perp\,i} : U \to \mathbb{R}$, and so $P_D^\perp(q)(v_q) = P_{D\,j}^{\perp\,i}(q)v_q^j\frac{\partial}{\partial q^i}(q)$ where $v_q = v_q^i\frac{\partial}{\partial q^i}(q)$. Hence we have

$$P_D^\perp(\gamma'(t)) = P_{D\,j}^{\perp\,i}(\gamma(t))\dot{q}^j(t)\frac{\partial}{\partial q^i}(\gamma(t)) = 0,$$

or

$$P_{D\,j}^{\perp\,i}(q(t))\dot{q}^j(t) = 0.$$

Consider the linearization about the same trajectory $(q(t) = \phi(q_0), \dot{q}(t) = 0, \ddot{q}(t) = 0, \lambda^i(t) = 0)$ using

$$\begin{cases} q(t) = \phi(q_0) + \delta q(t) \\ \dot{q}(t) = 0 + \delta\dot{q}(t) \\ \ddot{q}(t) = 0 + \delta\ddot{q}(t) \\ \lambda^i(t) = 0 + \delta\lambda^i(t). \end{cases}$$

Linearizing

$$P_{D\,j}^{\perp\,i}(q)\dot{q}^j = 0,$$

we get

$$\begin{aligned} \underbrace{\frac{\partial(P_{D\,j}^{\perp\,i}(q)\dot{q}^j)}{\partial q}\bigg|_{q=\phi(q_0),\dot{q}=0}}_{0}\delta q + \frac{\partial(P_{D\,j}^{\perp\,i}(q)\dot{q}^j)}{\partial \dot{q}}\bigg|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q} &= 0 \\ P_{D\,j}^{\perp\,i}(q)\Big|_{q=\phi(q_0)}\delta\dot{q}^j &= 0 \\ P_{D\,j}^{\perp\,i}(q_0)\delta\dot{q}^j &= 0. \end{aligned}$$

Hence the linearized equation is

$$P_{D\,j}^{\perp\,i}(q_0)\delta\dot{q}^j(t) = 0.$$

Now, we solve for the linearized constraint force. First, differentiate the linearized constraints to get

$$P^{\perp}_{D}{}^{i}_{j}(q_0)\delta\ddot{q}^j(t) = 0.$$

We had

$$\delta\ddot{q}^i(t) = -G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t) + \delta\lambda^i(t),$$

so

$$\begin{aligned}
\delta\lambda^i(t) &= \delta\ddot{q}^i(t) + G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) - G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\\
&= P^{\perp}_{D}{}^{i}_{s}(q_0)\left(\delta\ddot{q}^s(t) + G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right)\\
&= \underbrace{P^{\perp}_{D}{}^{i}_{s}(q_0)\delta\ddot{q}^s(t)}_{0} + P^{\perp}_{D}{}^{i}_{s}(q_0)\left(G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right)\\
&= P^{\perp}_{D}{}^{i}_{s}(q_0)\left(G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right).
\end{aligned}$$

Hence we have

$$\begin{aligned}
\delta\ddot{q}^i(t) &= -G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\\
&\quad + P^{\perp}_{D}{}^{i}_{s}(q_0)\left(G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right)\\
&= \delta^i_s\left(-G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right)\\
&\quad - P^{\perp}_{D}{}^{i}_{s}(q_0)\left(-G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right)\\
&= \Big(\underbrace{\delta^i_s - P^{\perp}_{D}{}^{i}_{s}(q_0)}_{P_D{}^i_s(q_0)}\Big)\left(-G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right)\\
&= -P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t).
\end{aligned}$$

In first order form, we have

$$\begin{pmatrix}\delta\dot{q}^1(t)\\ \vdots\\ \delta\dot{q}^n(t)\\ \delta\ddot{q}^1(t)\\ \vdots\\ \delta\ddot{q}^n(t)\end{pmatrix} = \begin{pmatrix} 0_{n\times n} & I_{n\times n}\\ \left[-P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0) + P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\right] & \left[P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\right]\end{pmatrix}\begin{pmatrix}\delta q^1(t)\\ \vdots\\ \delta q^n(t)\\ \delta\dot{q}^1(t)\\ \vdots\\ \delta\dot{q}^n(t)\end{pmatrix},$$

i.e.

$$\begin{pmatrix}\dot{q}(t)\\ \dot{v}(t)\end{pmatrix} = \begin{pmatrix}0 & \mathrm{id}_{T_{q_0}Q}\\ -P_D\circ\mathbb{G}^{\#}\circ\mathrm{Hess}V^{\flat}(q_0) + P_D\circ\mathbb{G}^{\#}\circ d_1F(0_{q_0}) & P_D\circ\mathbb{G}^{\#}\circ d_2F(0_{q_0})\end{pmatrix}\begin{pmatrix}q(t)\\ v(t)\end{pmatrix}.$$

This is exactly what we got using the linearization of a forced affine connection system when $dV(q_0) = 0$ (Proposition 4.4).

4.3.2. Direct Linearization of a Nonholonomic Mechanical System

Now, let us look at the general case. Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, and consider an equilibrium configuration $q_0 \in Q$ for Σ. The equations of motion are

$$\begin{cases} \overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -\underbrace{\mathrm{grad}V}_{\mathbb{G}^\# \circ dV}(\gamma(t)) + \mathbb{G}^\# \circ F(\gamma'(t)) + \lambda(t) \\ P_D^\perp(\gamma'(t)) = 0. \end{cases}$$

We want to linearize these equations about the equilibrium point 0_{q_0} and then use the linearized equations to solve for the linearized constraint force. We follow the same steps as before, except we must now linearize about the trajectory $(q(t) = \phi(q_0), \dot{q}(t) = 0, \ddot{q}(t) = 0, \lambda^i(t) = \lambda_0^i)$ using

$$\begin{cases} q(t) = \phi(q_0) + \delta q(t) \\ \dot{q}(t) = 0 + \delta\dot{q}(t) \\ \ddot{q}(t) = 0 + \delta\ddot{q}(t) \\ \lambda^i(t) = \lambda_0^i + \delta\lambda^i(t). \end{cases}$$

Linearizing

$$\ddot{q}^i + \overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q)\dot{q}^j\dot{q}^k = -G^{ij}(q)\frac{\partial V}{\partial q^j}(q) + G^{ij}(q)F_j(q,\dot{q}) + \lambda^i,$$

we get

$$\begin{aligned} &\delta\ddot{q}^i + \underbrace{\frac{\partial(\overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q)\dot{q}^j\dot{q}^k)}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}}_{0}\delta q + \underbrace{\frac{\partial(\overset{\mathbb{G}}{\Gamma}{}^i_{jk}(q)\dot{q}^j\dot{q}^k)}{\partial \dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}}_{0}\delta\dot{q} \\ &= -\frac{\partial(G^{ij}(q)\frac{\partial V}{\partial q^j}(q))}{\partial q}\Big|_{q=\phi(q_0)}\delta q + \frac{\partial(G^{ij}(q)F_j(q,\dot{q}))}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}\delta q + \frac{\partial(G^{ij}(q)F_j(q,\dot{q}))}{\partial \dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q} + \delta\lambda^i. \end{aligned}$$

Hence,

$$\begin{aligned} \delta\ddot{q}^i &= -\left(\frac{\partial G^{ij}(q)}{\partial q}\frac{\partial V}{\partial q^j}(q)\Big|_{q=\phi(q_0)} + G^{ij}(q)\frac{\partial\frac{\partial V}{\partial q^j}(q)}{\partial q}\Big|_{q=\phi(q_0)}\right)\delta q \\ &\quad + \left(\frac{\partial G^{ij}(q)}{\partial q}\underbrace{F_j(q,\dot{q})\Big|_{q=\phi(q_0),\dot{q}=0}}_{0} + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}\right)\delta q + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial\dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q} + \delta\lambda^i \\ &= -\frac{\partial G^{ij}(q)}{\partial q}\frac{\partial V}{\partial q^j}(q)\Big|_{q=\phi(q_0)}\delta q - G^{ij}(q)\frac{\partial\frac{\partial V}{\partial q^j}(q)}{\partial q}\Big|_{q=\phi(q_0)}\delta q + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial q}\Big|_{q=\phi(q_0),\dot{q}=0}\delta q \\ &\quad + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial\dot{q}}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q} + \delta\lambda^i \\ &= -\frac{\partial G^{ij}(q)}{\partial q^k}\frac{\partial V}{\partial q^j}(q)\Big|_{q=\phi(q_0)}\delta q^k - G^{ij}(q)\frac{\partial^2 V(q)}{\partial q^k\partial q^j}\Big|_{q=\phi(q_0)}\delta q^k + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial q^k}\Big|_{q=\phi(q_0),\dot{q}=0}\delta q^k \\ &\quad + G^{ij}(q)\frac{\partial F_j(q,\dot{q})}{\partial v^k}\Big|_{q=\phi(q_0),\dot{q}=0}\delta\dot{q}^k + \delta\lambda^i \\ &= -\frac{\partial G^{ij}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k - G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k + \delta\lambda^i. \end{aligned}$$

Hence the linearized equation is

$$\begin{aligned}\delta\ddot{q}^i(t) &= -\frac{\partial G^{ij}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) - G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t)\\ &\quad + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t) + \delta\lambda^i(t).\end{aligned}$$

In first order form, we have

$$\begin{pmatrix}\delta\dot{q}^1(t)\\ \vdots\\ \delta\dot{q}^n(t)\\ \delta\ddot{q}^1(t)\\ \vdots\\ \delta\ddot{q}^n(t)\end{pmatrix} = \begin{pmatrix} 0_{n\times n} & I_{n\times n}\\ \left[-\frac{\partial G^{ij}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0) - G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\right] & \left[G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\right]\end{pmatrix}\begin{pmatrix}\delta q^1(t)\\ \vdots\\ \delta q^n(t)\\ \delta\dot{q}^1(t)\\ \vdots\\ \delta\dot{q}^n(t)\end{pmatrix} + \begin{pmatrix}0\\ \vdots\\ 0\\ \delta\lambda^1(t)\\ \vdots\\ \delta\lambda^n(t)\end{pmatrix}.$$

Now, we have to linearize the constraint equations $P_D^{\perp}(\gamma'(t)) = 0$ about the trajectory $(q(t) = \phi(q_0), \dot{q}(t) = 0, \ddot{q}(t) = 0, \lambda^i(t) = \lambda^i_0)$. This gives us

$$P^{\perp}_D{}^i_j(q_0)\delta\dot{q}^j(t) = 0,$$

exactly as before.

Let us solve for the linearized constraint force. Differentiating, we get

$$P^{\perp}_D{}^i_j(q_0)\delta\ddot{q}^j(t) = 0.$$

We had

$$\begin{aligned}\delta\ddot{q}^i(t) &= -\frac{\partial G^{ij}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) - G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t)\\ &\quad + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t) + \delta\lambda^i(t),\end{aligned}$$

so

$$\begin{aligned}\delta\lambda^i(t) &= \delta\ddot{q}^i(t) + \frac{\partial G^{ij}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) + G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k(t) - G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t)\\ &\quad - G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\\ &= P^{\perp}_D{}^i_s(q_0)\bigg(\delta\ddot{q}^s(t) + \frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k(t)\\ &\quad - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\bigg)\\ &= \underbrace{P^{\perp}_D{}^i_s(q_0)\delta\ddot{q}^s(t)}_{0} + P^{\perp}_D{}^i_s(q_0)\bigg(\frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k(t)\\ &\quad - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\bigg)\\ &= P^{\perp}_D{}^i_s(q_0)\bigg(\frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k\partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t)\\ &\quad - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\bigg).\end{aligned}$$

Hence we have

$$
\begin{aligned}
\delta\ddot{q}^i(t) &= -\frac{\partial G^{ij}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) - G^{ij}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{ij}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t) \\
&\quad + P_D^{\perp}{}^i_s(q_0)\left(\frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) - G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right) \\
&= \delta^i_s\left(-\frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right) \\
&\quad - P_D^{\perp}{}^i_s(q_0)\left(-\frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\right) \\
&= \Big(\underbrace{\delta^i_s - P_D^{\perp}{}^i_s(q_0)}_{P_D{}^i_s(q_0)}\Big)\Big(-\frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) - G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) \\
&\quad + G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t)\Big) \\
&= -P_D{}^i_s(q_0)\frac{\partial G^{sj}}{\partial q^k}(q_0)\frac{\partial V}{\partial q^j}(q_0)\delta q^k(t) - P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial^2 V}{\partial q^k \partial q^j}(q_0)\delta q^k(t) \\
&\quad + P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial F_j}{\partial q^k}(0_{q_0})\delta q^k(t) + P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\delta\dot{q}^k(t).
\end{aligned}
$$

In first order form, we have

$$
\begin{pmatrix} \delta\dot{q}^1(t) \\ \vdots \\ \delta\dot{q}^n(t) \\ \delta\ddot{q}^1(t) \\ \vdots \\ \delta\ddot{q}^n(t) \end{pmatrix} = \begin{pmatrix} 0_{n\times n} & I_{n\times n} \\ \left[-P_D{}^i_s\frac{\partial G^{sj}}{\partial q^k}\frac{\partial V}{\partial q^j} - P_D{}^i_s G^{sj}\frac{\partial^2 V}{\partial q^k \partial q^j} + P_D{}^i_s G^{sj}\frac{\partial F_j}{\partial q^k}\right]\Big|_{0_{q_0}} & \left[P_D{}^i_s(q_0)G^{sj}(q_0)\frac{\partial F_j}{\partial v^k}(0_{q_0})\right] \end{pmatrix} \begin{pmatrix} \delta q^1(t) \\ \vdots \\ \delta q^n(t) \\ \delta\dot{q}^1(t) \\ \vdots \\ \delta\dot{q}^n(t) \end{pmatrix}.
$$

This is *not the same* as what we got using the linearization of a forced affine connection system.

We see here that the two approaches to linearization give us different answers. There is no ambiguity in the first approach since in that approach we solve for the constraint force first and then linearize the resulting equations of motion, which is exactly what we mean by linearization. The second approach is faulty; linearizing the system does not mean we can linearize the constraints. Now, it is interesting to ask: when do the two operations, of linearizing the equations and solving for Lagrange multipliers, commute? In other words, when does this second approach actually give us a valid linearization? Based on our calculations above, we have at least a sufficient condition. We know this holds at critical points of the potential function.

Chapter 5

Nonholonomic Mechanical Systems: Stability of Equilibria

Now, with our knowledge in nonholonomic mechanical systems and its linearizations, we try to explore the stability of its equilibria. Because the presence of constraints greatly complicate the nature of equilibria for these systems, we only scratch the surface of this topic here. The results in this chapter are by no means complete nor comprehensive; there remains much work to be done.

5.1. Stability via Linearization

We begin by looking at stability analysis via linearization. First, let us examine the structure of the linearization.

Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, and consider an equilibrium configuration $q_0 \in Q$ for Σ. The linearized equations of motion about the equilibrium point 0_{q_0} are

$$\begin{pmatrix} \dot{q}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{id}_{T_{q_0}Q} \\ -P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathrm{grad}V)(q_0) + P_D \circ \mathbb{G}^\# \circ d_1F(0_{q_0}) & P_D \circ \mathbb{G}^\# \circ d_2F(0_{q_0}) \end{pmatrix} \begin{pmatrix} q(t) \\ v(t) \end{pmatrix},$$

or

$$\begin{pmatrix} \dot{q}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{id}_{T_{q_0}Q} \\ -P_D(q_0) \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathrm{grad}V)(q_0) + P_D(q_0) \circ \mathbb{G}(q_0)^\# \circ d_1F(0_{q_0}) & P_D(q_0) \circ \mathbb{G}(q_0)^\# \circ d_2F(0_{q_0}) \end{pmatrix} \begin{pmatrix} q(t) \\ v(t) \end{pmatrix}.$$

Let $V = T_{q_0}Q$, $D = D(q_0)$, $M = \mathbb{G}(q_0)$, $P_D = P_D(q_0)$, $L = \overset{\mathbb{G}}{\nabla}(P_D \circ \mathrm{grad}V)(q_0)$, $F_1 = d_1F(0_{q_0})$, and $F_2 = d_2F(0_{q_0})$. Hence, we want to study the linear map

$$A = \begin{pmatrix} 0 & \mathrm{id}_V \\ -P_DL + P_DM^\#F_1 & P_DM^\#F_2 \end{pmatrix} : V \oplus V \to V \oplus V, \tag{5.1.1}$$

where V is a finite-dimensional $\mathbb{R}$-vector space, D is a subspace of V, M is an inner product on V, $P_D : V \to V$ is the orthogonal projection of V onto D, and L, F_1 and F_2 are linear maps from V to V.

Proposition 5.1 (Invariant Subspace of the State Space)**.** *Consider the linear map A from (5.1.1). We have that $D \oplus D$ is A-invariant, i.e. $A(D \oplus D) \subseteq D \oplus D$.*

Proof. Note that D is a subspace of V, so $D \oplus D$ is a subspace of $V \oplus V$. Given $(q, v) \in D \oplus D$, we have

$$\begin{aligned} A(q,v) &= \begin{pmatrix} 0 & \mathrm{id}_V \\ -P_D L + P_D M^{\#} F_1 & P_D M^{\#} F_2 \end{pmatrix} \begin{pmatrix} q \\ v \end{pmatrix} \\ &= \begin{pmatrix} v \\ -P_D L q + P_D M^{\#} F_1 q + P_D M^{\#} F_2 v \end{pmatrix} \in D \oplus D. \end{aligned}$$

□

Now, consider the decomposition $V = D \oplus D^{\perp}$. Let us write the linear map A from (5.1.1) as $A : D \oplus D^{\perp} \oplus D \oplus D^{\perp} \to D \oplus D^{\perp} \oplus D \oplus D^{\perp}$, where

$$\begin{aligned} A &= \begin{pmatrix} 0 & 0 & \mathrm{id}_D & 0 \\ 0 & 0 & 0 & \mathrm{id}_{D^{\perp}} \\ \pi_D(-P_D L + P_D M^{\#} F_1) i_D & \pi_D(-P_D L + P_D M^{\#} F_1) i_{D^{\perp}} & \pi_D(P_D M^{\#} F_2) i_D & \pi_D(P_D M^{\#} F_2) i_{D^{\perp}} \\ \pi_{D^{\perp}}(-P_D L + P_D M^{\#} F_1) i_D & \pi_{D^{\perp}}(-P_D L + P_D M^{\#} F_1) i_{D^{\perp}} & \pi_{D^{\perp}}(P_D M^{\#} F_2) i_D & \pi_{D^{\perp}}(P_D M^{\#} F_2) i_{D^{\perp}} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & \mathrm{id}_D & 0 \\ 0 & 0 & 0 & \mathrm{id}_{D^{\perp}} \\ -\pi_D P_D L i_D + \pi_D P_D M^{\#} F_1 i_D & -\pi_D P_D L i_{D^{\perp}} + \pi_D P_D M^{\#} F_1 i_{D^{\perp}} & \pi_D P_D M^{\#} F_2 i_D & \pi_D P_D M^{\#} F_2 i_{D^{\perp}} \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & \mathrm{id}_D & 0 \\ 0 & 0 & 0 & \mathrm{id}_{D^{\perp}} \\ -\pi_D L i_D + \pi_D M^{\#} F_1 i_D & -\pi_D L i_{D^{\perp}} + \pi_D M^{\#} F_1 i_{D^{\perp}} & \pi_D M^{\#} F_2 i_D & \pi_D M^{\#} F_2 i_{D^{\perp}} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

where $\pi_D, \pi_{D^{\perp}}, i_D, i_{D^{\perp}}$ are the obvious projection and inclusion maps. We define the **linearization restricted to the invariant subspace**, $A_D : D \oplus D \to D \oplus D$, as

$$A_D = \begin{pmatrix} 0 & \mathrm{id}_D \\ -\pi_D L i_D + \pi_D M^{\#} F_1 i_D & \pi_D M^{\#} F_2 i_D \end{pmatrix}.$$

Proposition 5.2 (Eigenvalues of the Linearization)**.** *The eigenvalues of A_D form a subset of the eigenvalues of A. All other eigenvalues of A are zero.*

Proof. Instead of writing A as a map from $D \oplus D^{\perp} \oplus D \oplus D^{\perp}$ to $D \oplus D^{\perp} \oplus D \oplus D^{\perp}$, let us write it as a map from $D \oplus D \oplus D^{\perp} \oplus D^{\perp}$ to $D \oplus D \oplus D^{\perp} \oplus D^{\perp}$, i.e.

$$A = \begin{pmatrix} 0 & \mathrm{id}_D & 0 & 0 \\ -\pi_D L i_D + \pi_D M^{\#} F_1 i_D & \pi_D M^{\#} F_2 i_D & -\pi_D L i_{D^{\perp}} + \pi_D M^{\#} F_1 i_{D^{\perp}} & \pi_D M^{\#} F_2 i_{D^{\perp}} \\ 0 & 0 & 0 & \mathrm{id}_{D^{\perp}} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We see from here that $D \oplus D$ is again A-invariant, and

$$\sigma(A) = \sigma(A_D) \cup \sigma \begin{pmatrix} 0 & \mathrm{id}_{D^{\perp}} \\ 0 & 0 \end{pmatrix},$$

so we are done.

We can also prove this directly. Suppose $(q_1, q_2, v_1, v_2) \in D \oplus D^\perp \oplus D \oplus D^\perp$ is an eigenvector not in $D \oplus D$. Hence $q_2 \neq 0$ or $v_2 \neq 0$, and there exists $\lambda \in \mathbb{C}$ such that $A(q_1, q_2, v_1, v_2) = \lambda(q_1, q_2, v_1, v_2)$. We have

$$\begin{cases} \lambda q_1 = v_1 \\ \lambda q_2 = v_2 \\ \lambda v_1 = Wq_1 + Xq_2 + Yv_1 + Zv_2 \\ \lambda v_2 = 0, \end{cases}$$

where $W = -\pi_D L i_D + \pi_D M^{\#} F_1 i_D$, $X = \pi_D M^{\#} F_2 i_D$, $Y = -\pi_D L i_{D^\perp} + \pi_D M^{\#} F_1 i_{D^\perp}$ and $Z = \pi_D M^{\#} F_2 i_{D^\perp}$. Now, if $v_2 \neq 0$, then $\lambda = 0$ implies that $v_1 = v_2 = 0$, which is a contradiction. So it must be that $q_2 \neq 0$. If $q_2 \neq 0$, then from $v_2 = \lambda q_2$ and $\lambda v_2 = 0$ we have $\lambda^2 q_2 = 0$, which means $\lambda = 0$. □

Proposition 5.3 (Eigenvectors of the Linearization)**.** *Let* $(q, v) \in V \oplus V$ *and* $\lambda \in \mathbb{C}$. *Then,* (q, v) *is an eigenvector of* A *with* λ *as the associated eigenvalue if and only if* $q \neq 0$ *and* $q \in Ker(\lambda^2 id_V + P_D L - P_D M^{\#} F_1 - \lambda P_D M^{\#} F_2)$.

Proof. Let $(q, v) \in V \oplus V$ be an eigenvector of A and let λ be the associated eigenvalue. Then,

$$\begin{pmatrix} 0 & \mathrm{id}_V \\ -P_D L + P_D M^{\#} F_1 & P_D M^{\#} F_2 \end{pmatrix} \begin{pmatrix} q \\ v \end{pmatrix} = \lambda \begin{pmatrix} q \\ v \end{pmatrix},$$

or

$$\begin{cases} v = \lambda q \\ -P_D L q + P_D M^{\#} F_1 q + P_D M^{\#} F_2 v = \lambda v. \end{cases}$$

Now, if $q = 0$, then $v = 0$ and so (q, v) cannot be an eigenvector. So it must be that $q \neq 0$. We have

$$\begin{aligned} -P_D L q + P_D M^{\#} F_1 q + P_D M^{\#} F_2 v &= \lambda v \\ -P_D L q + P_D M^{\#} F_1 q + \lambda P_D M^{\#} F_2 q &= \lambda^2 q \\ \lambda^2 q + P_D L q - P_D M^{\#} F_1 q - \lambda P_D M^{\#} F_2 q &= 0 \\ (\lambda^2 \mathrm{id}_V + P_D L - P_D M^{\#} F_1 - \lambda P_D M^{\#} F_2) q &= 0 \end{aligned}$$

so $q \in \mathrm{Ker}(\lambda^2 \mathrm{id}_V + P_D L - P_D M^{\#} F_1 - \lambda P_D M^{\#} F_2)$.

The converse is also clear from the above. □

From Proposition 5.2, we immediately see that the eigenvalues of A cannot be restricted to $\mathbb{C}_-$ since we always have eigenvalues at zero. Hence, we can never use Proposition A.8 to conclude local asymptotic stability of an equilibrium configuration for nonholonomic mechanical systems. This already reveals a severe limitation to the linearization approach.

Remark 5.4 (Trajectories of the Linearization). What do trajectories of the linearization look like? For an initial condition $(q_0, v_0) \in D \oplus D$, the trajectory remains inside $D \oplus D$. For an initial condition $(q_0, v_0) \in V \oplus V$ but not in $D \oplus D$:

1. If $v_0 \notin D$, then v_0 will have a component in $D^\perp$ which remains constant. Hence $q(t)$ never stops increasing along that direction in $D^\perp$ and so the trajectory is unstable.

2. If $q_0 \notin D$, then q_0 has a component in $D^{\perp}$. If $v_0 \in D$, then this component remains contant and so the trajectory will not go to zero. If $v_0 \notin D$, then like the first case, the trajectory is unstable.

Thus the linearized system as a whole is not stable.

It is interesting to ask whether it is enough to study the map A_D, and be able conclude local asymptotic stability from this since the trajectories that do appear to have physical meaning are those in $D \oplus D$. This is not true, however. Because the constraints only present restrictions on the velocity, we should really be looking at $V \oplus D$, not $D \oplus D$. The following is an example of a nonholonomic mechanical system in which the linearization restricted to the invariant subspace has eigenvalues with strictly negative real part, but the equilibrium configuration is not locally asymptotically stable.

Example 5.5. Consider $\Sigma = (Q, \mathbb{G}, V, F, D)$, a C^{∞}-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$, where

(i) $Q = \mathbb{R}^3$ with global coordinate charts $(\mathbb{R}^3, \phi = (x,y,z))$ and $(T\mathbb{R}^3, T\phi = (x,y,z,u,v,w))$,
(ii) $\mathbb{G} = dx \otimes dx + dy \otimes dy + dz \otimes dz$,
(iii) $V(q) = x^2 + y^2$,
(iv) $F(v_q) = -udx(q) - vdy(q) - wdz(q)$, and
(v) $D = \operatorname{span}\{X_1, X_2\}$, where $X_1 = \frac{\partial}{\partial x}$ and $X_2 = \frac{\partial}{\partial y} + (x+z)\frac{\partial}{\partial z}$.

We compute $[X_1, X_2] = \frac{\partial}{\partial z}$. Note that X_1 and X_2 belong to D, but $[X_1, X_2]$ does not. Hence D is not involutive and so is not integrable. Furthermore, $\{X_1, X_2, [X_1, X_2]\}$ is linearly independent for all $q \in Q$, so D is in fact totally nonholonomic.

Consider the equilibrium configuration $q_0 = (0,0,0)$. We calculate

$$[\mathrm{Hess}V(q_0)] = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad [\mathbb{G}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$[d_1F(0_{q_0})] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad [d_2F(0_{q_0})] = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad [P_D(q_0)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The linearization A is then

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -2 & 0 & 0 & -1 & 0 & 0 \\ 0 & -2 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and the restricted linearization A_D is

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 0 & -1 & 0 \\ 0 & -2 & 0 & -1 \end{pmatrix},$$

which has eigenvalues $\sigma(A_D) = \left\{-\frac{1}{2} - \frac{\sqrt{7}}{2}i, -\frac{1}{2} - \frac{\sqrt{7}}{2}i, -\frac{1}{2} + \frac{\sqrt{7}}{2}i, -\frac{1}{2} + \frac{\sqrt{7}}{2}i\right\}$. All eigenvalues of the map A_D have strictly negative real part. However, since the equilibrium configuration is not isolated (all points on the z-axis are equilibria), we cannot have local asymptotic stability.

We cannot conclude local asymptotic stability of an equilibrium configuration, but we can conclude, to a certain extent, instability. Consider the following definition.

Definition 5.6 (Instability of the Equilibrium Set)**.** Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$. Let $Q_0 \subseteq Q$ be the set of equilibrium configurations. An equilibrium configuration $q_0 \in Q$ is Q_0**-unstable** if there exists a neighbourhood $U \subseteq TQ$ of $Z(TQ_0)$ such that for every neighbourhood $V \subseteq TQ$ of 0_{q_0}, there is a point $v_q \in V$ such that the integral curve $\Phi_t^X(v_q) \notin U$ for some $t \geq 0$.

The following proposition is due to [6].

Proposition 5.7 (Instability via Linearization Analysis)**.** *Let $\Sigma = (Q, \mathbb{G}, V, F, D)$ be a C^∞-forced simple mechanical system with regular constraints such that $F(Z(TQ)) = Z(T^*Q)$. Let $q_0 \in Q$ be an equilibrium configuration for Σ and let $Q_0 \subseteq Q$ be the set of equilibrium configurations. Let $A_\Sigma(q_0)$ be the linearization of Σ at 0_{q_0}. If $\sigma(A_\Sigma(q_0)) \cap \mathbb{C}_+ \neq \emptyset$, then q_0 is Q_0-unstable.*

Proof. We require the following fact (see Theorem 6.1, Corollary 6.1 and the remark to Corollary 6.1 in [7]). Since the linearization $A_\Sigma(q_0)$ has an eigenvalue with positive real part, for some $\epsilon > 0$ there exists a solution $\gamma : (-\infty, 0] \to Q$ to (3.2.2) such that $\gamma'(t) \neq 0_{q_0}$ for all $t \leq 0$, and $\lim_{t\to-\infty} \|T\phi \circ \gamma'(t) - T\phi(0_{q_0})\| e^{-\epsilon t} = 0$ where we have chosen some coordinates. In other words, we can find a solution on the unstable manifold which converges to 0_{q_0} as we move back in time.

Note that we also have $\gamma'(t) \notin Z(TQ_0)$ for all $t \leq 0$. Let $U \subseteq TQ$ be a neighbourhood of $Z(TQ_0)$ such that $\gamma'(0) \notin U$. Now, for every neighbourhood $V \subseteq TQ$ of 0_{q_0}, there exists a $t < 0$ such that $\gamma'(t) \in V$ but $\gamma'(0) \notin U$. $\square$

Corollary 5.8 (Instability via Linearization Analysis in the Absence of Dissipation)**.** *In the absence of dissipation, if $P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ gradV)(q_0)$ has an eigenvalue with negative real part, then q_0 is Q_0-unstable.*

Proof. We have the linearization

$$A_\Sigma(q_0) = \begin{pmatrix} 0 & \mathrm{id}_{T_{q_0}Q} \\ -P_D \circ \overset{\mathbb{G}}{\nabla}(P_D \circ \mathrm{grad}V)(q_0) & 0 \end{pmatrix}.$$

The characteristic equation is

$$\begin{aligned}
&\det\left(s\begin{pmatrix}\mathrm{id}_{T_{q_0}Q} & 0\\ 0 & \mathrm{id}_{T_{q_0}Q}\end{pmatrix} - \begin{pmatrix}0 & \mathrm{id}_{T_{q_0}Q}\\ -P_D\circ\overset{\mathbb{G}}{\nabla}(P_D\circ\mathrm{grad}V)(q_0) & 0\end{pmatrix}\right)\\
&=\det\begin{pmatrix}s\cdot\mathrm{id}_{T_{q_0}Q} & -\mathrm{id}_{T_{q_0}Q}\\ P_D\circ\overset{\mathbb{G}}{\nabla}(P_D\circ\mathrm{grad}V)(q_0) & s\cdot\mathrm{id}_{T_{q_0}Q}\end{pmatrix}\\
&=\det\begin{pmatrix}0 & -\mathrm{id}_{T_{q_0}Q}\\ s^2\cdot\mathrm{id}_{T_{q_0}Q}+P_D\circ\overset{\mathbb{G}}{\nabla}(P_D\circ\mathrm{grad}V)(q_0) & s\cdot\mathrm{id}_{T_{q_0}Q}\end{pmatrix}\\
&=\det(s^2\cdot\mathrm{id}_{T_{q_0}Q}-(-P_D\circ\overset{\mathbb{G}}{\nabla}(P_D\circ\mathrm{grad}V)(q_0)))=0.
\end{aligned}$$

Hence the eigenvalues of $A_\Sigma(q_0)$ are the square roots of the eigenvalues of $-P_D\circ\overset{\mathbb{G}}{\nabla}(P_D\circ\mathrm{grad}V)(q_0)$. If $P_D\circ\overset{\mathbb{G}}{\nabla}(P_D\circ\mathrm{grad}V)(q_0)$ has an eigenvalue with negative real part, then $A_\Sigma(q_0)$ has an eigenvalue with positive real part and so q_0 is Q_0-unstable. $\square$

5.2. Stability via Lyapunov Methods

Now, we look at stability analysis via Lyapunov methods. Though our approach is elementary, interesting observations and questions arise. First, we prove that the total energy of a constrained mechanical system changes directly with the external force. The constraint forces have no overall effect on the total energy. Intuitively, this comes from the fact that constraint forces do no work on curves satisfying the equations of motion (3.2.2).

Proposition 5.9 (Time Derivative of Energy for Constrained Mechanical Systems)**.** *Let $(Q,\mathbb{G},V,F,D)$ be a C^∞-forced simple mechanical system with regular constraints, and consider a curve $\gamma: I\to Q$ satisfying the equations of motion (3.2.2). Then,*

$$\frac{dE(\gamma'(t))}{dt}=\langle F(\gamma'(t));\gamma'(t)\rangle.$$

Proof. We know that the total energy is

$$E(v_q)=\frac{1}{2}\mathbb{G}(q)(v_q,v_q)+V(q),$$

hence

$$E(\gamma'(t))=\frac{1}{2}\mathbb{G}(\gamma(t))(\gamma'(t),\gamma'(t))+V(\gamma(t)),$$

and so taking the time derivative we get

$$
\begin{aligned}
&\frac{dE(\gamma'(t))}{dt} \\
&= \frac{d}{dt}\left(\frac{1}{2}\mathbb{G}(\gamma(t))(\gamma'(t),\gamma'(t)) + V(\gamma(t))\right) \\
&= \frac{1}{2}\underbrace{(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\mathbb{G})}_{0}(\gamma(t))(\gamma'(t),\gamma'(t)) + \frac{1}{2}\mathbb{G}(\gamma(t))(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t),\gamma'(t)) \\
&\quad + \frac{1}{2}\mathbb{G}(\gamma(t))(\gamma'(t),\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)) + (\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}V)(\gamma(t)) \\
&= \mathbb{G}(\gamma(t))(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle \\
&= \mathbb{G}(\gamma(t))(P_D\circ\mathbb{G}^{\#}\circ F(\gamma'(t)) - P_D\circ\mathbb{G}^{\#}\circ dV(\gamma(t)) - (\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}P_D^{\perp})(\gamma'(t)),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle \\
&= \mathbb{G}(\gamma(t))(P_D\circ\mathbb{G}^{\#}\circ F(\gamma'(t)),\gamma'(t)) - \mathbb{G}(\gamma(t))(P_D\circ\mathbb{G}^{\#}\circ dV(\gamma(t)),\gamma'(t)) \\
&\quad - \mathbb{G}(\gamma(t))((\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}P_D^{\perp})(\gamma'(t)),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle \\
&= \mathbb{G}(\gamma(t))(\mathbb{G}^{\#}\circ F(\gamma'(t)),P_D^T(\gamma'(t))) - \mathbb{G}(\gamma(t))(\mathbb{G}^{\#}\circ dV(\gamma(t)),P_D^T(\gamma'(t))) \\
&\quad - \mathbb{G}(\gamma(t))((\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}P_D^{\perp})(\gamma'(t)),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle \\
&= \langle\mathbb{G}^{\flat}\circ\mathbb{G}^{\#}\circ F(\gamma'(t));P_D^T(\gamma'(t))\rangle - \langle\mathbb{G}^{\flat}\circ\mathbb{G}^{\#}\circ dV(\gamma(t));P_D^T(\gamma'(t))\rangle \\
&\quad - \mathbb{G}(\gamma(t))((\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}P_D^{\perp})(\gamma'(t)),\gamma'(t)) + \langle dV(\gamma(t));\gamma'(t)\rangle \\
&= \langle F(\gamma'(t));P_D^T(\gamma'(t))\rangle - \langle dV(\gamma(t));P_D^T(\gamma'(t))\rangle + \underbrace{\mathbb{G}(\gamma(t))(P_D^{\perp}(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)),\gamma'(t))}_{0} + \langle dV(\gamma(t));\gamma'(t)\rangle \\
&= \langle F(\gamma'(t));P_D^T(\gamma'(t))\rangle + \langle dV(\gamma(t));\gamma'(t) - P_D^T(\gamma'(t))\rangle \\
&= \langle F(\gamma'(t));P_D^T(\gamma'(t))\rangle + \langle dV(\gamma(t));(P_D^T)^{\perp}(\gamma'(t))\rangle \\
&= \langle F(\gamma'(t));P_D(\gamma'(t))\rangle + \langle dV(\gamma(t));\underbrace{P_D^{\perp}(\gamma'(t))}_{0}\rangle \\
&= \langle F(\gamma'(t));\gamma'(t)\rangle.
\end{aligned}
$$

□

Let us look at how Theorem 2.23 should be adapted, if possible. First of all, in order to apply the Lyapunov stability criteria, we need a function which is locally positive-definite about the equilibrium configuration. The obvious choice is the potential function. For dissipative nonholonomic mechanical systems, we know that equilibria can arise at configurations which are not critical points of the potential function. This already presents a huge limitation; in order to easily apply Lyapunov methods, we need to restrict ourselves to equilibria which are critical points. We can prove the following theorem, which raises some interesting questions.

Theorem 5.10. *Let $\Sigma = (Q, \mathbb{G}, V, F_{diss}, D)$ be a C^{∞}-simple mechanical system with dissipation subject to nonholonomic constraints, and consider $q_0 \in Q$ such that $dV(q_0) = 0$.*

(i) If V is locally positive-definite about q_0, then q_0 is stable.

(ii) If q_0 is an isolated local minimum for V, F_{diss} is strictly dissipative on D, and there exists a neighbourhood $U \subseteq Q$ of q_0 such that $dV(q) \notin ann(D)_q$ for all $q \in U \setminus \{q_0\}$, then q_0 is locally asymptotically stable.

Note: there exists a neighbourhood $U \subseteq Q$ of q_0 such that $dV(q) \notin \mathrm{ann}(D)_q$ for all $q \in U \setminus \{q_0\}$ if and only if $P_D(\mathbb{G}^\# \circ dV(q))$ is locally non-zero around q_0 except at q_0.

Proof. (i) This part of the proof is analogous to that of Theorem 2.23, however we only need to consider the manifold D, not the entire TQ. Hence, consider the energy function $E : D \to \mathbb{R}$, $E(v_q) = \frac{1}{2}\mathbb{G}(q)(v_q, v_q) + V(q)$. Without loss of generality, assume $V(q_0) = 0$.

Claim: E is a Lyapunov function for the associated vector field X for Σ at the point 0_{q_0}.

Proof of claim:

1. We have $E(0_{q_0}) = 0$.
2. $E(\gamma'(t))$ is non-increasing by Proposition 5.9. In other words, $\mathcal{L}_X E$ is negative-semidefinite about 0_{q_0}.
3. We need to show that there exists a neighbourhood $TU \cap D \subseteq D$ of 0_{q_0} such that for all $v_q \in TU \cap D \setminus \{0_{q_0}\}$, we have $E(v_q) > 0$. Since V is locally positive-definite about q_0, there exists a neighbourhood $U \subseteq Q$ of q_0 such that for all $q \in U \setminus \{q_0\}$, we have $V(q) > 0$. Now, for all $v_q \in TU \cap D \setminus Z(TU \cap D)$, we have $\mathbb{G}(q)(v_q, v_q) > 0$. Hence, for all $v_q \in TU \cap D \setminus \{0_{q_0}\}$, we have $E(v_q) > 0$. Hence E is a Lyapunov function for the associated vector field at the point 0_{q_0} as claimed.

Hence, by the Lyapunov stability criteria (Theorem A.16), q_0 is stable.

(ii) Since q_0 is an isolated local minimum for V, V is locally positive-definite about q_0 and so (i) applies. E is a Lyapunov function. Also, since q_0 is an isolated local minimum for V and there exists a neighbourhood $U \subseteq Q$ of q_0 such that $dV(q) \notin \mathrm{ann}(D)_q$ for all $q \in U \setminus \{q_0\}$, then there exists a neighbourhood $W \subseteq Q$ of q_0 such that for all $q \in W \setminus \{q_0\}$, we have $V(q) > V(q_0)$ and $P_D(\mathbb{G}^\# \circ dV(q)) \neq 0$. We have that E is positive-definite on $TW \cap D$ and $E(\gamma'(t))$ is non-increasing. Let $A = \{v_q \in TW \cap D \mid \langle F_{\mathrm{diss}}(v_q), v_q\rangle = 0\} = Z(TW \cap D)$.

Claim: $\{0_{q_0}\}$ is the only positively invariant set in A.

Proof of claim: Suppose $0_{q_1} \neq 0_{q_0}$ is another point. Now, unlike Theorem 2.23, having q_0 an isolated local minimum is not enough. We now require $dV(q) \notin \mathrm{ann}(D)_q$ for all $q \in U \setminus \{q_0\}$, which means q_0 is an isolated equilibrium configuration. Hence the requirement of q_0 being an isolated local minimum in Theorem 2.23 should actually be a requirement on the equilibria, i.e. q_0 should be an isolated equilibrium point. Since we assumed $dV(q) \notin \mathrm{ann}(D)_q$ for all $q \in U \setminus \{q_0\}$, $q_1 \neq q_0 \Rightarrow P_D(\mathbb{G}^\# \circ dV(q_1)) \neq 0$. Hence the solution will leave A, proving the claim.

Hence the corollary to LaSalle Invariance Principle (Corollary A.22) implies q_0 is locally asymptotically stable.

□

Remark 5.11 (Questions).

1. We know that isolated equilibria for nonholonomic mechanical systems do exist, but do isolated equilibria exist such that they are a local minimum for the potential function?
2. In fact, do locally asymptotically stable equilibria even exist for nonholonomic mechanical systems?

Chapter 6

Conclusions and Future Work

6.1. Conclusions

In this report, we studied the linearization of nonholonomic mechanical systems about equilibria and looked into the stability of its equilibria via both linearization and Lyapunov methods. Although there are still many unanswered questions, from this investigation we can summarize the following:

1. While the contraints tend to introduce a manifold of equilibria, equilibria can still be isolated and the set of equilibria may not form a submanifold of the configuration space.
2. The proper way to linearize the equations of motion of a nonholonomic mechanical system about an equilibrium point is to first solve for the constraint force before linearizing. Linearizing the constraints first before solving for the (linearized) constraint force in general will not give the same result. These two processes, however, do commute when linearizing about an equilibrium point which is also a critical point of the potential function.
3. Due to the presence of eigenvalues at zero, the linearization of a nonholonomic mechanical system can never be used to conclude local asymptotic stability, although it can be used to conclude a type of instability.
4. Lyapunov methods using energy tend to fail for the main reason that the potential function may no longer be positive-definite around the equilibrium.

6.2. Future Work

There remains much work to be done. Let us outline some interesting questions:

1. Can we find other conditions such that the two operations of linearization and solving for Lagrange multipliers commute when linearizing nonholonomic mechanical systems?
2. Is it possible to conclude stability (but not necessarily local asymptotic stability) by ignoring the eigenvalues at zero which arise due to the constraint and instead just studying the part of the linearization restricted to the invariant subspace?
3. Can we find other conditions for when the equilibria form a submanifold, or when they are isolated?

4. How can we adapt the Lyapunov methods to work in a more straightforward way? For example, if we assume the equilibria form a submanifold, it is possible to consider stability with respect to not just a point, but the entire equilibrium manifold. This is studied in [8], for example.

Appendix A

Nonlinear Systems and Stability Analysis

This appendix is more or less a quick summary of nonlinear systems and stability analysis as found in [3].

Definition A.1 (Nonlinear System). Let M be a C^∞-manifold and $X \in \Gamma^\infty(TM)$. The **nonlinear system** associated with X is the system of first order ODEs specified by $\gamma'(t) \triangleq T_t\gamma(\frac{\partial}{\partial t}(t)) = X(\gamma(t))$, where $\gamma : I \to M$. It is often convenient to simply refer to X as the nonlinear system, with the understanding that what we actually mean is the underlying differential equation.

Definition A.2 (Equilibrium Point). Let M be a C^∞-manifold and $X \in \Gamma^\infty(TM)$. A point $x_0 \in M$ is an **equilibrium point for** X if the trivial curve $\gamma : \mathbb{R} \to M$, $\gamma(t) = x_0$ is an integral curve for X.

Proposition A.3 (Characterization of Equilibrium Points). *Let M be a C^∞-manifold and $X \in \Gamma^\infty(TM)$. A point $x_0 \in M$ is an equilibrium point for X if and only if $X(x_0) = 0$.*

Definition A.4 (Convergence of a Curve). Let M be a C^∞-manifold, $S \subseteq M$, and consider $\gamma : [t_0, \infty) \to M$.

(i) γ **approaches** S **as** $t \to \infty$ if, for all neighbourhoods $U \subseteq M$ of S, there exists $T \geq t_0$ such that for all $t > T$, we have $\gamma(t) \in U$. This is denoted by $\gamma(t) \to S$ as $t \to \infty$.

(ii) For $x_0 \in M$ and $S = \{x_0\}$, γ **converges to** x_0 **as** $t \to \infty$ if γ approaches S as $t \to \infty$. This is denoted by $\lim_{t\to\infty} \gamma(t) = x_0$.

Definition A.5 (Stability Notions). Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$, and consider $x_0 \in M$, an equilibrium point for X.

(i) x_0 is **stable** if, for all neighbourhoods $U \subseteq M$ of x_0, there exists a neighbourhood $W \subseteq U$ of x_0 such that for all $x \in W$, the integral curve $t \mapsto \Phi_t^X(x)$ takes values in U for $t \geq 0$.

(ii) x_0 is **unstable** if x_0 is not stable.

(iii) x_0 is **locally asymptotically stable** if x_0 is stable and there exists a neighbourhood $U \subseteq M$ of x_0 such that for all $x \in U$, the integral curve $t \mapsto \Phi_t^X(x)$ converges to x_0 as $t \to \infty$.

(iv) x_0 is **globally asymptotically stable** if x_0 is stable and for all $x \in M$, the integral curve $t \mapsto \Phi_t^X(x)$ converges to x_0 as $t \to \infty$.

Definition A.6 (Linear System)**.** Let V be a finite-dimensional $\mathbb{R}$-vector space and $A \in L(V;V)$, where $L(V;V)$ is the set of linear maps from V to V. The **linear system** associated with A is the system of first order linear ODEs specified by $\dot{x}(t) = A(x(t))$, where $x : I \to V$. Note that $0 \in V$ is always an equilibrium point.

We denote the set of eigenvalues by σ, the algebraic multiplicity of an eigenvalue λ by $m_a(\lambda)$, and the geometric multiplicity of an eigenvalue λ by $m_g(\lambda)$.

Theorem A.7 (Stability of Linear Systems)**.** *Let V be a finite-dimensional $\mathbb{R}$-vector space and $A \in L(V;V)$.*

(i) If $\sigma(A) \cap \mathbb{C}_+ \neq \emptyset$, then 0 is unstable.
(ii) $\sigma(A) \subseteq \mathbb{C}_-$ if and only if 0 is globally asympotically stable.
(iii) If $\sigma(A) \subseteq \overline{\mathbb{C}}_-$ and for all $\lambda \in \sigma(A) \cap i\mathbb{R}$ we have $m_a(\lambda) = m_g(\lambda)$, then 0 is stable.
(iv) If $\sigma(A) \subseteq \overline{\mathbb{C}}_-$ and there exists $\lambda \in \sigma(A) \cap i\mathbb{R}$ such that $m_a(\lambda) > m_g(\lambda)$, then 0 is unstable.

Proposition A.8 (Stability from Linear Stability)**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$, and $x_0 \in M$, an equilibrium point for X. Consider $A_X(x_0) \in L(T_{x_0}M; T_{x_0}M)$, the linearization of X at x_0 (see Proposition C.8).*

(i) If $\sigma(A_X(x_0)) \subseteq \mathbb{C}_-$, then x_0 is locally asymptotically stable.
(ii) If $\sigma(A_X(x_0)) \cap \mathbb{C}_+ \neq \emptyset$, then x_0 is unstable.

Definition A.9 (More Stability Notions)**.** Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$ and $x_0 \in M$, an equilibrium point for X. Consider $A_X(x_0) \in L(T_{x_0}M; T_{x_0}M)$, the linearization of X at x_0.

(i) x_0 is **linearly stable** if 0 is stable for the linearization $A_X(x_0)$.
(ii) x_0 is **linearly asymptotically stable** if 0 is asymptotically stable for the linearization $A_X(x_0)$.
(iii) x_0 is **linearly unstable** if x_0 is not linearly stable.
(iv) x_0 is **spectrally stable** if $\sigma(A_X(x_0)) \subseteq \overline{\mathbb{C}}_-$.
(v) x_0 is **spectrally unstable** if x_0 is not spectrally stable.

Note A.10 (Some Immediate Implications)**.** 1. linear asymptotic stability $\Rightarrow$ local asymptotic stability
2. spectral instability $\Rightarrow$ instability

Proposition A.11 (Time Derivative of a Function Evaluated Along an Integral Curve)**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$ and $\psi \in C^\infty(M)$. Then,*

$$\frac{d}{dt}\psi(\Phi_t^X(x)) = \mathcal{L}_X\psi(\Phi_t^X(x)).$$

Proof. Consider $(U, \phi = (x^1, \ldots, x^n))$, a local chart of M, and let $X\Big|_U = X^i \frac{\partial}{\partial x^i}$. Let

$$\begin{array}{rccc} \gamma : & I & \to & M \\ & t & \mapsto & \Phi_t^X(x), \end{array}$$

and $\phi \circ \gamma(t) = (\gamma^1(t), \dots, \gamma^n(t))$. Then,

$$\begin{aligned}
\mathcal{L}_X\psi(\Phi_t^X(x)) &= \mathcal{L}_X\psi(\gamma(t)) \\
&= d\psi(\gamma(t))(X(\gamma(t))) \\
&= \frac{\partial\psi}{\partial x^j}(\gamma(t))dx^j(\gamma(t))\left(\dot{\gamma}^i(t)\frac{\partial}{\partial x^i}(\gamma(t))\right) \\
&= \frac{\partial\psi}{\partial x^i}(\gamma(t))\dot{\gamma}^i(t) \\
&= \frac{d}{dt}(\psi \circ \phi^{-1} \circ \phi \circ \gamma(t)) \\
&= \frac{d}{dt}(\psi \circ \gamma(t)) \\
&= \frac{d}{dt}\psi(\Phi_t^X(x)).
\end{aligned}$$

□

Definition A.12 (Locally Positive-definite Function)**.** Let M be a C^∞-manifold, $\psi \in C^\infty(M)$ and $x_0 \in M$.

(i) ψ is **locally positive-definite about** x_0 if $\psi(x_0) = 0$ and there exists a neighbourhood $U \subseteq M$ of x_0 such that for all $x \in U \setminus \{x_0\}$, we have $\psi(x) > 0$.

(ii) ψ is **locally positive-semidefinite about** x_0 if $\psi(x_0) = 0$ and there exists a neighbourhood $U \subseteq M$ of x_0 such that for all $x \in U$, we have $\psi(x) \geq 0$.

Definition A.13 (Sublevel Set)**.** Let M be a C^∞-manifold, $\psi \in C^\infty(M)$, $x_0 \in M$ and $L \in \mathbb{R}$.

(i) Define the **L-sublevel set of** ψ by $\psi^{-1}(\leq L) \overset{\Delta}{=} \psi^{-1}((-\infty, L]) = \{x \in M \mid \psi(x) \leq L\}$.

(ii) Define $\psi^{-1}(\leq L, x_0)$ as the connected component of $\psi^{-1}(\leq L)$ containing x_0 (if $\psi(x_0) \leq L$), or $\emptyset$ (if $\psi(x_0) > L$).

Lemma A.14 (Positive-definiteness and Existence of a Compact Sublevel Set)**.** *Let M be a C^∞-manifold, $\psi \in C^\infty(M)$ and $x_0 \in M$. If ψ is locally positive-definite about x_0, then for all neighbourhoods $U \subseteq M$ of x_0, there exists $\alpha > 0$ such that $\psi^{-1}(\leq \alpha, x_0) \subseteq U$ is compact.*

Definition A.15 (Lyapunov Function)**.** Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$, $x_0 \in M$ and $\psi \in C^\infty(M)$. ψ is a **Lyapunov function for** X **about** x_0 if ψ is locally positive-definite about x_0 and $\mathcal{L}_X\psi$ is locally negative-semidefinite about x_0.

Theorem A.16 (Lyapunov Stability Criteria)**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$ and $x_0 \in M$ an equilibrium point for X.*

(i) If there exists a Lyapunov function ψ for X about x_0, then x_0 is stable.

(ii) If there exists a Lyapunov function ψ for X about x_0 and $\mathcal{L}_X\psi$ is locally negative-definite about x_0, then x_0 is locally asymptotically stable.

Definition A.17 (Invariant Set)**.** Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$ and $A \subseteq M$.

(i) A is **X-invariant** if, for all $x \in A$, the integral curve $t \mapsto \Phi_t^X(x)$ takes values in A for all admissible $t \in \mathbb{R}$.

(ii) A is **positively X-invariant** if, for all $x \in A$, the integral curve $t \mapsto \Phi_t^X(x)$ takes values in A for all admissible $t \in \mathbb{R}_{\geq 0}$.

Lemma A.18 (X-invariance and Upper Limit of Definition of an Integral Curve)**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$ and $A \subseteq M$. If A is compact and positively X-invariant, then for all $x \in A$, $\sigma_+(X, x) = \infty$.*

Definition A.19 (Positive Limit Set)**.** Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$ and $x \in M$. The **positive limit set of x for X** is

$$\Omega(X, x) = \{y \in M \mid \exists \{t_k\}_{k \in \mathbb{N}} \subseteq \mathbb{R} : t_k < t_{k+1} \forall k \in \mathbb{N}, \lim_{k \to \infty} t_k = \infty, \lim_{k \to \infty} \Phi_{t_k}^X(x) = y\}.$$

Proposition A.20 (X-invariance of Positive Limit Sets)**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$ and $A \subseteq M$ compact and positively X-invariant. If $x \in A$, then $\Omega(X, x) \subseteq A$ is nonempty, compact, and positively X-invariant. Furthermore, $t \mapsto \Phi_t^X(x)$ approaches $\Omega(X, x)$ as $t \to \infty$.*

Theorem A.21 (LaSalle Invariance Principle)**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$, $A \subseteq M$ compact and positively X-invariant, and $\psi \in C^\infty(M)$ such that for all $x \in A$, we have $\mathcal{L}_X\psi(x) \leq 0$. Let B be the largest positively X-invariant set contained in $\{x \in A \mid \mathcal{L}_X\psi(x) = 0\}$.*

(i) For all $x \in A$, the integral curve $t \mapsto \Phi_t^X(x)$ approaches B as $t \to \infty$.

(ii) If B consists of a finite number of isolated points, then for all $x \in A$, the integral curve $t \mapsto \Phi_t^X(x)$ converges to a point in B as $t \to \infty$.

Corollary A.22 (Corollary of LaSalle Invariance Principle)**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$, $x_0 \in M$ an equilibrium point for X, and $\psi \in C^\infty(M)$ a Lyapunov function for X about x_0. Let $U \subseteq M$, a neighbourhood of x_0 on which ψ is positive-definite and $\mathcal{L}_X\psi$ is negative-semidefinite. Let $C = \{x \in U \mid \mathcal{L}_X\psi(x) = 0\}$. If $\{x_0\}$ is the only positively X-invariant set in C, then x_0 is locally asymptotically stable.*

Appendix B

Affine Connections and Covariant Derivatives

Definition B.1 (Affine Connection)**.** Let M be a C^∞-manifold. A **C^∞-affine connection** on M is an object $\nabla : \Gamma^\infty(TM) \times \Gamma^\infty(TM) \to \Gamma^\infty(TM)$, where notationally we write $\nabla_X Y$ instead of $\nabla(X, Y)$, such that

(i) ∇ is $\mathbb{R}$-bilinear,

(ii) for all $X, Y \in \Gamma^\infty(TM)$ and $f \in C^\infty(M)$, we have $\nabla_{fX} Y = f \nabla_X Y$, and

(iii) for all $X, Y \in \Gamma^\infty(TM)$ and $f \in C^\infty(M)$, we have $\nabla_X fY = (\mathcal{L}_X f)Y + f\nabla_X Y$.

$\nabla_X Y$ is called the **covariant derivative of Y with respect to X.**

Remark B.2 (An Affine Connection Introduces Additional Structure)*.* Unlike the Lie derivative, property (2) of Definition B.1 provides additional structure to the manifold.

Note B.3 (Covariant Derivative in Coordinates)**.** Let ∇ be a C^∞-affine connection on a C^∞-manifold M, and let $X, Y \in \Gamma^\infty(TM)$. Consider a local chart $(U, \phi = (x^1, \ldots, x^n))$ of M. Write $X\Big|_U = X^i \frac{\partial}{\partial x^i}$ and $Y\Big|_U = Y^i \frac{\partial}{\partial x^i}$. Then,

$$\begin{aligned}
\nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial x^i}} Y^j \frac{\partial}{\partial x^j} \\
&= X^i \nabla_{\frac{\partial}{\partial x^i}} Y^j \frac{\partial}{\partial x^j} \\
&= X^i \left(\left(\mathcal{L}_{\frac{\partial}{\partial x^i}} Y^j \right) \frac{\partial}{\partial x^j} + Y^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right) \\
&= \frac{\partial Y^i}{\partial x^j} X^j \frac{\partial}{\partial x^i} + \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} X^j Y^k.
\end{aligned}$$

Note that $\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k}$ is a vector field, hence let us write $\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} = \Gamma^i_{jk} \frac{\partial}{\partial x^i}$. Now, we have the coordinate formula

$$\nabla_X Y = \left(\frac{\partial Y^i}{\partial x^j} X^j + \Gamma^i_{jk} X^j Y^k \right) \frac{\partial}{\partial x^i}.$$

The functions $\Gamma^i_{jk} : U \to \mathbb{R}$ are called the **Christoffel symbols for ∇ in the chart** (U, ϕ). Note that if we are given Christoffel symbols for an affine connection, then we know the structure of the affine connection, at least locally.

Note B.4 (Covariant Derivative of a Vector Field with respect to a Tangent Vector)**.** At a point $x \in M$, we have $\nabla_X Y(x) = \left(\frac{\partial Y^i}{\partial x^j}(x)X^j(x) + \Gamma^i_{jk}(x)X^j(x)Y^k(x)\right)\frac{\partial}{\partial x^i}(x)$. Observe that the value of $\nabla_X Y(x)$ depends on X only on its value at x, i.e. $X(x)$. Hence, we define the **covariant derivative of** Y **with respect to** $v_x \in TM$ such that in coordinates $(U, \phi = (x^1, \ldots, x^n))$ around x we have

$$\nabla_{v_x} Y(x) = \left(\frac{\partial Y^i}{\partial x^j}(x)v_x^j + \Gamma^i_{jk}(x)v_x^j Y^k(x)\right)\frac{\partial}{\partial x^i}(x).$$

where $v_x = v_x^i \frac{\partial}{\partial x^i}(x)$. It is sometimes convenient to simply write $\nabla_{v_x} Y$ instead $\nabla_{v_x} Y(x)$, with the understanding that $\nabla_{v_x} Y$ is actually just a tangent vector at x, not a vector field.

Note B.5 (Covariant Derivative along a Curve)**.** Let ∇ be a C^∞-affine connection on a C^∞-manifold M, let $Y \in \Gamma^\infty(TM)$, and consider a C^∞-curve $\gamma : I \to M$. Given $t \in I$, we know that $\gamma'(t) \stackrel{\Delta}{=} T_t\gamma(\frac{\partial}{\partial t}(t)) \in T_{\gamma(t)}M$. So in suitable coordinates $(U, \phi = (x^1, \ldots, x^n))$, we have

$$\nabla_{\gamma'(t)} Y(\gamma(t)) = \left(\frac{\partial Y^i}{\partial x^j}(\gamma(t))\dot\gamma^j(t) + \Gamma^i_{jk}(\gamma(t))\dot\gamma^j(t)Y^k(\gamma(t))\right)\frac{\partial}{\partial x^i}(\gamma(t)),$$

where $\phi \circ \gamma(t) = (\gamma^1(t), \ldots, \gamma^n(t))$ and $\gamma'(t) = \dot\gamma^i(t)\frac{\partial}{\partial x^i}(\gamma(t))$. Notice that if we let $\overline{Y}(t) = Y(\gamma(t))$, we have

$$\nabla_{\gamma'(t)} \overline{Y}(t) = \left(\frac{d\overline{Y}^i}{dt}(t) + \Gamma^i_{jk}(\gamma(t))\dot\gamma^j(t)\overline{Y}^k(t)\right)\frac{\partial}{\partial x^i}(\gamma(t)).$$

Now, given a **vector field** $\overline{Y}$ **along a curve** γ, i.e. $\overline{Y}(t) \in T_{\gamma(t)}M$ for all t, we can define the **covariant derivative of a vector field along a curve** such that in coordinates we have the expression above. In particular, γ' is itself a vector field along the curve γ. In this case, we get

$$\nabla_{\gamma'(t)} \gamma'(t) = \left(\ddot\gamma^i(t) + \Gamma^i_{jk}(\gamma(t))\dot\gamma^j(t)\dot\gamma^k(t)\right)\frac{\partial}{\partial x^i}(\gamma(t)),$$

or in compact form,

$$\nabla_{\gamma'(t)} \gamma'(t) = \left(\ddot\gamma^i + \Gamma^i_{jk}\dot\gamma^j\dot\gamma^k\right)\frac{\partial}{\partial x^i}(\gamma(t)).$$

Remark B.6 (Acceleration). Essentially, $\nabla_{\gamma'(t)}\gamma'(t)$ gives us a coordinate-invariant way to express the familiar yet geometrically complicated notion of acceleration.

In Appendix C, we will cover additional details regarding $\nabla_{\gamma'(t)}\gamma'(t)$. Here, let us move on to a more general topic: covariant derivatives of tensor fields.

Note B.7 (Covariant Derivatives of Tensor Fields)**.** Consider a C^∞-manifold M. Let us look at how to define a differentiation operator on (r, s)-tensor fields, which we denote by D^r_s.

1. $(0,0)$-tensor fields: The simplest case is $D_0^0 : C^\infty(M) \to C^\infty(M)$. Note that elements in $C^\infty(M)$ are $(0,0)$-tensor fields.
   Given $X \in \Gamma^\infty(TM)$, D_0^0 can be defined by the Lie derivative
   $$D_0^0 f \triangleq \mathcal{L}_X f,$$
   or by the covariant derivative
   $$D_0^0 f \triangleq \nabla_X f \triangleq \mathcal{L}_X f.$$
   In this case, they are the same.
2. $(1,0)$-tensor fields: For $D_0^1 : \Gamma^\infty(TM) \to \Gamma^\infty(TM)$, we have the Lie bracket
   $$D_0^1 Y \triangleq [X, Y],$$
   or the typical covariant derivative
   $$D_0^1 Y \triangleq \nabla_X Y.$$
   Since we are interested in covariant derivatives, we shall use $D_0^0 f \triangleq \nabla_X f$ and $D_0^1 Y \triangleq \nabla_X Y$.
3. $(0,1)$-tensor fields: Now, we want to find the covariant derivative of $(0,1)$-tensor fields,
   $$D_1^0 : \Gamma^\infty(T^*M) \to \Gamma^\infty(T^*M).$$
   To do this, we insist that the product rule be satisfied, i.e. given $\alpha \in \Gamma^\infty(T^*M)$ and $Y \in \Gamma^\infty(TM)$, we have $\alpha(Y) \in C^\infty(M)$, so
   $$D_0^0(\alpha(Y)) = (D_1^0 \alpha)(Y) + \alpha(D_0^1 Y),$$
   or
   $$\nabla_X(\alpha(Y)) = (\nabla_X \alpha)(Y) + \alpha(\nabla_X Y),$$
   which means
   $$(\nabla_X \alpha)(Y) = \mathcal{L}_X(\alpha(Y)) - \alpha(\nabla_X Y).$$
   This shows how the covariant derivative of a $(0,1)$-tensor field acts on vector fields.
   Let $(U, \phi = (x^1, \ldots, x^n))$ be a local chart of M. We can write this out in coordinates as follows:
   $$\nabla_X \alpha = (\nabla_X \alpha)_i dx^i,$$
   where

$$\begin{aligned}
(\nabla_X\alpha)_i &= (\nabla_X\alpha)\left(\frac{\partial}{\partial x^i}\right)\\
&= \mathcal{L}_X\left(\alpha\left(\frac{\partial}{\partial x^i}\right)\right) - \alpha\left(\nabla_X\frac{\partial}{\partial x^i}\right)\\
&= \mathcal{L}_X\left(\alpha_j dx^j\left(\frac{\partial}{\partial x^i}\right)\right) - \alpha_j dx^j\left(\nabla_X\frac{\partial}{\partial x^i}\right)\\
&= \mathcal{L}_X(\alpha_i) - \alpha_p dx^p\left(\Gamma^k_{ji}X^j\frac{\partial}{\partial x^k}\right)\\
&= \frac{\partial\alpha_i}{\partial x^j}dx^j\left(X^k\frac{\partial}{\partial x^k}\right) - \alpha_k\Gamma^k_{ji}X^j\\
&= \frac{\partial\alpha_i}{\partial x^j}X^j - \alpha_k\Gamma^k_{ij}X^j.
\end{aligned}$$

It can be shown that $\nabla_X\alpha$ is indeed a $(0,1)$-tensor field.

4. $(1,1)$-tensor fields: Given $t \in \Gamma^\infty(T^1_1(TM))$, $\alpha \in \Gamma^\infty(T^*M)$ and $Y \in \Gamma^\infty(TM)$, we have $t(\alpha, Y) \in C^\infty(M)$. Hence, applying product rule,

$$\nabla_X(t(\alpha,Y)) = (\nabla_X t)(\alpha,Y) + t(\nabla_X\alpha, Y) + t(\alpha, \nabla_X Y),$$

i.e.

$$(\nabla_X t)(\alpha,Y) = \nabla_X(t(\alpha,Y)) - t(\nabla_X\alpha, Y) - t(\alpha, \nabla_X Y).$$

Let $(U, \phi = (x^1,\dots,x^n))$ be a local chart of M, and write

$$t = t^i_j\frac{\partial}{\partial x^i}\otimes dx^j,$$

and

$$\nabla_X t = (\nabla_X t)^i_j\frac{\partial}{\partial x^i}\otimes dx^j.$$

Hence,

$$\begin{aligned}
(\nabla_X t)^i_j &= (\nabla_X t)\left(dx^i, \frac{\partial}{\partial x^j}\right)\\
&= \nabla_X\left(t\left(dx^i,\frac{\partial}{\partial x^j}\right)\right) - t\left(\nabla_X dx^i, \frac{\partial}{\partial x^j}\right) - t\left(dx^i, \nabla_X\frac{\partial}{\partial x^j}\right)\\
&= \mathcal{L}_X(t^i_j) - t\left(-\Gamma^i_{qk}X^k dx^q, \frac{\partial}{\partial x^j}\right) - t\left(dx^i, \Gamma^k_{pj}X^p\frac{\partial}{\partial x^k}\right)\\
&= \frac{\partial t^i_j}{\partial x^k}X^k - t^r_s\frac{\partial}{\partial x^r}\left(-\Gamma^i_{qk}X^k dx^q\right)dx^s\left(\frac{\partial}{\partial x^j}\right) - t^r_s\frac{\partial}{\partial x^r}(dx^i)dx^s\left(\Gamma^k_{pj}X^p\frac{\partial}{\partial x^k}\right)\\
&= \frac{\partial t^i_j}{\partial x^k}X^k + t^q_j\Gamma^i_{qk}X^k - t^i_k\Gamma^k_{pj}X^p.
\end{aligned}$$

5. (r,s)-tensor fields: For the general case, given $t \in \Gamma^\infty(T^r_s(TM))$, $\alpha^1,\dots,\alpha^r \in \Gamma^\infty(T^*M)$, and $Y_1,\dots,Y_s \in \Gamma^\infty(TM)$, we have

$$\begin{aligned}(\nabla_X t)(\alpha^1,\dots,\alpha^r,Y_1,\dots,Y_s) = &\mathcal{L}_X(t(\alpha^1,\dots,\alpha^r,Y_1,\dots,Y_s))\\ &-\sum_{i=1}^r t(\alpha^1,\dots,\nabla_X\alpha^i,\dots,\alpha^r,Y_1,\dots,Y_s)\\ &-\sum_{j=1}^s t(\alpha^1,\dots,\alpha^r,Y_1,\dots,\nabla_X Y_j,\dots,Y_s).\end{aligned}$$

In coordinates $(U,\phi=(x^1,\dots,x^n))$, we have

$$(\nabla_X t)^{i_1\dots i_r}_{j_1\dots j_s} = \frac{\partial t^{i_1\dots i_r}_{j_1\dots j_s}}{\partial x^k}X^k + \sum_{\rho=1}^r \Gamma^{i_\rho}_{ki} t^{i_1\dots i\dots i_r}_{j_1\dots j_s} X^k - \sum_{\sigma=1}^s \Gamma^j_{kj_\sigma} t^{i_1\dots i_r}_{j_1\dots j\dots j_s} X^k.$$

Note B.8 (Covariant Differential)**.** In general, $\nabla_X t(x)$ depends on X only on its value at x, i.e. $X(x)$. We call such objects **tensorial** in X. Hence we can define the covariant derivative at a point using a tangent vector v_x, and write $\nabla t(v_x)$ instead of $\nabla_{v_x} t$.

Appendix C

The Geodesic Spray and Lifts of Vector Fields to the Tangent Bundle

Definition C.1 (Geodesic)**.** Let ∇ be an C^∞-affine connection on a C^∞-manifold M. A curve $\gamma : I \to M$ is a **geodesic** if $\nabla_{\gamma'(t)}\gamma'(t) = 0$.

Definition C.2 (Geodesic Spray)**.** Let ∇ be an C^∞-affine connection on a C^∞-manifold M. The **geodesic spray of** ∇ is a vector field on TM, $S : TM \to TTM$, defined such that the integral curves of S projected onto M (using the canonical projection) are geodesics of ∇. It can be shown that such a vector field is well-defined.

Note C.3 (Geodesic Spray in Coordinates)**.** Let ∇ be an C^∞-affine connection on a C^∞-manifold M. Consider a local chart $(U, \phi = (x^1, \ldots, x^n))$ of M and corresponding local chart $(TU, T\phi = (x^1, \ldots, x^n, v^1, \ldots, v^n))$ of TM. The geodesic spray in coordinates is

$$S\Big|_{TU}(v_x) = v^i \frac{\partial}{\partial x^i}(v_x) - \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i}(v_x),$$

where Γ^i_{jk} are the Christoffel symbols for ∇ in the chart (U, ϕ).

Now, let us look at how vector fields can be lifted to be defined on the tangent bundle. It will be helpful to introduce the following construction. Consider a C^∞-manifold M. For a local chart (U, ϕ) around $x \in M$, we define the map $\theta_{x,(U,\phi)} : T_xM \to \mathbb{R}^n$ such that $\theta_{x,(U,\phi)}([\gamma]_x) = \frac{d}{dt}\Big|_{t=0} \phi \circ \gamma(t)$.

Definition C.4 (Vertical Lift)**.** Let M be a C^∞-manifold and consider the tangent bundle TM. Given $v_x \in TM$, define the **vertical lift by** v_x as

$$\begin{array}{rccl} \mathrm{vlft}_{v_x} : & T_xM & \to & T_{v_x}TM \\ & X_x & \mapsto & [t \mapsto v_x + tX_x]_{v_x}. \end{array}$$

Now, given $X \in \Gamma^\infty(TM)$, define the **vertical lift of** X, $\mathrm{vlft}(X) \in \Gamma^\infty(TTM)$, by

$$\begin{array}{rccl} \mathrm{vlft}(X) : & TM & \to & TTM \\ & v_x & \mapsto & \mathrm{vlft}_{v_x}(X(x)). \end{array}$$

Note C.5 (Vertical Lift in Coordinates)**.** Consider a local chart $(U, \phi = (x^1, \ldots, x^n))$ of M and the corresponding local chart $(TU, T\phi = (x^1, \ldots, x^n, v^1, \ldots, v^n))$ of TM. Let $X\Big|_U(x) = X^i(x)\frac{\partial}{\partial x^i}(x)$. Then,

$$\begin{aligned}
TT\phi \circ \mathrm{vlft}(X)\Big|_{TU}(v_x) &= (T\phi(v_x), \frac{d}{dt}\Big|_{t=0} T\phi(v_x + tX(x)))) \\
&= (T\phi(v_x), \frac{d}{dt}\Big|_{t=0} (\phi(x), \theta_{x,(U,\phi)}(v_x + tX(x)))) \\
&= (T\phi(v_x), \frac{d}{dt}\Big|_{t=0} (\phi(x), \theta_{x,(U,\phi)}(v_x) + t\theta_{x,(U,\phi)}(X(x)))) \\
&= (T\phi(v_x), 0, \ldots, 0, \theta_{x,(U,\phi)}(X(x)))) \\
&= (T\phi(v_x), 0, \ldots, 0, X^1(x), \ldots, X^n(x))),
\end{aligned}$$

and so

$$\mathrm{vlft}(X)\Big|_{TU}(v_x) = X^i(x)\frac{\partial}{\partial v^i}(v_x).$$

Note that by the same calculations, we also have

$$\mathrm{vlft}_{v_x}\Big|_{TU}(X_x) = X^i_x \frac{\partial}{\partial v^i}(v_x),$$

where $X_x = X^i_x \frac{\partial}{\partial x^i}(x)$.

Definition C.6 (Tangent Lift)**.** Given $X \in \Gamma^\infty(TM)$, recall that the flow along X by t is a mapping $\Phi^X_t : M \to M$. Define the **tangent lift of** X, $X^T \in \Gamma^\infty(TTM)$, by

$$\begin{array}{rccl}
X^T: & TM & \to & TTM \\
& v_x & \mapsto & [t \mapsto T_x\Phi^X_t(v_x)]_{v_x}.
\end{array}$$

Note C.7 (Tangent Lift in Coordinates)**.** Consider a local chart $(U, \phi = (x^1, \ldots, x^n))$ of M and the corresponding local chart $(TU, T\phi = (x^1, \ldots, x^n, v^1, \ldots, v^n))$ of TM. Let $X\Big|_U(x) = X^i(x)\frac{\partial}{\partial x^i}(x)$. Then,

$$\begin{aligned}
TT\phi \circ X^T\Big|_{TU}(v_x) &= (T\phi(v_x), \frac{d}{dt}\Big|_{t=0} (T\phi \circ T_x\Phi^X_t(v_x))) \\
&= (T\phi(v_x), \frac{d}{dt}\Big|_{t=0} (\phi \circ \Phi^X_t(x), \theta_{\Phi^X_t(x),(U,\phi)} \circ T_x\Phi^X_t(v_x))).
\end{aligned}$$

Let $\phi \circ \Phi^X_t(x) = (\gamma^1(t), \ldots, \gamma^n(t))$. Then,

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=0} \phi \circ \Phi^X_t(x) &= (\dot\gamma^1(0), \ldots, \dot\gamma^n(0)) \\
&= (X^1 \circ \phi^{-1}(\gamma^1(0), \ldots, \gamma^n(0)), \ldots, X^n \circ \phi^{-1}(\gamma^1(0), \ldots, \gamma^n(0))) \\
&= (X^1(x), \ldots, X^n(x)).
\end{aligned}$$

Let $v_x = [\tau \mapsto \lambda(\tau)]_x$ and $\phi \circ \lambda(\tau) = (\lambda^1(\tau), \dots, \lambda^n(\tau))$. Then,

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=0} \theta_{\Phi_t^X(x),(U,\phi)} \circ T_x\Phi_t^X(v_x) &= \frac{d}{dt}\Big|_{t=0} \theta_{\Phi_t^X(x),(U,\phi)}([\Phi_t^X \circ \lambda(\tau)]_{\Phi_t^X(x)}) \\
&= \frac{d}{dt}\Big|_{t=0} \frac{d}{d\tau}\Big|_{\tau=0} (\phi \circ \Phi_t^X \circ \lambda(\tau)) \\
&= \frac{d}{dt}\Big|_{t=0} \frac{d}{d\tau}\Big|_{\tau=0} (\phi \circ \Phi_t^X \circ \phi^{-1} \circ \phi \circ \lambda(\tau)) \\
&= \frac{d}{dt}\Big|_{t=0} \left(\frac{d}{d\tau}\Big|_{\tau=0} (\phi \circ \Phi_t^X \circ \phi^{-1} \circ \phi \circ \lambda(\tau)) \right) \\
&= \frac{d}{dt}\Big|_{t=0} \left(D(\phi \circ \Phi_t^X \circ \phi^{-1})\Big|_{\phi(x)} \circ \frac{d}{d\tau}\Big|_{\tau=0} (\phi \circ \lambda(\tau)) \right) \\
&= \frac{d}{dt}\Big|_{t=0} \left(D(\phi \circ \Phi_t^X \circ \phi^{-1})\Big|_{\phi(x)} (v^1, \dots, v^n) \right) \\
&= \left(\frac{d}{dt}\Big|_{t=0} D(\phi \circ \Phi_t^X \circ \phi^{-1})\Big|_{\phi(x)} \right) (v^1, \dots, v^n) \\
&= D\left(\frac{d}{dt}\Big|_{t=0} \left(\phi \circ \Phi_t^X \circ \phi^{-1}\right) \right) \Big|_{\phi(x)} (v^1, \dots, v^n) \\
&= D(X^1 \circ \phi^{-1}, \dots, X^n \circ \phi^{-1})\Big|_{\phi(x)} (v^1, \dots, v^n) \\
&= \begin{pmatrix} \frac{\partial (X^1 \circ \phi^{-1})}{\partial x^1}\Big|_{\phi(x)} & \cdots & \frac{\partial (X^1 \circ \phi^{-1})}{\partial x^n}\Big|_{\phi(x)} \\ \vdots & \ddots & \vdots \\ \frac{\partial (X^n \circ \phi^{-1})}{\partial x^1}\Big|_{\phi(x)} & \cdots & \frac{\partial (X^n \circ \phi^{-1})}{\partial x^n}\Big|_{\phi(x)} \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \\
&= \left(\frac{\partial (X^1 \circ \phi^{-1})}{\partial x^i}\Big|_{\phi(x)} v^i, \dots, \frac{\partial (X^n \circ \phi^{-1})}{\partial x^i}\Big|_{\phi(x)} v^i \right) \\
&= \left(\frac{\partial X^1}{\partial x^i}(x) v^i, \dots, \frac{\partial X^n}{\partial x^i}(x) v^i \right).
\end{aligned}$$

Hence, we have

$$TT\phi \circ X^T\Big|_{TU}(v_x) = (T\phi(v_x), X^1(x), \dots, X^n(x), \frac{\partial X^1}{\partial x^i}(x)v^i, \dots, \frac{\partial X^n}{\partial x^i}(x)v^i),$$

and so

$$X^T\Big|_{TU}(v_x) = X^i(x)\frac{\partial}{\partial x^i}(v_x) + \frac{\partial X^i}{\partial x^j}(x)v^j\frac{\partial}{\partial v^i}(v_x).$$

Proposition C.8 (Tangent Lift at an Equilibrium Point [3])**.** *Let M be a C^∞-manifold, $X \in \Gamma^\infty(TM)$, and $x_0 \in M$ such that $X(x_0) = 0$. Recall the tangent bundle projection $\pi_{TM} : TM \to M$. Then, for all $v_{x_0} \in T_{x_0}M$, we have*

$$T_{v_{x_0}}\pi_{TM}(X^T(v_{x_0})) = 0.$$

In addition, there exists a unique $A_X(x_0) \in L(T_{x_0}M; T_{x_0}M)$ such that for all $v_{x_0} \in T_{x_0}M$, we have

$$X^T(v_{x_0}) = vlft_{v_{x_0}}(A_X(x_0) \cdot v_{x_0}).$$

*$A_X(x_0)$ is called the **linearization of** X **at** x_0.*

Note C.9 (Tangent Lift and Linearization)**.** The tangent lift is essentially the linearization of a vector field which is valid over the entire manifold. At an equilibrium point, with the tangent lift written in coordinates, only the Jacobian term remains, and thus there exists a linearization which is a linear map such that the tangent lift can be written in terms of a vertical lift.

Appendix D

Vector Bundles

In this appendix, we give a brief introduction to vector bundles, referring to [3]. The basic idea of a vector bundle is to attach a vector space to each point of some base manifold.

Definition D.1 (Local Model for a Vector Bundle)**.** Given $U \subseteq \mathbb{R}^n$ open, we call $U \times \mathbb{R}^k$ a **local model for a vector bundle**.

Definition D.2 (Local Vector Bundle Map)**.** Let $g : U \times \mathbb{R}^k \to V \times \mathbb{R}^l$, where $U \times \mathbb{R}^k$ and $V \times \mathbb{R}^l$ are local models for vector bundles.

(i) g is a **C^∞-local vector bundle map** if $g(x, v) = (g_1(x), g_2(x)v)$, where $g_1 : U \to V$ and $g_2 : U \to L(\mathbb{R}^k; \mathbb{R}^l)$ are C^∞.

(ii) g is a **C^∞-local vector bundle isomorphism** if $g(x, v) = (g_1(x), g_2(x)v)$, where $g_1 : U \to V$ is a C^∞-diffeomorphism and $g_2 : U \to L(\mathbb{R}^k; \mathbb{R}^l)$ is C^∞ such that for all $x \in U$, $g_2(x)$ is a vector space isomorphism.

Definition D.3 (Vector Bundle)**.** Let $A = \{(U_\alpha, \phi_\alpha)\}_{\alpha \in \Lambda}$ be an atlas for a set S. We call S a **C^∞-vector bundle** if, for all $\alpha \in \Lambda$, $\phi_\alpha(U_\alpha)$ is a local model for a vector bundle, and, for all $\alpha, \beta \in \Lambda$ such that $U_\alpha \cap U_\beta \neq \emptyset$,

$$\phi_\alpha \circ \phi_\beta^{-1}\Big|_{\phi_\beta(U_\alpha \cap U_\beta)} : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$$

is a C^∞-local vector bundle isomorphism. We call A a **C^∞-vector bundle atlas**. Two C^∞-vector bundle atlases A_1 and A_2 are **equivalent** if $A_1 \cup A_2$ is a C^∞-vector bundle atlas. A **C^∞-vector bundle structure** is an equivalence class of such atlases. A chart in one of these atlases is called an **admissible vector bundle chart**.

Definition D.4 (Base Space)**.** Let V be a vector bundle. The **base space of** V is defined as

$$B = \{v \in V \mid \exists \text{ an admissible vector bundle chart } (U, \phi) \text{ such that } \phi(v) = (x, 0) \in \phi(U)\}.$$

Note that this is well-defined.

Definition D.5 (Vector Bundle Projection)**.** The **vector bundle projection** is defined as $\pi : V \to B$ such that for all $v \in V$, we have $\pi(v) = \phi^{-1}(\mathrm{pr}_1(\phi(v)), 0)$, where (U, ϕ) is a local vector bundle chart for V around v. Note that this is also well-defined.

Definition D.6. (i) The base space B, when thought of as a submanifold of V, is called the **zero section** and is denoted $Z(V)$.
(ii) Given $b \in B$, the **fiber over** b is the set $V_b = \pi^{-1}(b)$.

Definition D.7 (Section of a Vector Bundle)**.** Let V be a C^∞-vector bundle, $\pi : V \to B$ the vector bundle projection, and consider a map $\xi : B \to V$. We call ξ a **C^∞-section of V** if ξ is C^∞ and $\pi \circ \xi = \mathrm{id}_B$. Define $\Gamma^\infty(V) \triangleq \{\xi : B \to V \mid \xi \text{ is a } C^\infty\text{-section of } V\}$.

Definition D.8 (Subbundle of a Vector Bundle)**.** Let V be a C^∞-vector bundle, $\pi : V \to B$ the vector bundle projection, and consider $W \subseteq V$. We call W a **C^∞-vector subbundle of** V if, for all $b \in B$, there exists a local chart (U, ϕ) of B and an admissible vector bundle chart $(\pi^{-1}(U), \psi)$ for V such that
(i) $b \in U$,
(ii) $\psi : \pi^{-1}(U) \to \phi(U) \times \mathbb{R}^l \times \mathbb{R}^{k-l}$,
(iii) $\phi \circ \pi \circ \psi^{-1}(x, u, v) = x$, where $\phi \circ \pi \circ \psi^{-1} : \phi(U) \times \mathbb{R}^l \times \mathbb{R}^{k-l} \to \phi(U)$, and
(iv) $\psi(\pi^{-1}(U) \cap W) = \phi(U) \times \mathbb{R}^l \times \{0\}$.

Note D.9. 1. W is a submanifold of V.
2. The idea of a vector subbundle is that one can smoothly select a subspace $W_b = V_b \cap W$ from each fiber V_b.
3. Define the **rank of** W **at** b as $\mathrm{rank}(W_b) = \dim(W_b)$.
4. The rank of a subbundle is locally constant. This can be too much of a restriction, so we consider generalized subbundles.

Definition D.10 (Generalized Subbundle of a Vector Bundle)**.** Let V be a C^∞-vector bundle, $\pi : V \to B$ the vector bundle projection, and consider $W \subseteq V$. We call W a **C^∞-generalized subbundle of** V if, for all $b_0 \in B$,
(i) $V_{b_0} \cap W$ is a subspace, and
(ii) there exists a neighbourhood $U \subseteq B$ of b_0 and a family $\{\xi_a\}_{a \in A}$ of C^∞-sections of $V\Big|_U$ (called **local generators of** W) such that for all $b \in U$, we have $V_b \cap W = \mathrm{span}_{\mathbb{R}}\{\xi_a(b) \mid a \in A\}$.

Definition D.11. Let V be a C^∞-vector bundle, and consider a C^r-generalized subbundle $W \subseteq V$ of V.
(i) Given $b \in B$, b is a **regular point of** W if there exists a neighbourhood $U \subseteq B$ of b such that the rank of the fiber is constant on U.
(ii) Given $b \in B$, b is a **singular point of** W if b is not a regular point of W.
(iii) W is **regular** if b is a regular point of W for all $b \in B$.

Note D.12. A regular C^∞-generalized subbundle is equivalently a C^∞-subbundle.

Definition D.13 (Vector Bundle Map)**.** Given C^∞-vector bundles V and W, with $\pi : V \to B$ and $\sigma : W \to A$, consider a C^∞-map $f : V \to W$. We call f a **C^∞-vector bundle map** if there exists a C^∞-map $f_0 : B \to A$ such that $f_0 \circ \pi = \sigma \circ f$ and $f\Big|_{V_b} \in L(V_b; W_{f_0(b)})$. If f is a C^∞-diffeomorphism, we call it a **C^∞-vector bundle isomorphism**.

Appendix E

Distributions, Codistributions and Integrability

Definition E.1 (Distribution)**.** Let M be a C^∞-manifold.

(i) We call D a **distribution on** M if, for all $x \in M$, $D(x) \subseteq T_xM$ is a subspace.

(ii) A distribution D on M is C^∞ if D is a C^∞-generalized subbundle of TM.

(iii) A C^∞-distribution D on M is **regular** if D is a C^∞-subbundle of TM.

Definition E.2 (Codistribution)**.** Let M be a C^∞-manifold.

(i) We call Λ a **codistribution on** M if, for all $x \in M$, $\Lambda(x) \subseteq T_x^*M$ is a subspace.

(ii) A codistribution Λ on M is C^∞ if Λ is a C^∞-generalized subbundle of T^*M.

(iii) A C^∞ codistribution Λ on M is **regular** if Λ is a C^∞-subbundle of T^*M.

Note E.3. For a distribution D or codistribution Λ on M, we will often write D_x and Λ_x instead of $D(x)$ and $\Lambda(x)$.

Definition E.4 (Involutivity)**.** (i) Given a C^∞-distribution D on M and $X \in \Gamma^\infty(TM)$, we say that X **belongs to** D if, for all $x \in M$, we have $X(x) \in D(x)$.

(ii) Given a C^∞-distribution D on M, we say that D is **involutive** if, for all $X_1, X_2 \in \Gamma^\infty(TM)$ such that X_1, X_2 belong to D, we have that $[X_1, X_2]$ belongs to D.

Definition E.5 (Local Integral Manifold)**.** Consider a C^∞-manifold M, and a C^∞ distribution D on M. Given $x_0 \in M$, and a C^∞-immersed submanifold S of a neighbourhood U of x_0, S is a **local integral manifold** if, for all $x \in S$, we have $T_xS \subseteq D(x)$.

Definition E.6 (Maximal Integral Manifold)**.** (i) Given S, a local integral manifold, S is **maximal** if, for all $x \in S$, we have $T_xS = D(x)$.

(ii) Given S, a maximal local integral manifold through x_0, S is the **maximal integral manifold for** D **through** x_0 if it contains any maximal local integral manifold through x_0.

Definition E.7 (Integrability)**.** D is **integrable** if, for all $x \in M$, there exists a maximal local integral manifold through x.

Theorem E.8 (Frobenius)**.** *A regular C^∞-distribution D is integrable if and only if D is involutive.*

Proposition E.9. *Let D be an integrable C^∞-distribution on M and $x_0 \in M$. If x_0 is a regular point for D, then there exists a local chart (U, ϕ) of M around x_0 such that for all $x \in U$, we have*

$$D(x) = span\left\{\frac{\partial}{\partial x^1}(x), \dots, \frac{\partial}{\partial x^k}(x)\right\}.$$

Conversely, if such a local chart exists, then $D\Big|_U$ is integrable.

References

[1] A. Baksa. On the stability of various equilibrium and stationary motion of nonholonomic systems. *Publications de l'Institut Mathématique*, 14(28):9–14, 1972.

[2] O. Bottema. On the small vibrations of non-holonomic systems. 1949.

[3] F. Bullo and A. D. Lewis. *Geometric Control of Mechanical Systems*. Number 49 in Texts in Applied Mathematics. Springer, 2005.

[4] L. Cai. On the stability of the equilibrium state and small oscillations of non-holonomic systems. *Dynamics and Stability of Systems: An International Journal*, 9(1):3–17, 1994.

[5] V. Čović, M. Vesković, and A. Obradović. On the instability of equilibrium of nonholonomic systems with nonhomogeneous constraints. *Mathematical and Computer Modelling*, 51(9–10):1097–1106, 2010.

[6] G. DeMarco. Linearization of nonholonomic systems at equilibrium points. *Rendiconti del Seminario Matematico*, 57(3):187–198, 1999.

[7] P. Hartman. *Ordinary Differential Equations*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2nd edition, 2002.

[8] R. Hui and A. A. Goldenberg. Stability of non-holonomic systems. *IEEE/RSJ International Workshop on Intelligent Robots and Sytems*, 1991.

[9] A. Y. Krasinskii. The stability and stabilization of the equilibrium positions of non-holonomic systems. *Journal of Applied Mathematics and Mechanics*, 52(2):152–158, 1988.

[10] J. M. Lee. *Introduction to Smooth Manifolds*. Number 218 in Graduate Texts in Mathematics. Springer, 2013.

[11] J. I. Neimark and N. A. Fufaev. Stability of the equilibrium states of nonholonomic systems. *Journal of Applied Mathematics and Mechanics*, 29(1):48–56, 1965.

[12] J. I. Neimark and N. A. Fufaev. *Dynamics of Nonholonomic Systems*, volume 33 of *Translations of Mathematical Monographs*. American Mathematical Society, 1972.

[13] M. Vesković and V. Čović. On the instability of equilibrium of a mechanical system with nonconservative forces. *Theoretical and Applied Mechanics*, 31(3-4):411–424, 2005.

[14] E. T. Whittaker. *A Treatise on Analytical Dynamics of Particles and Rigid Bodies*. Cambridge University Press, 4th edition, 1961.