REDUCTION, LINEARIZATION, AND STABILITY OF RELATIVE EQUILIBRIA FOR MECHANICAL SYSTEMS ON RIEMANNIAN MANIFOLDS

FRANCESCO BULLO AND ANDREW D. LEWIS

ABSTRACT. Consider a Riemannian manifold equipped with an infinitesimal isometry. For this setup, a unified treatment is provided, solely in the language of Riemannian geometry, of techniques in reduction, linearization, and stability of relative equilibria. In particular, for mechanical control systems, an explicit characterization is given for the manner in which reduction by an infinitesimal isometry, and linearization along a controlled trajectory "commute." As part of the development, relationships are derived between the Jacobi equation of geodesic variation and concepts from reduction theory, such as the curvature of the mechanical connection and the effective potential. As an application of our techniques, fiber and base stability of relative equilibria are studied. The paper also serves as a tutorial of Riemannian geometric methods applicable in the intersection of mechanics and control theory.

1. INTRODUCTION

Mechanical systems with symmetry have been a focus of an enormous research effort during the past few decades. This reflects of the importance of the notion of symmetry in physics. Of particular importance are those trajectories of a dynamical system that are also orbits for the symmetry group of the problem; these are relative equilibria. The stability of relative equilibria has both theoretical and practical importance. From a theoretical point of view, the relative equilibria, and their associated stability analysis, often give important insight into global behavior of solutions. In practical applications, relative equilibria arise in such diverse areas as fluid mechanics and underwater vehicle dynamics. In such problems, one often desires stability of a given relative equilibrium. Should such a relative equilibrium be naturally unstable, one must then develop ways of stabilizing it using control theory.

This paper is concerned with the analysis of relative equilibria for simple mechanical systems, i.e., those that are Lagrangian with kinetic energy minus potential energy Lagrangians. More specifically, in the paper we explore some of the Riemannian geometry associated with a system with symmetry in general, and a relative equilibrium in particular. Of course, much work has been done in this area, so let us locate our work in this body of literature. When talking about reduction using symmetry (as opposed to the more specific discussion of relative equilibria), one can work in a Hamiltonian or Lagrangian setting. The Hamiltonian setting is perhaps the more developed, going back to work of Arnol'd [1966], and the symplectic reduction work of Marsden and Weinstein [1974] and Meyer [1973]. This work has been presented in a fully developed manner in the books [Abraham and Marsden 1978, Guillemin and Sternberg 1984, Marsden 1992, Marsden and Ratiu 1999], for example. More recent research on Hamiltonian reduction theory is concerned with so-called singular reduction theory, where the regularity assumptions of the earlier work are relaxed. We refer the reader to [Ortega and Ratiu 2004] for the literature in this area. The Lagrangian theory of reduction is more recent, and we refer to the presentation in the book [Marsden and Ratiu 1999], and to the papers [Cendra, Marsden, Pekarsky, and Ratiu 2003, Cendra, Marsden, and Ratiu 2001] as representative of the work in this area.

In the Hamiltonian theory of reduction using symmetry, the symplectic, or more generally Poisson, structure plays the prominent role. Indeed, much of the work in this area is concerned with general symplectic manifolds rather than cotangent bundles. On the Lagrangian side, the emphasis has been on the reduction of variational principles, the motivation for this being that variational principles are fundamental for Lagrangian mechanics. In this paper we focus on Lagrangians that are of the kinetic energy minus potential energy form. For such Lagrangians, an important role is played by the kinetic energy Riemannian metric and its attendant Levi-Civita connection. With this as motivation, we study reduction and relative equilibria strictly in terms of Riemannian geometry. We do not wish to assert that the Riemannian geometry approach we give here is superior to the variational approach; we too believe in the primacy of the variational principle in Lagrangian mechanics. However, the extra structure offered by the Riemannian metric does lead one to naturally ask whether a theory of reduction based purely on Riemannian geometry is possible and, if so, revealing. We show that it is possible, and we believe that it is revealing. This is presented in Section 3.

Another emphasis of this paper is linearization. We start this in a rather general way. Since the stabilization of relative equilibria is of interest [Bullo

2000, Jalnapurkur and Marsden 2000, 2001], we develop our theory of linearization in the setting of control systems. One of the main contributions of the paper is to explain the geometry that connects the unreduced and reduced linearizations. Since the unreduced linearization is about a trajectory and not an equilibrium point, we first present a rather general geometric theory of linearization of a control-affine system about an arbitrary controlled trajectory. We then specialize this to so-called affine connection control systems; these are control-affine systems whose state manifold is the tangent bundle of a configuration manifold, whose drift vector field is the geodesic spray for an affine connection, and whose control vector fields are vertical lifts of vector fields on the configuration manifold. As is explained in [Bullo and Lewis 2004, Chapter 4], affine connection control systems arise in the modeling of a large class of mechanical systems, including those with nonholonomic constraints. The linearization of an affine connection control system is related to the Jacobi equation of geodesic variation, and we explicitly develop this relationship in Proposition 5.6; this association is not surprising, but it does not appear to have been presented before. We then specialize the affine connection control system setup to the special case where the affine connection is the Levi-Civita connection. In Theorem 5.10 we explicitly show how the general linearization for an affine connection control system is related to the linearization of the reduced system. Again, it is not surprising that this should be possible. However, the explicit development does not seem to have been presented, and is not entirely trivial.

The final topic of the paper is stability of relative equilibria. We present here two stability theorems, one linear and one nonlinear, the latter relying on the Lyapunov Stability Criterion and the LaSalle Invariance Principle. As with our study of linearization, one of the essential contributions here is an understanding, in the context of Riemannian geometry, of the relationship between the unreduced linearized energy and the reduced linearized energy. Here again, that there should be a relationship is expected, but the explicit development, culminating in Proposition 6.6, is new and not completely trivial. In particular, we utilize the Sasaki metric in our presentation, something that has not been done before to the best of our knowledge.

2. Simple mechanical (control) systems with symmetry

In this section we review some well-known concepts in geometric mechanics, and mechanical systems with symmetry. We refer the reader to [Bloch 2003, Bullo and Lewis 2004, Marsden and Ratiu 1999] for additional discussion. Our presentation is in the context of systems with controls, since the development of Sections 3 and 5 is also done in this context. Our notation mostly follows [Bullo and Lewis 2004].

Notation. In this section we quickly present some of the concepts and notation we use.

If U and V are $\mathbb{R}$-vector spaces, $L(\mathsf{U};\mathsf{V})$ denotes the set of linear maps from U to V.

Throughout the paper we work with manifolds and geometric object that are of class C^∞, unless stated to the contrary. If M is a manifold, then $C^\infty(\mathsf{M})$ denotes the set of C^∞-functions on M. The tangent bundle of M is denoted by $\pi_{\mathsf{TM}}\colon \mathsf{TM} \to \mathsf{M}$. If $(x^1,\dots,x^n)$ are coordinates for M, then natural tangent bundle coordinates for TM are denoted by $((x^1,\dots,x^n),(v^1,\dots,v^n))$, or sometimes more compactly by $(\boldsymbol{x},\boldsymbol{v})$. The derivative of a map $\phi\colon \mathsf{M} \to \mathsf{N}$ is denoted by $T\phi\colon \mathsf{TM} \to \mathsf{TN}$, and the restriction of the derivative to the tangent space $\mathsf{T}_x\mathsf{M}$ is denoted by $T_x\phi$. If $\pi\colon \mathsf{E} \to \mathsf{B}$ is a vector bundle, then $\Gamma^\infty(\mathsf{E})$ denotes the set of C^∞-sections of E. A vector field $X\colon \mathsf{E} \to \mathsf{TE}$ on the total space of a vector bundle $\pi\colon \mathsf{E} \to \mathsf{B}$ is a **linear vector field** over a vector field X_0 on B if X is π-related to X_0 and if the following diagram commutes:

$$
\begin{array}{ccc}
\mathsf{E} & \xrightarrow{\ X\ } & \mathsf{TE} \\
\pi \downarrow & & \downarrow T\pi \\
\mathsf{B} & \xrightarrow[\ X_0\ ]{} & \mathsf{TB}
\end{array}
$$

The zero section of a vector bundle E is denoted by $Z(\mathsf{E})$. If B is a $(0,2)$-tensor field on a manifold M, then $B^\flat\colon \mathsf{TM} \to \mathsf{T}^*\mathsf{M}$ is the vector bundle map defined by $\langle B(u_x);v_x\rangle = B(v_x,u_x)$ for $u_x,v_x \in \mathsf{TM}$. If $B^\flat$ in an isomorphism, then its inverse is denoted by $B^\sharp$. If $\mathbb{G}$ is a Riemannian metric on M and if $f \in C^\infty(\mathsf{M})$, then $\mathrm{grad} f$ denotes the vector field $\mathrm{grad} f = \mathbb{G}^\sharp \circ \boldsymbol{d}f$. Here $\boldsymbol{d}f$ is the differential of f. Also, $\|\cdot\|_{\mathbb{G}}$ denotes the norm on the fibers of TM defined by $\mathbb{G}$.

Typically, $I \subset \mathbb{R}$ will denote an interval. If $\gamma\colon I \to \mathsf{M}$ is a differentiable curve on M, then its tangent vector field is denoted by $\gamma'\colon I \to \mathsf{TM}$. A curve $\gamma\colon I \to \mathsf{M}$ is **locally absolutely continuous** (abbreviated **LAC**) if $f \circ \gamma$ is locally absolutely continuous for every $f \in C^\infty(\mathsf{M})$. If γ' is LAC, then γ is **locally absolutely differentiable** (abbreviated **LAD**).

For a vector field X, the flow of X is denoted by $(t,x) \mapsto \Phi^X_{0,t}(x)$, so that the integral curve through x is $t \mapsto \Phi^X_{0,t}(x)$. We will be considering time-varying vector fields at various points in the paper, and since we wish to allow fairly general time-dependence, we should be precise about how we do this. To this end, by a **time-dependent vector field** on M we shall mean a map $X\colon I \times \mathsf{M} \to \mathsf{TM}$ with the following properties:

(1) for each $t \in I$, the map $X_t\colon x \mapsto X(t,x)$ defines a C^∞-vector field;

(2) for each $x \in \mathsf{M}$, the map $t \mapsto X(t,x)$ is measurable (meaning its components are measurable in some, and so any, set of coordinates);

(3) for each $k \in \natural$, each collection $X_1, \ldots, X_k$ of C^∞-vector fields on M, each C^∞-one-form α on M, and each compact subset $K \subset \mathsf{M}$, there exists a positive locally integrable function $\psi \colon I \to \mathbb{R}$ such that

$$|\mathscr{L}_{X_1} \cdots \mathscr{L}_{X_k} \langle \alpha; X_t \rangle (x)| \leq \psi(t), \qquad x \in K.$$

The usual Carathéodory theory shows that time-dependent vector fields defined in this fashion have LAC integral curves, and that their flows are of class C^∞ with respect to initial conditions. We shall use the notation $(t,x) \mapsto \Phi_{0,t}^X(x)$ to denote the flow of a time-dependent vector field X, i.e., we use the same notation for the flow of both time-independent and time-dependent vector fields.

2.1. **Simple mechanical (control) systems.** A ***forced simple mechanical control system*** is a 5-tuple $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F} = \{F^1, \ldots, F^m\})$ where

(1) Q is the ***configuration manifold***, here assumed to be of pure dimension n,

(2) $\mathbb{G}$ is the ***kinetic energy metric***, which is a Riemannian metric on Q,

(3) V is the ***potential energy***, a function on Q,

(4) $F \colon \mathsf{TQ} \to \mathsf{T}^*\mathsf{Q}$ is a bundle map over id_Q, which is the ***external force***, and

(5) $\{F^1, \ldots, F^m\}$ are the ***control forces***, which are one-forms on Q.

The presentation in [Bullo and Lewis 2004] also includes an extra piece of data in the form of a set $U \subset \mathbb{R}^m$ where the control takes values. However, for the purposes of this paper, this is not important, so we do not include it (that is to say, we take $U = \mathbb{R}^m$). The equations governing a simple mechanical control system are

$$\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -\mathrm{grad}\,V(\gamma(t)) + \mathbb{G}^\sharp \circ F(\gamma'(t)) + \sum_{a=1}^m u^a(t)\mathbb{G}^\sharp \circ F^a(\gamma(t)), \quad (1)$$

where $\overset{\mathbb{G}}{\nabla}$ is the Levi-Civita connection associated with $\mathbb{G}$. We also denote by $\overset{\mathbb{G}}{\Gamma}{}^i_{jk}$ the Christoffel symbols for $\overset{\mathbb{G}}{\nabla}$. A ***controlled trajectory*** for Σ is thus a pair (γ, u) where $u \colon I \to \mathbb{R}^m$ is locally integrable and defined on an interval $I \subset \mathbb{R}$, and where $\gamma \colon I \to \mathsf{Q}$ is LAD and satisfies (1).

Simple mechanical control systems are examples of a more general class of systems which we now introduce. A ***forced affine connection control system*** is a 4-tuple $(\mathsf{Q}, \nabla, Y, \mathscr{Y} = \{Y_1, \ldots, Y_m\})$, where

(1) Q is a manifold as above,

(2) ∇ is an affine connection on Q,
(3) $Y \colon \mathsf{T}\mathsf{Q} \to \mathsf{T}\mathsf{Q}$ is a bundle map over id_{Q}, and
(4) $\{Y_1, \ldots, Y_m\}$ are vector fields on Q.

The equations governing such a system are

$$\nabla_{\gamma'(t)}\gamma'(t) = Y(\gamma'(t)) + \sum_{a=1}^{m} u^a(t) Y_a(\gamma(t)).$$

We can define the notion of a controlled trajectory for an affine connection system in the same way as we did for a forced simple mechanical control system. Moreover, note that a forced simple mechanical control system is also a forced affine connection control system by taking $\nabla = \overset{\mathbb{G}}{\nabla}$, $Y = \mathrm{grad}V \circ \pi_{\mathsf{T}\mathsf{Q}} + \mathbb{G}^{\sharp} \circ F$. In [Bullo and Lewis 2004, Chapter 4] it is shown that affine connection control systems model a large class of mechanical systems, including those with nonholonomic constraints.

Parts of the paper will be concerned with systems without forces, so let us introduce notation for these. A **forced simple mechanical system** is a quadruple $(\mathsf{Q}, \mathbb{G}, V, F)$, where all the data are as for a forced simple mechanical control system. Of course, the governing equations are

$$\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -\mathrm{grad}V(\gamma(t)) + \mathbb{G}^{\sharp} \circ F(\gamma'(t)).$$

If the external force F is omitted from list of data, the resulting triple $(\mathsf{Q}, \mathbb{G}, V)$ is a **simple mechanical system**.

2.2. **Systems with symmetry.** Now let us add symmetry to the above formulations. In this paper we shall only consider a single infinitesimal symmetry, rather than the more usual situation where one considers symmetry of the system under a Lie group G. The extension of the results in the paper to this more general setup of a Lie group symmetry is currently ongoing.

Thus we consider a forced simple mechanical system $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F} = \{F^1, \ldots, F^m\})$ and a vector field X on Q having the following properties:

(1) X is an infinitesimal isometry for $\mathbb{G}$, i.e., $\mathscr{L}_X \mathbb{G} = 0$;
(2) $\mathscr{L}_X V = 0$;
(3) $\mathscr{L}_X(F(Y)) = F(\mathscr{L}_X Y)$ for all vector fields Y on Q;
(4) $\mathscr{L}_X F^a = 0$ for $a \in \{1, \ldots, m\}$.

A vector field X having these properties is an **infinitesimal symmetry** for Σ. The infinitesimal symmetry X is **complete** if X is a complete vector field. A vector field having only property 1 is an **infinitesimal isometry** for $\mathbb{G}$.

Let us give some useful properties of infinitesimal isometries of $\mathbb{G}$.

Proposition 2.1 (Characterization of infinitesimal isometry). *Let $(\mathsf{Q}, \mathbb{G})$ be a Riemannian manifold. The following statements hold:*

(i) the vector field X is an infinitesimal isometry if and only if the covariant differential $\overset{\mathbb{G}}{\nabla} X$ is skew-symmetric with respect to $\mathbb{G}$; that is, for all $Y, Z \in \Gamma^\infty(\mathsf{T}\mathsf{Q})$,

$$\mathbb{G}(Y, \overset{\mathbb{G}}{\nabla}_Z X) + \mathbb{G}(Z, \overset{\mathbb{G}}{\nabla}_Y X) = 0; \tag{2}$$

(ii) if the vector field X is an infinitesimal isometry, then the function $q \mapsto \|X\|_{\mathbb{G}}^2(q)$ is X-invariant and satisfies

$$\tfrac{1}{2}\mathrm{grad}\,\|X\|_{\mathbb{G}}^2 = -\overset{\mathbb{G}}{\nabla}_X X. \tag{3}$$

Proof. Regarding part (i), we compute, for $Y, Z \in \Gamma^\infty(\mathsf{T}\mathsf{Q})$,

$$\begin{aligned}
\mathbb{G}(Y, \overset{\mathbb{G}}{\nabla}_Z X) + \mathbb{G}(Z, \overset{\mathbb{G}}{\nabla}_Y X) &= \mathbb{G}(Y, \overset{\mathbb{G}}{\nabla}_X Z + [Z, X]) + \mathbb{G}(Z, \overset{\mathbb{G}}{\nabla}_X Y + [Y, X]) \\
&= \mathscr{L}_X(\mathbb{G}(Y, Z)) - \mathbb{G}(Y, [X, Z]) - \mathbb{G}(Z, [X, Y]) \\
&= (\mathscr{L}_X \mathbb{G})(Y, Z).
\end{aligned}$$

The result follows since this equality holds for all $Y, Z \in \Gamma^\infty(\mathsf{T}\mathsf{Q})$.

To prove equation (3), let $Z \in \Gamma^\infty(\mathsf{T}\mathsf{Q})$ and compute

$$\mathbb{G}(Z, \mathrm{grad}\,\|X\|_{\mathbb{G}}^2) = \mathscr{L}_Z \|X\|_{\mathbb{G}}^2 = 2\mathbb{G}(X, \overset{\mathbb{G}}{\nabla}_Z X) = -2\mathbb{G}(Z, \overset{\mathbb{G}}{\nabla}_X X).$$

To show that $q \mapsto \|X\|_{\mathbb{G}}^2(q)$ is X-invariant, note that

$$\mathscr{L}_X \|X\|_{\mathbb{G}}^2 = -2\mathbb{G}(X, \overset{\mathbb{G}}{\nabla}_X X) = 0.$$

$\square$

Associated to an infinitesimal isometry X for $\mathbb{G}$ is the **X-momentum map** $J_X \colon \mathsf{T}\mathsf{Q} \to \mathbb{R}$ defined by $J_X(v_q) = \mathbb{G}(X(q), v_q)$. One can show [Bullo and Lewis 2004, Theorem 5.69] that the X-momentum map satisfies the evolution equation

$$\frac{\mathrm{d}}{\mathrm{d}t} J_X(\gamma'(t)) = \left\langle \bar{F}_u(t, \gamma'(t)); X(\gamma(t)) \right\rangle$$

along a controlled trajectory (γ, u), where

$$\bar{F}_u(t, v_q) = F(v_q) + \sum_{a=1}^{} u^a(t) F^a(v_q).$$

In particular, if $F = 0$ and $u^a(t) = 0$, $a \in \{1, \ldots, m\}$, for all t, then the X-momentum is conserved; this is the Noether Conservation Law in this particular setup.

2.3. **Relative equilibria.** For mechanical systems with symmetry, relative equilibria are important since they often give initial insights into the behavior of a system. In control theory, relative equilibria are often elementary mechanical motions that can be used as a basis for, for example, motion planning. For additional discussion we refer to [Bloch 2003, Marsden 1992, Marsden and Ratiu 1999].

Definition 2.2 (Relative equilibrium). *Let X be a complete infinitesimal symmetry for $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ and let $\chi \colon \mathbb{R} \to \mathsf{Q}$ be a maximal integral curve of X.*

    (i) *The curve χ is a **relative equilibrium** for Σ if it is a solution to the equation of motion (27).*
    (ii) *The relative equilibrium χ is **regular** if χ is an embedding.*   ●

To characterize relative equilibria, it is convenient to have the following notion. We define the **energy** for a forced simple mechanical control system $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ by $E(v_q) = \frac{1}{2}\mathbb{G}(v_q, v_q) + V(q)$.

Definition 2.3 (Effective energy). *Let X be a complete infinitesimal symmetry for $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$, and let $J_X \colon \mathsf{TQ} \to \mathbb{R}$ be the associated momentum map.*

    (i) *The **effective potential** $V_X \colon \mathsf{Q} \to \mathbb{R}$ is the function $V_X(q) = V(q) - \frac{1}{2}\|X\|_{\mathbb{G}}^2(q)$.*
    (ii) *The **effective energy** $E_X \colon \mathsf{TQ} \to \mathbb{R}$ is the function $E_X(v_q) = E(v_q) - J_X(v_q)$.*   ●

The following result characterizes the effective energy.

Lemma 2.4. $E_X(v_q) = V_X(q) + \frac{1}{2}\|v_q - X(q)\|_{\mathbb{G}}^2$.

Proof. The lemma is a consequence of the following "completing the square" computation:

$$
\begin{aligned}
E(v_q) - J_X(v_q) &= V(q) + \tfrac{1}{2}\|v_q\|_{\mathbb{G}}^2 - \mathbb{G}(v_q, X(q)) \\
&= V(q) - \tfrac{1}{2}\|X\|_{\mathbb{G}}^2(q) + \tfrac{1}{2}\|X\|_{\mathbb{G}}^2(q) + \tfrac{1}{2}\|v_q\|_{\mathbb{G}}^2 - \mathbb{G}(v_q, X(q)) \\
&= V_X(q) + \tfrac{1}{2}\|v_q - X(q)\|_{\mathbb{G}}^2,
\end{aligned}
$$

as desired.   □

Using this result, we can characterize relative equilibria as follows.

Proposition 2.5 (Existence of relative equilibria). *Consider a forced simple mechanical system $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ and let X be a complete infinitesimal symmetry for Σ. Then a maximal integral curve $\chi \colon \mathbb{R} \to \mathsf{Q}$ of X is a relative equilibrium for Σ if and only if $\boldsymbol{d}V_X(\chi(t)) = F(\chi'(t))$ for some (and hence*

for all) $t \in \mathbb{R}$. *In particular, if* $F(\chi'(t)) = 0$ *for some (and hence for all)* $t \in \mathbb{R}$, *then* χ *is a relative equilibrium for* Σ *if and only if* V_X *has a critical point at* $\chi(t)$ *for some (hence for all)* $t \in \mathbb{R}$.

Proof. We proceed in a direct way. Assume that χ is a maximal integral curve of X. The curve χ satisfies the equations of motion (27) if and only if

$$
\begin{aligned}
0 &= \overset{\mathbb{G}}{\nabla}_{\chi'(t)} \chi'(t) + \mathrm{grad}V(\chi(t)) - \mathbb{G}^\sharp \circ F(\chi'(t)) \\
&= \overset{\mathbb{G}}{\nabla}_{X(\chi(t))} X(\chi(t)) + \mathrm{grad}V(\chi(t)) - \mathbb{G}^\sharp \circ F(\chi'(t)) \\
&= \mathrm{grad}\big(-\tfrac{1}{2} \|X\|_{\mathbb{G}}^2 + V \big)(\chi(t)) - \mathbb{G}^\sharp \circ F(\chi'(t)) \\
&= \mathbb{G}^\sharp \circ \boldsymbol{d}V_X(\chi(t)) - \mathbb{G}^\sharp \circ F(\chi'(t)).
\end{aligned}
$$

Therefore, χ is a relative equilibrium if and only if $\boldsymbol{d}V_X(\chi(t)) = F(\chi'(t)$ for all $t \in \mathbb{R}$. Because V_X is X-invariant, $\boldsymbol{d}V_X(\chi(t))$ is independent of t. Since X is an infinitesimal symmetry for Σ, $\mathscr{L}_X F(X) = F[X, X] = 0$. Therefore, $F(\chi'(t))$ is also independent of t, and so the result follows. $\qquad \square$

3. Reduction by an infinitesimal symmetry

In this section we shall investigate the dynamics of a simple mechanical system with a complete infinitesimal symmetry. As mentioned in the introduction, our treatment is limited in scope; for a more comprehensive treatment we refer the reader to the related work in [Cendra, Marsden, and Ratiu 2001].

3.1. **Preliminary constructions.** We begin with an assumption on the nature of an infinitesimal isometry X that will hold throughout the paper.

Assumption 3.1. *Let X be a complete infinitesimal symmetry for a simple mechanical system* $(\mathsf{Q}, \mathbb{G}, V)$. *We assume that the set of X-orbits is a manifold, say* B, *and that the projection* $\pi_\mathsf{B} \colon \mathsf{Q} \to \mathsf{B}$ *is a surjective submersion.•*

In Figure 1 we provide an illustration of how the reader should think about Assumption 3.1.

Now we introduce some useful definitions. The ***vertical distribution*** VQ is the distribution on Q generated by the vector field X. The ***horizontal distribution*** HQ is the $\mathbb{G}$-orthogonal complement of VQ so that $\mathsf{TQ} = \mathsf{HQ} \oplus \mathsf{VQ}$.

At each $q \in \mathsf{Q}$, the linear map $T_q\pi_\mathsf{B} \colon \mathsf{T}_q\mathsf{Q} \to \mathsf{T}_{\pi_\mathsf{B}(q)}\mathsf{B}$ is a surjection. Therefore, the map $T_q\pi_\mathsf{B}|\mathsf{H}_q\mathsf{Q} \colon \mathsf{H}_q\mathsf{Q} \to \mathsf{T}_{\pi_\mathsf{B}(q)}\mathsf{B}$ is a linear isomorphism. The ***horizontal lift*** of the tangent vector $v_b \in \mathsf{T}_b\mathsf{B}$ at point $q \in \mathsf{Q}$ is the tangent vector in $\mathsf{T}_q\mathsf{Q}$ given by

$$
\mathrm{hlft}_q(v_b) = \left(T_q\pi_\mathsf{B}|\mathsf{H}_q\mathsf{Q} \right)^{-1}(v_b). \tag{4}
$$
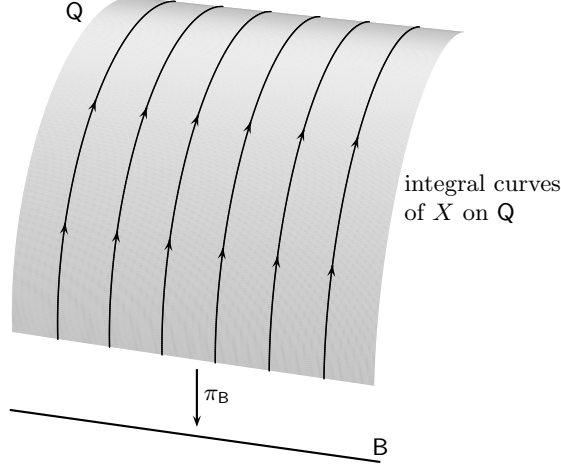
FIGURE 1. The bundle of X-orbits

Furthermore, the **horizontal lift** of the vector field $Y \in \Gamma^\infty(\mathsf{T}\mathsf{B})$ is the vector field $\mathrm{hlft}(Y) \in \Gamma^\infty(\mathsf{T}\mathsf{Q})$ defined by $\mathrm{hlft}(Y)(q) = \mathrm{hlft}_q(Y(\pi_\mathsf{B}(q)))$. This vector field is X-invariant and takes values in $\mathsf{H}\mathsf{Q}$. If $\eta \colon I \to \mathsf{B}$ is a C^1-curve with $\eta(0) = b_0$, and if $q_0 \in \pi_\mathsf{B}^{-1}(b_0)$, then the **horizontal lift** of η through q_0 is denoted by $\mathrm{hlft}_{q_0}(\eta) \colon I \to \mathsf{Q}$.

Next, we note that is possible to project certain X-invariant objects from Q onto B in natural ways. For example, given the Riemannian metric $\mathbb{G}$ on Q, we define the Riemannian metric $\mathbb{G}_\mathsf{B}$ on B, called the **projected metric**, by

$$\mathbb{G}_\mathsf{B}(b)(v_b, w_b) = \mathbb{G}(q)(\mathrm{hlft}(v_b), \mathrm{hlft}(w_b)),$$

for $b \in \mathsf{B}$, $v_b, w_b \in \mathsf{T}_b\mathsf{B}$, and $q \in \pi_\mathsf{B}^{-1}(b)$. Let $\overset{\mathbb{G}_\mathsf{B}}{\nabla}$ be the Levi-Civita affine connection on the Riemannian manifold $(\mathsf{B}, \mathbb{G}_\mathsf{B})$ and, given a function $f \colon \mathsf{B} \to \mathbb{R}$, let $\mathrm{grad}_\mathsf{B} f$ be its gradient with respect to the metric $\mathbb{G}_\mathsf{B}$. Given an X-invariant function $h \colon \mathsf{Q} \to \mathbb{R}$, let $h_\mathsf{B} \colon \mathsf{B} \to \mathbb{R}$ be its projection defined by $h_\mathsf{B}(b) = h(q)$, where $\pi_\mathsf{B}(q) = b$. One can then show that

$$T\pi_\mathsf{B} \circ \mathrm{grad} h = \mathrm{grad}_\mathsf{B} h_\mathsf{B} \circ \pi_\mathsf{B}. \tag{5}$$

We will also find the following result useful.

Lemma 3.2. *For all* $Y, Z \in \Gamma^\infty(\mathsf{T}\mathsf{B})$, $T\pi_\mathsf{B} \circ \left(\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y)} \mathrm{hlft}(Z)\right) = \overset{\mathbb{G}_\mathsf{B}}{\nabla}_Y Z \circ \pi_\mathsf{B}$.

Proof. Let $W, Y, Z \in \Gamma^\infty(\mathsf{TB})$. The Koszul formula defining the Levi-Civita affine connection gives

$$
\begin{aligned}
\mathbb{G}(\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y)}\mathrm{hlft}(Z), \mathrm{hlft}(W)) = {} & \tfrac{1}{2}\big(\mathscr{L}_{\mathrm{hlft}(Y)}(\mathbb{G}(\mathrm{hlft}(Z), \mathrm{hlft}(W))) \\
& + \mathscr{L}_{\mathrm{hlft}(Z)}(\mathbb{G}(\mathrm{hlft}(W), \mathrm{hlft}(Y))) - \mathscr{L}_{\mathrm{hlft}(W)}(\mathbb{G}(\mathrm{hlft}(Y), \mathrm{hlft}(Z))) \\
& + \mathbb{G}([\mathrm{hlft}(Y), \mathrm{hlft}(Z)], \mathrm{hlft}(W)) - \mathbb{G}([\mathrm{hlft}(Y), \mathrm{hlft}(W)], \mathrm{hlft}(Z)) \\
& \qquad\qquad - \mathbb{G}([\mathrm{hlft}(Z), \mathrm{hlft}(W)], \mathrm{hlft}(Y))\big).
\end{aligned}
$$

Since W is π_{B}-related to $\mathrm{hlft}(W)$, and similarly for Y and Z, it follows that $[W, Y]$ is π_{B}-related to $[\mathrm{hlft}(W), \mathrm{hlft}(Y)]$, and similarly for the other Lie brackets. Since the function $q \mapsto \mathbb{G}(\mathrm{hlft}(W)(q), \mathrm{hlft}(Y)(q))$ is X-invariant, and since the vector fields $\mathrm{hlft}(W)$, $\mathrm{hlft}(Y)$, and $\mathrm{hlft}(Z)$ are X-invariant, we have

$$
\mathscr{L}_{\mathrm{hlft}(Z)}(\mathbb{G}(\mathrm{hlft}(W), \mathrm{hlft}(Y))) = \mathscr{L}_Z\mathbb{G}_{\mathsf{B}}(W, Y) \circ \pi_{\mathsf{B}},
$$

and similarly for similar terms, permuting W, Y, and Z. Also,

$$
\mathbb{G}(\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y)}\mathrm{hlft}(Z), \mathrm{hlft}(W)) = \mathbb{G}_{\mathsf{B}} \circ \pi_{\mathsf{B}}(T\pi_{\mathsf{B}} \circ (\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Z)}\mathrm{hlft}(Y)), W \circ \pi_{\mathsf{B}}).
$$

Thus we conclude that

$$
\begin{aligned}
\mathbb{G}_{\mathsf{B}} \circ {} & \pi_{\mathsf{B}}(T\pi_{\mathsf{B}} \circ (\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y)}\mathrm{hlft}(Z)), W \circ \pi_{\mathsf{B}}) \\
= {} & \tfrac{1}{2}\big(\mathscr{L}_Y(\mathbb{G}_{\mathsf{B}}(Z, W)) \circ \pi_{\mathsf{B}} + \mathscr{L}_Z(\mathbb{G}_{\mathsf{B}}(W, Y)) \circ \pi_{\mathsf{B}} - \mathscr{L}_W(\mathbb{G}_{\mathsf{B}}(Y, Z)) \circ \pi_{\mathsf{B}} \\
& + \mathbb{G}_{\mathsf{B}}([Y, Z], W) \circ \pi_{\mathsf{B}} - \mathbb{G}_{\mathsf{B}}([Y, W], Z) \circ \pi_{\mathsf{B}} - \mathbb{G}_{\mathsf{B}}([Z, W], Y) \circ \pi_{\mathsf{B}}\big).
\end{aligned}
$$

Applying the Koszul formula to $\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}$, we get the desired conclusion. $\qquad\square$

Let us now return to the study of a simple mechanical system $(\mathsf{Q}, \mathbb{G}, V)$ with infinitesimal symmetry X. We need to introduce two final concepts. First, for $\lambda \in \mathbb{R}$, we define the ***parameterized effective potential*** $V_{X,\lambda}^{\mathrm{eff}} \colon \mathsf{Q} \to \mathbb{R}$ (note the slight difference with the previously introduced effective potential V_X) by

$$
V_{X,\lambda}^{\mathrm{eff}}(q) = V(q) - \frac{\lambda^2}{2}\|X\|_{\mathbb{G}}^2(q),
$$

and the ***parameterized amended potential*** $V_{X,\lambda}^{\mathrm{amd}} \colon \mathsf{Q} \to \mathbb{R}$ by

$$
V_{X,\lambda}^{\mathrm{amd}}(q) = V(q) + \frac{\lambda^2}{2}\|X\|_{\mathbb{G}}^{-2}(q).
$$

The parameterized effective potential and the parameterized amended potential are X-invariant functions on Q. Second, we define the ***gyroscopic***

tensor C_X as the $(1,1)$-tensor field on B such that, for any $q \in \pi_{\mathsf{B}}^{-1}(b)$ and $v_b \in \mathsf{T}_b\mathsf{B}$,

$$C_X(v_b) = -2T_q\pi_{\mathsf{B}}\left(\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}_q(v_b)}X(q)\right).$$

One can show that C_X is well-defined in the sense that the choice of q is immaterial, and that C_X is skew-symmetric with respect to $\mathbb{G}_{\mathsf{B}}$; that is, for all $v_b, w_B \in \mathsf{T}_b\mathsf{B}$,

$$\mathbb{G}_{\mathsf{B}}(v_b, C_X(w_b)) + \mathbb{G}_{\mathsf{B}}(w_b, C_X(v_b)) = 0.$$

The skew-symmetry of C_X is an immediate consequence of the skew-symmetry of $\overset{\mathbb{G}}{\nabla}X$.

3.2. **The reduced dynamics.** We are finally ready to state the main result of this section.

Theorem 3.3 (Reduced dynamics). *Let* $(\mathsf{Q}, \mathbb{G}, V)$ *be a simple mechanical system with a complete infinitesimal symmetry X satisfying Assumption 3.1. The following statements about the C^1-curves $\gamma\colon I \to \mathsf{Q}$, $\eta\colon I \to \mathsf{B}$, $v\colon I \to \mathbb{R}$, and $\mu\colon I \to \mathbb{R}$ are equivalent:*

   (i) γ *satisfies*

$$\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -\,\mathrm{grad}V(\gamma(t)),$$
$$\gamma'(0) = v_{q_0} \in \mathsf{T}_{q_0}\mathsf{Q},$$

   and, in turn, η, μ, and v are defined by $\eta(t) = \pi_{\mathsf{B}} \circ \gamma(t)$, $\mu(t) = J_X \circ \gamma'(t)$, and $v(t) = (J_X \circ \gamma'(t))(\|X\|_{\mathbb{G}}^{-2} \circ \gamma(t))$, respectively;

   (ii) η *and* v *together satisfy*

$$\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{\eta'(t)}\eta'(t) = -\,\mathrm{grad}_{\mathsf{B}}\left(V_{X,v(t)}^{\mathrm{eff}}\right)_{\mathsf{B}}(\eta(t)) + v(t)C_X(\eta'(t)),$$
$$\dot{v}(t) = -\,\frac{v(t)}{(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t))}\langle \boldsymbol{d}(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t)); \eta'(t)\rangle,$$
$$\eta'(0) = T_{q_0}\pi_{\mathsf{B}}(v_{q_0}), \quad v(0) = J_X(v_{q_0})\|X\|_{\mathbb{G}}^{-2}(q_0),$$

   and, in turn, γ and μ are defined by $\gamma(t) = \Phi_{0,t}^{vX}(\mathrm{hlft}_{q_0}(\eta)(t))$, and $\mu(t) = v(t)((\|X\|_{\mathbb{G}}^2)_{\mathsf{B}} \circ \eta(t))$, respectively;

   (iii) η *and* μ *together satisfy*

$$\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{\eta'(t)}\eta'(t) = -\,\mathrm{grad}_{\mathsf{B}}\left(V_{X,\mu(t)}^{\mathrm{amd}}\right)_{\mathsf{B}}(\eta(t)) + \frac{\mu(t)}{(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t))}C_X(\eta'(t)),$$
$$\dot{\mu}(t) = 0,$$
$$\eta'(0) = T_{q_0}\pi_{\mathsf{B}}(v_{q_0}), \quad \mu(0) = J_X(v_{q_0}),$$

and, in turn, v and γ are defined by $v(t) = \mu(t)((\|X\|_{\mathbb{G}}^{-2})_{\mathsf{B}} \circ \eta(t))$ and $\gamma(t) = \Phi_{0,t}^{vX}(\mathrm{hlft}_{q_0}(\eta)(t))$, respectively.

Proof. Let us prove that fact (i) implies, and is implied by, fact (ii). Let $\{Y_1, \ldots, Y_{n-1}\}$ be a family of vector fields that forms a basis for each tangent space in an open set $\mathcal{U} \subset \mathsf{B}$. We write $\gamma'(t) = w^k(t)\mathrm{hlft}(Y_k)(\gamma(t)) + v(t)X(\gamma(t))$ for some functions $w^k \colon I \to \mathbb{R}$, $k \in \{1, \ldots, n-1\}$, and compute

$$\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = \dot{w}^k(t)\mathrm{hlft}(Y_k)(\gamma(t)) + \dot{v}(t)X(\gamma(t))$$
$$+ w^k(t)\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\mathrm{hlft}(Y_k)(\gamma(t)) + v(t)\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}X(\gamma(t)).$$

The equality $\eta = \pi_{\mathsf{B}} \circ \gamma$ implies that $\eta'(t) = \mathsf{T}_{\gamma(t)}\pi_{\mathsf{B}}(\gamma'(t))$, which in coordinates reads $\dot{\eta}^k(t) = w^k(t)$, $k \in \{1, \ldots, n-1\}$. Now we project the previous equation onto TB to obtain

$$T_{\gamma(t)}\pi_{\mathsf{B}}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)\right) = \ddot{\eta}^k(t)Y_k(t) + \dot{\eta}^k(t)T_{\gamma(t)}\pi_{\mathsf{B}}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\mathrm{hlft}(Y_k)(\gamma(t))\right)$$
$$+ v(t)T_{\gamma(t)}\pi_{\mathsf{B}}\left(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}X(\gamma(t))\right)$$
$$= \ddot{\eta}^k(t)Y_k(\gamma(t))$$
$$+ \dot{\eta}^k(t)T_{\gamma(t)}\pi_{\mathsf{B}}\left(\dot{\eta}^j(t)\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y_j)}\mathrm{hlft}(Y_k)(\gamma(t)) + v(t)\overset{\mathbb{G}}{\nabla}_X\mathrm{hlft}(Y_k)(\gamma(t))\right)$$
$$+ v(t)T_{\gamma(t)}\pi_{\mathsf{B}}\left(\dot{\eta}^j(t)\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y_j)}X(\gamma(t)) + v(t)\overset{\mathbb{G}}{\nabla}_X X(\gamma(t))\right)$$
$$= \left(\ddot{\eta}^k(t)Y_k(t) + \dot{\eta}^k(t)\dot{\eta}^j(t)\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{Y_j}Y_k(\eta(t))\right)$$
$$+ 2v(t)\dot{\eta}^k(t)T_{\gamma(t)}\pi_{\mathsf{B}}\left(\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y_k)}X(\gamma(t))\right) + v^2(t)T_{\gamma(t)}\pi_{\mathsf{B}}\left(\overset{\mathbb{G}}{\nabla}_X X(\gamma(t))\right),$$

where we have used Lemma 3.2 and the X-invariance of $\mathrm{hlft}(Y_k)$, $k \in \{1, \ldots, n-1\}$, i.e., $\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(Y_k)}X = \overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}(X)}Y_k$. From Proposition 2.1, equation (5), and the definition of C_X, we obtain

$$T_{\gamma(t)}\pi_{\mathsf{B}}(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t)) =$$
$$= \overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{\eta'(t)}\eta'(t) - v(t)C_X(\eta'(t)) - v^2(t)\mathrm{grad}_{\mathsf{B}}\left(\tfrac{1}{2}\|X\|_{\mathbb{G}}^2\right)_{\mathsf{B}}(\eta(t)).$$

Next, note that $\mu = J_X \circ \gamma' = v(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}} \circ \eta$. Therefore, the vertical component of γ' satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu(t) = \dot{v}(t)(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t)) + v(t)\langle \boldsymbol{d}(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t)); \eta'(t)\rangle.$$

These computations show the equivalence of statements (i) and (ii). The equivalence between statements (ii) and (iii) follows, after some bookkeeping, from the equality

$$\operatorname{grad}_{\mathsf{B}}\big(\,\|X\|_{\mathbb{G}}^{2}\,\big)_{\mathsf{B}} = -\big(\,\|X\|_{\mathbb{G}}^{4}\,\big)_{\mathsf{B}}\operatorname{grad}_{\mathsf{B}}\big(\,\|X\|_{\mathbb{G}}^{-2}\,\big)_{\mathsf{B}}.$$

$\square$

Remark 3.4 (Forced simple mechanical systems). *It is immediate to extend the results in the theorem to the setting of forced simple mechanical systems. The uncontrolled external force F will in general appear in both the horizontal and the vertical equations, so that the equations in Theorem 3.3(ii) read*

$$\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{\eta'(t)}\eta'(t) =$$
$$= -\operatorname{grad}_{\mathsf{B}}\big(V_{X,v(t)}^{\mathrm{eff}}\big)_{\mathsf{B}}(\eta(t)) + v(t)C_X(\eta'(t)) + T\pi_{\mathsf{B}}\big(\mathbb{G}^{\sharp}(F(t,\gamma'(t)))\big),$$
$$\dot{v}(t) =$$
$$= -\frac{v(t)\langle \boldsymbol{d}(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(\eta(t));\eta'(t)\rangle}{(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(\eta(t))} + \frac{1}{(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(\eta(t))}\,\langle F(t,\gamma'(t));X(\gamma(t))\rangle.$$

4. Some tangent bundle geometry

In the preceding section, we developed the Riemannian geometry of reduction by an infinitesimal symmetry. In the section following this one, we shall linearize about a relative equilibrium. To develop this linearization theory, we need a few ideas concerning the geometry of tangent bundles. The main idea is the use of Ehresmann connections to develop an explicit link between the standard theory of linearization and the Jacobi equation of geodesic variation. Much of what we say here can be found in the book of Yano and Ishihara [1973].

4.1. Tangent lifts of vector fields.

Let X be a vector field on a manifold M. The ***tangent lift*** of X is the vector field X^T on TM defined by

$$X^T(v_x) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\big(T_x\Phi_{0,t}^X(v_x)\big).$$

If X is time-dependent, then its tangent lift is defined by $X^T(t,x) = X_t^T(x)$, where X_t is the vector field on M defined by $X_t(x) = X(t,x)$. One may verify in coordinates that

$$X^T = X^i \frac{\partial}{\partial x^i} + \frac{\partial X^i}{\partial x^j}v^j \frac{\partial}{\partial v^i}. \tag{6}$$

From this coordinate expression, we may immediately assert a few useful facts.

Remarks 4.1 (Properties of the tangent lift). (1) X^T *is a linear vector field on* TM *over* X.

(2) *Since* X^T *is* π_{TM}*-related to* X, *if* $t \mapsto \Upsilon(t)$ *is an integral curve for* X^T, *then this curve projects to an integral curve for* X. *Thus integral curves for* X^T *may be thought of as vector fields along integral curves for* X.

(3) *Let* $x \in \mathsf{M}$ *and let* γ *be the integral curve for* X *with initial condition* x *at time* $t = a$. *Let* $v_{1,x}, v_{2,x} \in \mathsf{T}_x\mathsf{M}$ *with* Υ_1 *and* Υ_2 *the integral curves for* X^T *with initial conditions* $v_{1,x}$ *and* $v_{2,x}$, *respectively, at time* $t = a$. *Then* $t \mapsto \alpha_1 \Upsilon_1(t) + \alpha_2 \Upsilon_2(t)$ *is the integral curve for* X^T *with initial condition* $\alpha_1 v_{1,x} + \alpha_2 v_{2,x}$, *for* $\alpha_1, \alpha_2 \in \mathbb{R}$. *That is to say, the family of integral curves for* X^T *that project to* γ *is a* $\dim(\mathsf{M})$*-dimensional vector space.*

(4) *One may think of* X^T *as the "linearization" of* X *in the following sense. Let* $\gamma\colon I \to \mathsf{M}$ *be the integral curve of* X *through* $x \in \mathsf{M}$ *at time* $t = a$, *and let* $\Upsilon\colon I \to \mathsf{TM}$ *be the integral curve of* X^T *with initial condition* $v_x \in \mathsf{T}_x\mathsf{M}$ *at time* $t = a$. *Choose a variation* $\sigma\colon I \times J \to$ *of* γ *with the following properties:*

   (a) J *is an interval for which* $0 \in \mathrm{int}(J)$;

   (b) $s \mapsto \sigma(t, s)$ *is differentiable for* $t \in I$;

   (c) *for* $s \in J$, $t \mapsto \sigma(t, s)$ *is the integral curve of* X *through* $\sigma(a, s)$ *at time* $t = a$;

   (d) $\sigma(t, 0) = \gamma(t)$ *for* $t \in I$;

   (e) $v_x = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}\sigma(a, s)$.

We then have $\Upsilon(t) = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}\sigma(t, s)$. *Thus* $X^T(v_x)$ *measures the "variation" of solutions of* X *when perturbed by initial conditions lying in the direction of* v_x. *In cases where* M *has additional structure, as we shall see, we can make more precise statements about the meaning of* X^T. $\bullet$

4.2. **Two Ehresmann connections associated to an affine connection.** First let us recall the notion of an Ehresmann connection on a fiber bundle $\pi\colon \mathsf{M} \to \mathsf{B}$. On such a fiber bundle, the ***vertical subbundle*** is given by $\mathsf{VM} = \ker(T\pi)$. An ***Ehresmann connection*** is then a subbundle HM that is complementary to VM, i.e., $\mathsf{TM} = \mathsf{HM} \oplus \mathsf{VM}$.

Now let ∇ be an affine connection on Q. If S is the geodesic spray defined by an affine connection ∇ on Q, then there is a natural Ehresmann connection HTQ on $\pi_{\mathsf{TQ}}\colon \mathsf{TQ} \to \mathsf{Q}$. This can be described in several ways, and we refer to [Yano and Ishihara 1973] for some of these. For our purposes, it suffices to write a basis for HTQ in local coordinates:

$$\mathrm{hlft}\Big(\frac{\partial}{\partial q^i}\Big) = \frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma_{ik}^j + \Gamma_{ki}^j)v^k \frac{\partial}{\partial v^j}, \qquad i \in \{1, \ldots, n\},$$

where Γ^i_{jk}, $i, j, k \in \{1, \ldots, n\}$ are the Christoffel symbols for ∇. This defines "hlft" as the horizontal lift map for the connection we describe here. (Note that we make an abuse of notation here by using "hlft" both for the horizontal lift on $\pi_{\mathsf{B}} \colon \mathsf{Q} \to \mathsf{B}$ and on $\pi_{\mathsf{TQ}} \colon \mathsf{TQ} \to \mathsf{Q}$.) We can also define the **vertical lift** map by

$$\mathrm{vlft}\left(\frac{\partial}{\partial q^i}\right) = \frac{\partial}{\partial v^i}.$$

This definition can be written intrinsically as $\mathrm{vlft}_{v_q}(w_q) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}(v_q + tw_q)$, defining an isomorphism from $\mathsf{T}_q\mathsf{Q}$ to $\mathsf{V}_{v_q}\mathsf{TQ}$. Note that we have $S(v_q) = \mathrm{hlft}_{v_q}(v_q)$. Also note that this splitting $\mathsf{T}_{v_q}\mathsf{TQ} \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}$ extends the natural splitting $\mathsf{T}_{0_q}\mathsf{TQ} \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}$ that one has on the zero section away from the zero section. We depict the situation in Figure 2 to give the reader
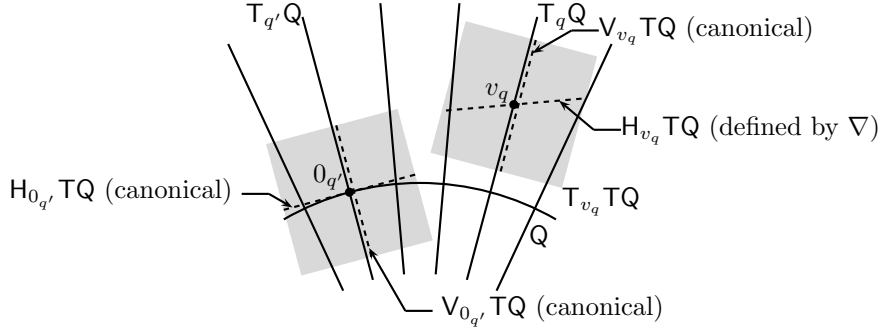
FIGURE 2. A depiction of the Ehresmann connection on $\pi_{\mathsf{TQ}} \colon \mathsf{TQ} \to \mathsf{Q}$ associated with an affine connection on Q

some intuition for what is going on.

In this section we "lift" the preceding construction to construct an Ehresmann connection on $\pi_{\mathsf{TTQ}} \colon \mathsf{TTQ} \to \mathsf{TQ}$. This requires an affine connection on TQ. It turns out that there are various ways of lifting an affine connection on Q to its tangent bundle, and the one suited to our purposes is defined as follows [Yano and Ishihara 1973].

Lemma 4.2. *If ∇ is an affine connection on Q, then there exists a unique affine connection ∇^T on TQ satisfying*

$$\nabla^T_{X^T} Y^T = (\nabla_X Y)^T$$

*for vector fields X and Y on Q. This affine connection is the **tangent lift** of ∇.*

Now we may construct the Ehresmann connection on $\pi_{\mathsf{TTQ}}\colon \mathsf{TTQ} \to \mathsf{TQ}$ using the affine connection ∇^T. Let us denote this connection by $\mathsf{H}(\mathsf{TTQ})$. Note that this connection provides a splitting

$$\mathsf{T}_{X_{v_q}}\mathsf{TTQ} \simeq \mathsf{T}_{v_q}\mathsf{TQ} \oplus \mathsf{T}_{v_q}\mathsf{TQ} \tag{7}$$

for $X_{v_q} \in \mathsf{T}_{v_q}\mathsf{TQ}$. Also, the Ehresmann connection HTQ on $\pi_{\mathsf{TQ}}\colon \mathsf{TQ} \to \mathsf{Q}$ described above gives a splitting $\mathsf{T}_{v_q}\mathsf{TQ} \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}$. Therefore, we have the resulting splitting

$$\mathsf{T}_{X_{v_q}}\mathsf{TTQ} \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}. \tag{8}$$

In this splitting, the first two components are the horizontal subspace and the second two components are the vertical subspace. Within each pair, the first part is horizontal and the second is vertical.

Let us now write a basis of vector fields on TTQ that is adapted to the splitting (8). To obtain a coordinate expression for the Ehresmann connection on $\pi_{\mathsf{TTQ}}\colon \mathsf{TTQ} \to \mathsf{TQ}$, we use the coordinate expression for ∇^T. We can then write a basis that is adapted to the splitting of $\mathsf{T}_{v_q}\mathsf{TQ}$. Let us skip the messy intermediate computations, and simply present the local bases since these are all we shall need. The resulting basis vector field for $\mathsf{H}(\mathsf{TTQ})$ are

$$\mathrm{hlft}^T\Big(\frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma_{ik}^j + \Gamma_{ki}^j)v^k\frac{\partial}{\partial v^j}\Big) = \frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma_{ik}^j + \Gamma_{ki}^j)v^k\frac{\partial}{\partial v^j}$$

$$- \frac{1}{2}(\Gamma_{ik}^j + \Gamma_{ki}^j)u^k\frac{\partial}{\partial u^j} - \frac{1}{2}\Big(\frac{\partial\Gamma_{i\ell}^j}{\partial q^k}u^\ell v^k + \frac{\partial\Gamma_{\ell i}^j}{\partial q^k}u^\ell v^k + (\Gamma_{ik}^j + \Gamma_{ki}^j)w^k$$

$$- \frac{1}{2}(\Gamma_{i\ell}^k + \Gamma_{\ell i}^k)(\Gamma_{km}^j + \Gamma_{mk}^j)u^m v^\ell\Big)\frac{\partial}{\partial w^j}, \qquad i \in \{1,\dots,n\},$$

$$\mathrm{hlft}^T\Big(\frac{\partial}{\partial v^i}\Big) = \frac{\partial}{\partial v^i} - \frac{1}{2}(\Gamma_{ik}^j + \Gamma_{ki}^j)u^k\frac{\partial}{\partial w^j}, \qquad i \in \{1,\dots,n\},$$

with the first n basis vectors forming a basis for the horizontal part of $\mathsf{H}_{X_{v_q}}(\mathsf{TTQ})$, and the second n vectors forming a basis for the vertical part of $\mathsf{H}_{X_{v_q}}(\mathsf{TTQ})$, with respect to the splitting $\mathsf{H}_{X_{v_q}}(\mathsf{TTQ}) \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}$. Note that we use the notation hlft^T to refer to the horizontal lift for the connection on $\pi_{\mathsf{TTQ}}\colon \mathsf{TTQ} \to \mathsf{TQ}$. We also denote by vlft^T the vertical lift on this vector bundle.

We may easily derive a basis for the vertical subbundle of $\pi_{\mathsf{TTQ}}\colon \mathsf{TTQ} \to \mathsf{TQ}$ that adapts to the splitting of $\mathsf{T}_{v_q}\mathsf{TQ} \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}$. We may verify

that the vector fields

$$\mathrm{vlft}^T\Big(\frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})v^k\frac{\partial}{\partial v^j}\Big) = \frac{\partial}{\partial u^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})v^k\frac{\partial}{\partial w^j},$$
$$i \in \{1,\dots,n\},$$

$$\mathrm{vlft}^T\Big(\frac{\partial}{\partial v^i}\Big) = \frac{\partial}{\partial w^i}, \qquad i \in \{1,\dots,n\},$$

have the property that the first n vectors span the horizontal part of $\mathsf{V}_{X_{v_q}}\mathsf{TTQ}$, and the second n span the vertical part of $\mathsf{V}_{X_{v_q}}\mathsf{TTQ}$.

4.3. The Jacobi equation and the tangent lift of the geodesic spray.
As we have stated several times already, we will be looking at control-affine systems whose drift vector field is the geodesic spray S for an affine connection. One way to frame the objective of this section is to think about how one might represent S^T in terms of objects defined on Q, even though S^T is itself a vector field on TTQ. That this ought to be possible seems reasonable as all the information used to describe S^T is contained in the affine connection ∇ on Q, along with some canonical tangent bundle geometry. It turns out that it *is* possible to essentially represent S^T on Q, but to do so requires some effort. What is more, it is perhaps not immediately obvious how one should proceed.

To understand the meaning of S^T, consider the following construction. Let $\gamma\colon I \to \mathsf{Q}$ be a geodesic for the affine connection ∇. Let $\sigma\colon I \times J \to \mathsf{Q}$ be a variation of γ. Thus

(1) J is an interval for which $0 \in \mathrm{int}(J)$,
(2) $s \mapsto \sigma(t,s)$ is differentiable for $t \in I$,
(3) for $s \in J$, $t \mapsto \sigma(t,s)$ is a geodesic of ∇, and
(4) $\sigma(t,0) = \gamma(t)$ for $t \in I$.

If one defines $\xi(t) = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}\sigma(t,s)$, then it can be shown (see Theorem 1.2 in Chapter VIII of volume 2 of [Kobayashi and Nomizu 1963]) that ξ satisfies the **Jacobi equation**:

$$\nabla^2_{\gamma'(t)}\xi(t) + R(\xi(t),\gamma'(t))\gamma'(t) + \nabla_{\gamma'(t)}(T(\xi(t),\gamma'(t))) = 0,$$

where T is the torsion tensor and R is the curvature tensor for ∇. Thus the Jacobi equation tells us how geodesics vary along γ as we vary their initial conditions.

With this and Remark 4.1–4 as backdrop, we expect there to indeed be a concrete relationship between S^T and the Jacobi equation. This relationship involves the Ehresmann connections on $\pi_{\mathsf{TQ}}\colon \mathsf{TQ} \to \mathsf{Q}$ and $\pi_{\mathsf{TTQ}}\colon \mathsf{TTQ} \to \mathsf{TQ}$ presented in the preceding section.

First recall that the connection $\mathsf{H}(\mathsf{TTQ})$ on $\pi_{\mathsf{TTQ}}\colon \mathsf{TTQ} \to \mathsf{TQ}$ and the connection HTQ on $\pi_{\mathsf{TQ}}\colon \mathsf{TQ} \to \mathsf{Q}$ combine to give a splitting

$$\mathsf{T}_{X_{v_q}}\mathsf{TTQ} \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q},$$

where $X_{v_q} \in \mathsf{T}_{v_q}\mathsf{TQ}$. Here we maintain our convention that the first two components refer to the horizontal component for a connection $\mathsf{H}(\mathsf{TTQ})$ on $\pi_{\mathsf{TTQ}}\colon \mathsf{TTQ} \to \mathsf{TQ}$, and the second two components refer to the vertical component. Using the splitting (7), let us write $X_{v_q} \in \mathsf{T}_{v_q}\mathsf{TQ}$ as $u_{v_q} \oplus w_{v_q}$ for some $u_{v_q}, w_{v_q} \in \mathsf{T}_q\mathsf{Q}$. Note that we depart from the usual notation of writing tangent vectors in $\mathsf{T}_q\mathsf{Q}$ with a subscript of q, instead using the subscript v_q. This abuse of notation is necessary (and convenient) to reflect the fact that these vectors depend on where we are in TQ, and not just in Q.

We may use this representation of S^T to obtain a refined relationship between solutions of the Jacobi equation and integral curves of S^T. To do so, we first prove a simple lemma. We state a more general form of this lemma than we shall immediately use, but the extra generality will be useful in Section 5.

Lemma 4.3. *Let Y be a time-dependent vector field on Q, suppose that $\gamma\colon I \to \mathsf{Q}$ is the* LAD *curve satisfying $\nabla_{\gamma'(t)}\gamma'(t) = Y(t, \gamma(t))$, and denote by $\Upsilon\colon I \to \mathsf{TQ}$ the tangent vector field of γ (i.e., $\Upsilon = \gamma'$). Let $X\colon I \to \mathsf{TTQ}$ be an* LAC *vector field along Υ, and denote $X(t) = X_1(t) \oplus X_2(t) \in \mathsf{T}_{\gamma(t)}\mathsf{Q} \oplus \mathsf{T}_{\gamma(t)}\mathsf{Q} \simeq \mathsf{T}_{\Upsilon(t)}\mathsf{TQ}$. Then the tangent vector field to the curve $t \mapsto X(t)$ is given by $\gamma'(t) \oplus Y(t, \gamma(t)) \oplus \tilde{X}_1(t) \oplus \tilde{X}_2(t)$, where*

$$\tilde{X}_1(t) = \nabla_{\gamma'(t)}X_1(t) + \tfrac{1}{2}T(X_1(t), \gamma'(t)),$$
$$\tilde{X}_2(t) = \nabla_{\gamma'(t)}X_2(t) + \tfrac{1}{2}T(X_2(t), \gamma'(t)).$$

Proof. In coordinates, the curve $t \mapsto X(t)$ has the form

$$(q^i(t), \dot{q}^j(t), X_1^k(t), X_2^\ell(t) - \tfrac{1}{2}(\Gamma_{mr}^\ell + \Gamma_{rm}^\ell)\dot{q}^m(t)X_1^r(t)).$$

The tangent vector to this curve is then given a.e. by

$$\dot{q}^i\frac{\partial}{\partial q^i} + (Y^i - \Gamma_{jk}^i\dot{q}^j\dot{q}^k)\frac{\partial}{\partial v^i} + \dot{X}_1^i\frac{\partial}{\partial u^i} + \Big(\dot{X}_2^i - \frac{1}{2}\frac{\partial\Gamma_{jk}^i}{\partial q^\ell}\dot{q}^k\dot{q}^\ell X_1^j - \frac{1}{2}\frac{\partial\Gamma_{kj}^i}{\partial q^\ell}\dot{q}^k\dot{q}^\ell X_1^j$$
$$- \frac{1}{2}(\Gamma_{jk}^i + \Gamma_{kj}^i)(Y^k - \Gamma_{\ell m}^k\dot{q}^\ell\dot{q}^m)X_1^j - \frac{1}{2}(\Gamma_{jk}^i + \Gamma_{kj}^i)\dot{q}^k\dot{X}_1^j\Big)\frac{\partial}{\partial w^i}.$$

A straightforward computation shows that this tangent vector field has the representation

$$\gamma'(t) \oplus Y(t, \gamma(t)) \oplus \big(\nabla_{\gamma'(t)}X_1(t) + \tfrac{1}{2}T(X_1(t), \gamma'(t))\big)\oplus$$
$$\oplus\big(\nabla_{\gamma'(t)}X_2(t) + \tfrac{1}{2}T(X_2(t), \gamma'(t))\big),$$

which proves the lemma. □ □

We may now prove our main result that relates the integral curves of S^T with solutions to the Jacobi equation.

Theorem 4.4 (Tangent lift of $\boldsymbol{S}$ and Jacobi equation). *Let ∇ be an affine connection on Q with S the corresponding geodesic spray. Let $\gamma\colon I \to \mathsf{Q}$ be a geodesic with $t \mapsto \Upsilon(t) \triangleq \gamma'(t)$ the corresponding integral curve of S. Let $a \in I$, $u, w \in \mathsf{T}_{\gamma(a)}\mathsf{Q}$, and define vector fields $U, W\colon I \to \mathsf{T}\mathsf{Q}$ along γ by asking that $t \mapsto U(t) \oplus W(t) \in \mathsf{T}_{\gamma(t)}\mathsf{Q} \oplus \mathsf{T}_{\gamma(t)}\mathsf{Q} \simeq \mathsf{T}_{\Upsilon(t)}\mathsf{T}\mathsf{Q}$ be the integral curve of S^T with initial conditions $u \oplus w \in \mathsf{T}_{\gamma(a)}\mathsf{Q} \oplus \mathsf{T}_{\gamma(a)}\mathsf{Q} \simeq \mathsf{T}_{\Upsilon(a)}\mathsf{T}\mathsf{Q}$. Then U and W have the following properties:*

(i) U satisfies the Jacobi equation

$$\nabla^2_{\gamma'(t)}U(t) + R(U(t), \gamma'(t))\gamma'(t) + \nabla_{\gamma'(t)}(T(U(t), \gamma'(t))) = 0;$$

(ii) $W(t) = \nabla_{\gamma'(t)}U(t) + \frac{1}{2}T(U(t), \gamma'(t))$.

Proof. This follows most directly, although very messily, from a coordinate computation using Lemma 4.3. We refer the reader to [Bullo and Lewis 2005]. □

4.4. **The Sasaki metric.** When the constructions of this section are applied in the case when ∇ is the Levi-Civita affine connection associated with a Riemannian metric $\mathbb{G}$ on Q, there is an important additional construction that can be made.

Definition 4.5 (Sasaki metric). *Let $(\mathsf{Q}, \mathbb{G})$ be a Riemannian manifold, and for $v_q \in \mathsf{T}\mathsf{Q}$, let $\mathsf{T}_{v_q}\mathsf{T}\mathsf{Q} \simeq \mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}$ denote the splitting defined by the affine connection $\overset{\mathbb{G}}{\nabla}$. The **Sasaki metric** is the Riemannian metric $\mathbb{G}^T$ on $\mathsf{T}\mathsf{Q}$ given by*

$$\mathbb{G}^T(u^1_{v_q} \oplus w^1_{v_q}, u^2_{v_q} \oplus w^2_{v_q}) = \mathbb{G}(u^1_{v_q}, u^2_{v_q}) + \mathbb{G}(w^1_{v_q}, w^2_{v_q}). \qquad \bullet$$

The Sasaki metric was introduced by Sasaki [1958, 1962]. Much research has been made into the properties of the Sasaki metric, beginning with the work of Sasaki who studied the curvature, geodesics, and Killing vector fields of the metric. Some of these results are also given in [Yano and Ishihara 1973].

5. Linearization along relative equilibria

This long section contains many of the essential results in the paper. Our end objective is to relate the linearization of equilibria of the reduced equations of Theorem 3.3 to the linearization along the associated *relative* equilibria of the unreduced system. We aim to perform this linearization in

a control theoretic setting so that our constructions will be useful not just for investigations of stability, but also for stabilization. To properly understand this process, we begin in Section 5.1 with linearization of general control-affine systems about general controlled trajectories. Then we specialize this discussion in Section 5.2 to the linearization of a general affine connection control system about a general controlled trajectory. We then, in Section 5.3, finally specialize to the case of interest, namely the linearization of the unreduced equations along a relative equilibrium. Here the main result is Theorem 5.10 which gives the geometry associated with linearization along a relative equilibrium. In particular this result, or more precisely its proof, makes explicit the relationship between the affine differential geometric concepts arising in the linearization of Section 5.2 and the usual concepts arising in reduction of mechanical systems, such as the curvature of the mechanical connection and the effective potential. Then, in Section 5.4 we turn to the linearization, in the standard sense, of the reduced equations about an equilibrium. The main result here is Theorem 5.13 which links the reduced and unreduced linearizations.

We let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ be a forced simple mechanical control system with F time-independent, let X be a complete infinitesimal symmetry of Σ, and let $\chi \colon \mathbb{R} \to \mathsf{Q}$ be a regular relative equilibrium for $(\mathsf{Q}, \mathbb{G}, V, F)$. Thus χ is an integral curve for X that is also an uncontrolled trajectory for the system. We let B denote the set of X-orbits, and following Assumption 3.1, we assume that B is a smooth manifold for which $\pi_{\mathsf{B}} \colon \mathsf{Q} \to \mathsf{B}$ is a surjective submersion. We let $Y_a = \mathbb{G}^\sharp \circ F^a$, $a \in \{1, \ldots, m\}$, and if $u \colon I \to \mathbb{R}^m$ is a locally integrable control, we denote

$$Y_u(t, q) = \sum_{a=1}^m u^a(t) Y_a(q), \qquad t \in I, \ q \in \mathsf{Q}, \tag{9}$$

for brevity. Define vector fields $Y_{\mathsf{B},a}$, $a \in \{1, \ldots, m\}$, by $Y_{\mathsf{B},a}(b) = T_q \pi_{\mathsf{B}}(Y_a(q))$ for $q \in \pi_{\mathsf{B}}^{-1}(b)$. Since the vector fields $\mathscr{Y}$ are X-invariant, this definition is independent of $q \in \pi_{\mathsf{B}}^{-1}(b)$. Similarly to (9), we denote

$$Y_{\mathsf{B},u}(t, b) = \sum_{a=1}^m u^a(t) Y_{\mathsf{B},a}(b).$$

In Theorem 3.3 we showed that the reduced system has $\mathsf{TB} \times \mathbb{R}$ as its state space, and satisfies the equations

$$\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{\eta'(t)}\eta'(t) = -\operatorname{grad}_{\mathsf{B}}\left(V^{\mathrm{eff}}_{X,v(t)}\right)_{\mathsf{B}}(\eta(t)) + v(t)C_X(\eta'(t))$$
$$+ T\pi_{\mathsf{B}} \circ \mathbb{G}^\sharp \circ F(\gamma'(t)) + Y_{\mathsf{B},u}(t,\eta(t))),$$
$$\dot{v}(t) = -\frac{v(t)\langle \boldsymbol{d}(\|X\|^2_{\mathbb{G}})_{\mathsf{B}}(\eta(t)); \eta'(t)\rangle}{(\|X\|^2_{\mathbb{G}})_{\mathsf{B}}(\eta(t))} + \frac{\langle F(\gamma'(t)); X(\gamma(t))\rangle}{(\|X\|^2_{\mathbb{G}})_{\mathsf{B}}(\eta(t))} \quad (10)$$
$$+ \frac{\mathbb{G}(Y_u(t,\gamma(t)), X(\gamma(t)))}{(\|X\|^2_{\mathbb{G}})_{\mathsf{B}}(\eta(t))},$$

where (γ, u) is the controlled trajectory, $\eta = \pi_{\mathsf{B}} \circ \gamma$, and v is defined by $\operatorname{ver}(\gamma'(t)) = v(t)X(\gamma(t))$. The relative equilibrium χ corresponds to the equilibrium point $(T\pi_{\mathsf{B}}(\chi'(0)), 1)$ of the reduced equations (10). Therefore, linearization of the relative equilibrium χ could be *defined* to be the linearization of the equations (10) about the equilibrium point $(T\pi_{\mathsf{B}}(\chi'(0)), 1)$. This is one view of linearization of relative equilibria. Another view is that, since χ is a trajectory for the unreduced system, we could linearize along it in the manner described when describing the Jacobi equation. In this section we shall see how these views of linearization of relative equilibria tie together. We build up to this by first considering linearization in more general settings.

5.1. **Linearization of a control-affine system along a controlled trajectory.** In order to talk about linearization along a relative equilibrium, we first discuss linearization along a general controlled trajectory. In order to do this, it is convenient to first consider the general control-affine case, then specialize to the mechanical setting.

Let us first recall that a ***control-affine system*** is a pair $(\mathsf{M}, \mathscr{C} = \{f_0, f_1, \ldots, f_m\})$ where M is a manifold and $f_0, f_1, \ldots, f_m$ are vector fields on M. The ***drift vector field*** is f_0 and the ***control vector fields*** are $f_1, \ldots, f_m$. The governing equations for a time-dependent control-affine system $(\mathsf{M}, \mathscr{C})$ are then

$$\gamma'(t) = f_0(\gamma(t)) + \sum_{a=1}^{m} u^a(t)f_a(\gamma(t)),$$

for a ***controlled trajectory*** (γ, u). We shall also consider the notion of a ***time-dependent control-affine system***, by which we mean a control-affine system where the vector fields $f_0, f_1, \ldots, f_m$ are time-dependent.

We shall also require the standard notion of a ***linear control system***, by which we mean a triple (V, A, B), where V is a finite-dimensional $\mathbb{R}$-vector space, $A \in L(\mathsf{V}; \mathsf{V})$, and $B \in L(\mathbb{R}^m; \mathsf{V})$. The equations governing a

linear control system are

$$\dot{x}(t) = A(x(t)) + B(u(t)).$$

Suppose that we have a (time-independent) control-affine system $(\mathsf{M}, \mathscr{C} = \{f_0, f_1, \ldots, f_m\})$, and a controlled trajectory (γ_0, u_0) defined on an interval I. We wish to linearize the system about this controlled trajectory. Linearization is to be done with respect to both state and control. Thus, speaking somewhat loosely for a moment, to compute the linearization, one should first fix the control at u_0 and linearize with respect to state, then fix the state and linearize with respect to control, and then add the results to obtain the linearization. Let us now be more formal about this.

If we fix the control at u_0, we obtain the time-dependent vector field f_{u_0} on M defined by

$$f_{u_0}(t, x) = f_0(x) + \sum_{a=1}^{m} u_0^a(t) f_a(x).$$

We call f_{u_0} the **reference vector field** for the controlled trajectory (γ_0, u_0). The linearization of the reference vector field is exactly described by its tangent lift, as discussed in Remark 4.1–4. Thus one component of the linearization is $f_{u_0}^T$. The other component is computed by fixing the state, say at x, and linearizing with respect to the control. Thus we consider the map

$$\mathbb{R} \times \mathbb{R}^m \ni (t, u) \mapsto f_0(x) + \sum_{a=1}^{m} (u_0^a(t) + u^a) f_a(x) \in \mathsf{T}_x \mathsf{M},$$

and differentiate this with respect to u at $u = \mathbf{0}$. The resulting map from $\mathsf{T}_\mathbf{0} \mathbb{R}^m \simeq \mathbb{R}^m$ to $\mathsf{T}_{f_{u_0}(t,x)}(\mathsf{T}_x \mathsf{M}) \simeq \mathsf{T}_x \mathsf{M}$ is simply given by

$$u \mapsto \sum_{a=1}^{m} u^a f_a(x).$$

In order to add the results of the two computations, we regard $\mathsf{T}_x \mathsf{M}$ as being identified with $\mathsf{V}_{f_{u_0}(t,x)} \mathsf{T} \mathsf{M}$. Thus the linearization with respect to the control yields the linearized control vector fields $\mathrm{vlft}(f_a)$, $a \in \{1, \ldots, m\}$. In this way, we arrive at the time-dependent control-affine system $\Sigma^T(\gamma_0, u_0) = (\mathsf{T} \mathsf{M}, \{f_{u_0}^T, \mathrm{vlft}(f_1), \ldots, \mathrm{vlft}(f_a)\})$, whose controlled trajectories (ξ, u) satisfy

$$\xi'(t) = f_{u_0}^T(t, \xi(t)) + \sum_{a=1}^{m} u^a(t) \mathrm{vlft}(f_a)(\xi(t)). \tag{11}$$

The following result gives an important property of these controlled trajectories.

Lemma 5.1. *For every locally integrable control $t \mapsto u(t)$, the time-dependent vector field*

$$(t, v_x) \mapsto f_{u_0}^T(t, v_x) + \sum_{a=1}^{m} u^a(t)\mathrm{vlft}(f_a)(v_x)$$

is a linear vector field over f_{u_0}.

Proof. This is easily proved in coordinates. $\qquad\qquad\square$

From Remark 4.1–1, we then know that, if (ξ, u) is a controlled trajectory for $\Sigma^T(\gamma_0, u_0)$, then $\pi_{\mathsf{TM}} \circ \xi$ is an integral curve for f_{u_0}. In particular, if (ξ, u) is a controlled trajectory for $\Sigma^T(\gamma_0, u_0)$ that satisfies $\pi_{\mathsf{TM}} \circ \xi(t) = \gamma_0(t)$ for some $t \in I$, then ξ is a vector field along γ_0.

To formally define the linearization along (γ_0, u_0), we need an additional concept, following Sussmann [1997].

Definition 5.2 (Differential operator along a curve). *Let M be a manifold, let $\gamma\colon I \to \mathsf{M}$ be an LAC curve, and let $\pi\colon \mathsf{E} \to \mathsf{M}$ be a vector bundle. A **differential operator** in E along γ assigns, to each LAC section ξ of E along γ, a locally integrable section $\mathscr{L}(\xi)$ along γ, and the assignment has the property that, if $f \in C^\infty(\mathsf{M})$ and if $\Xi \in \Gamma^\infty(\mathsf{E})$, then*

$$\mathscr{L}(f \circ \gamma(\Xi \circ \gamma))(t) = f \circ \gamma(t)\mathscr{L}(\Xi \circ \gamma)(t) + (\mathscr{L}_{\gamma'(t)}f)(\gamma(t))\Xi \circ \gamma(t). \qquad \bullet$$

Thus a differential operator simply "differentiates" sections of E along γ, with the differentiation rule satisfying the usual derivation property with respect to multiplication with respect to functions. Sussmann [1997] shows that, in coordinates $(x^1, \ldots, x^n)$ for M, if $t \mapsto (\xi^1(t), \ldots, \xi^k(t))$ are the fiber components of the local representative of an LAC section ξ of E, then the fiber components of the local representative of $\mathscr{L}(\xi)$ satisfy

$$(\mathscr{L}(\xi))^a(t) = \dot{\xi}^a(t) + \sum_{b=1}^{k} L_b^a(t)\xi^b(t), \qquad a \in \{1, \ldots, k\},$$

for some locally integrable functions $t \mapsto L_a^b(t)$, $a, b \in \{1, \ldots, k\}$. If $\gamma\colon I \to \mathsf{M}$ is an integral curve of a time-dependent vector field X, then there is a naturally induced differential operator in TM along γ, denoted by $\mathscr{L}^{X,\gamma}$, and defined by

$$\mathscr{L}^{X,\gamma}(\xi) = [X_t, \Xi](\gamma(t)), \qquad \text{a.e. } t \in I,$$

where Ξ is a vector field satisfying $\xi = \Xi \circ \gamma$, and where X_t is the vector field defined by $X_t(x) = X(t, x)$. In coordinates this differential operator satisfies

$$\mathscr{L}^{X,\gamma}(\xi)^i(t) = \dot{\xi}^i(t) - \frac{\partial X^i}{\partial x^j}(\gamma(t))\xi^j(t), \qquad i \in \{1, \ldots, n\}.$$

This differential operator is sometimes referred to as the "Lie drag" (see [Crampin and Pirani 1986, Section 3.5]).

A coordinate computation readily verifies the following result, and we refer to [Lewis and Tyner 2003, Sussmann 1997] for details.

Proposition 5.3 (Tangent lift and differential operator). *Let $X\colon I \times M \to \mathsf{T}M$ be a time-dependent vector field, let $v_{x_0} \in \mathsf{T}_x M$, let $t_0 \in I$, and let $\gamma\colon I \to M$ be the integral curve of X satisfying $\gamma(t_0) = x_0$. For a vector field ξ along γ satisfying $\xi(t_0) = v_{x_0}$, the following statements are equivalent:*

> *(i) ξ is an integral curve for X^T;*
> *(ii) there exists a variation σ of X along γ such that $\frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}\sigma(t,s) = \xi(t)$ for each $t \in I$;*
> *(iii) $\mathscr{L}^{X,\gamma}(\xi) = 0$.*

With the preceding as motivation, we can define the linearization of a control-affine system.

Definition 5.4 (Linearization of a control-affine system). *Let $\Sigma = (\mathsf{M}, \mathscr{C} = \{f_0, f_1, \ldots, f_m\})$ be a control-affine system with (γ_0, u_0) a controlled trajectory. The **linearization** of Σ about (γ_0, u_0) is given by $\{\mathscr{L}_\Sigma(\gamma_0, u_0), b_{\Sigma,1}(\gamma_0, u_0), \ldots, b_{\Sigma,m}(\gamma_0, u_0)\}$, where*

> *(i) $\mathscr{L}_\Sigma(\gamma_0, u_0)$ is the differential operator in $\mathsf{T}\mathsf{M}$ along γ_0 defined by*
> $$\mathscr{L}_\Sigma(\gamma_0, u_0) = \mathscr{L}^{f_{u_0}, \gamma_0},$$
> *and*
> *(ii) $b_{\Sigma,a}$, $a \in \{1, \ldots, m\}$, are the vector fields along γ_0 defined by*
> $$b_{\Sigma,a}(\gamma_0, u_0)(t) = \mathrm{vlft}(f_a(\gamma_0(t))), \qquad a \in \{1, \ldots, m\}. \qquad\bullet$$

The equations governing the linearization are

$$\mathscr{L}_\Sigma(\gamma_0, u_0)(\xi)(t) = \sum_{a=1}^{m} u^a(t) b_{\Sigma,a}(\gamma_0, u_0),$$

which are thus equations for a vector field ξ along γ_0. By Proposition 5.3, these equations are exactly the restriction to image(γ_0) of the equations for the time-dependent control-affine system in (11). In the special case where $f_0(x_0) = 0_{x_0}$, $u_0 = \mathbf{0}$, $\gamma_0 = x_0$ for some $x_0 \in \mathsf{M}$, one can readily check that we recover the linearization of the system at x_0 in the usual sense.

5.2. **Linearization of a forced affine connection control system along a controlled trajectory.** After beginning our discussion of linearization in the context of control-affine systems, we next specialize to affine connection control systems. We let $\Sigma = (\mathsf{Q}, \nabla, Y, \mathscr{Y})$ be a forced

affine connection control system. In this section, we make the following assumption about the external force Y.

Assumption 5.5 (External force for linearization). *Assume that the vector force Y is time-independent and decomposable as $Y(v_q) = \overline{Y}_0(q) + \overline{Y}_1(v_q)$, where $\overline{Y}_0$ is a basic vector force and where $\overline{Y}_1$ is a $(1,1)$-tensor field.* •

This assumption will allow us to model potential forces and Rayleigh dissipative forces, and so makes the development useful for stabilization using PD control as in [Bullo 2000] (see also [Bullo and Lewis 2005]). The governing equations for the system are

$$\nabla_{\gamma'(t)}\gamma'(t) = \overline{Y}_0(\gamma(t)) + \overline{Y}_1(\gamma'(t)) + \sum_{a=1}^{m} u^a(t) Y_a(\gamma(t)).$$

To linearize these equations about *any* controlled trajectory (γ_0, u_0), following the development in the preceding section, we first need to compute the tangent lift for the time-dependent vector field S_{u_0} on TQ defined by

$$S_{u_0}(t, v_q) = S(v_q) + \mathrm{vlft}(\overline{Y}_0(q) + \overline{Y}_1(v_q) + Y_{u_0}(t, q)),$$

where Y_{u_0} is defined as in (9). The Jacobi equation contains the essential features of the tangent lift of S. We recall the notation from Section 4.3 where points in TTQ are written as $u_{v_q} \oplus w_{v_q}$, relative to the splitting defined by the Ehresmann connection on $\pi_{\mathsf{TTQ}} \colon \mathsf{TTQ} \to \mathsf{TQ}$. The following result gives the linearization along (γ_0, u_0) using the Ehresmann connection of Section 4.3.

Proposition 5.6 (State linearization of ACCS). *Let $\Sigma = (\mathsf{Q}, \nabla, Y, \mathscr{Y})$ be a forced simple mechanical control system where Y satisfies Assumption 5.5, let (γ_0, u_0) be a controlled trajectory for Σ defined on I, and let $t \mapsto \Upsilon_0(t) = \gamma_0'(t)$ be the tangent vector field of γ_0. For $a \in I$, let $u, w \in \mathsf{T}_{\gamma_0(a)}\mathsf{Q}$, and define vector fields $U, W \colon I \to \mathsf{TQ}$ along γ_0 by asking that $t \mapsto U(t) \oplus W(t) \in \mathsf{T}_{\gamma_0(t)}\mathsf{Q} \oplus \mathsf{T}_{\gamma_0(t)}\mathsf{Q} \simeq \mathsf{T}_{\Upsilon_0(t)}\mathsf{TQ}$ be the integral curve of $S_{u_0}^T$ with initial conditions $u \oplus w \in \mathsf{T}_{\gamma_0(a)}\mathsf{Q} \oplus \mathsf{T}_{\gamma_0(a)}\mathsf{Q} \simeq \mathsf{T}_{\Upsilon_0(a)}\mathsf{TQ}$. Then U and W satisfy the equations*

$$W(t) = \nabla_{\gamma_0'(t)} U(t) + \tfrac{1}{2} T(U(t), \gamma_0'(t)),$$

$$\nabla^2_{\gamma_0'(t)} U(t) + R(U(t), \gamma_0'(t))\gamma_0'(t) + \nabla_{\gamma_0'(t)}(T(U(t), \gamma_0'(t)))$$
$$= \nabla_{U(t)}(\overline{Y}_0 + Y_{u_0})(\gamma_0(t)) + (\nabla_{U(t)}\overline{Y}_1)(\gamma_0'(t)) + \overline{Y}_1(\nabla_{\gamma_0'(t)} U(t)).$$

Proof. Let us denote $X_{u_0} = \overline{Y}_0 + Y_{u_0}$, for brevity. A computation in coordinates readily shows that the tangent lift of the vertical lift of X_{u_0} is given by

$$\mathrm{vlft}(Y_{u_0})^T(u_{v_q} \oplus w_{v_q}) = 0 \oplus X_{u_0}(q) \oplus 0 \oplus (\nabla X_{u_0}(u_{v_q}) + \tfrac{1}{2} T(X_{u_0}(q), u_{v_q})).$$

A coordinate computation also gives

$$\text{vlft}(\overline{Y}_1)^T(u_{v_q} \oplus w_{v_q}) = 0 \oplus \overline{Y}_1(v_q) \oplus 0$$
$$\oplus \left(\nabla_{u_{v_q}} \overline{Y}_1(v) + \overline{Y}_1(w_{v_q}) + \tfrac{1}{2} T(\overline{Y}_1(v_q), u_{v_q}) + \tfrac{1}{2} \overline{Y}_1(T(u_{v_q}, v_q)) \right).$$

The tangent lift of S is given by Theorem 4.4 as

$$S^T(u_{v_q} \oplus w_{v_q}) = v_q \oplus 0 \oplus w_{v_q} \oplus \left(-R(u_{v_q}, v_q)v_q - \tfrac{1}{2}(\nabla_{v_q} T)(u_{v_q}, v_q) \right.$$
$$\left. + \tfrac{1}{4} T(T(u_{v_q}, v_q), v_q) \right).$$

Thus, using Lemma 4.3, we have that U and W satisfy

$$\nabla_{\gamma_0'(t)} U(t) + \tfrac{1}{2} T(U(t), \gamma_0'(t)) = W(t),$$
$$\nabla_{\gamma_0'(t)} W(t) + \tfrac{1}{2} T(W(t), \gamma_0'(t)) = -R(U(t), \gamma_0'(t))\gamma_0'(t)$$
$$- \tfrac{1}{2}(\nabla_{\gamma_0'(t)} T)(U(t), \gamma_0'(t)) + \tfrac{1}{4} T(T(U(t), \gamma_0'(t)), \gamma_0'(t))$$
$$+ \nabla X_{u_0}(U(t)) + \tfrac{1}{2} T(X_{u_0}(t, \gamma_0(t)), U(t))$$
$$+ \nabla_{U(t)} \overline{Y}_1(\gamma_0'(t)) + \overline{Y}_1(W(t)) + \tfrac{1}{2} T(\overline{Y}_1(\gamma_0'(t)), U(t))$$
$$+ \tfrac{1}{2} \overline{Y}_1(T(U(t), \gamma_0'(t))).$$

The first of the equations is the first equation in the statement of the proposition. Differentiating this first equation, and substituting the second, gives the second equation in the statement of the proposition, after some simplification. $\qquad\square$

Next we linearize with respect to the controls. This is simpler, and, following the procedure in the preceding section, gives the control vector fields $\text{vlft}(\text{vlft}(Y_a))$, $a \in \{1, \ldots, m\}$. Thus, we arrive at the time-dependent control-affine system $\Sigma^T(\gamma_0, u_0) = (\mathsf{TTQ}, \{S_{u_0}^T, \text{vlft}(\text{vlft}(Y_1)), \ldots, \text{vlft}(\text{vlft}(Y_m)))$. With respect to the splitting defined by the Ehresmann connection associated with ∇, it is easy to verify that

$$\text{vlft}(\text{vlft}(Y_a))(u_q \oplus w_q) = 0 \oplus 0 \oplus 0 \oplus Y_a(q).$$

If we write a controlled trajectory for $\Sigma^T(\gamma_0, u_0)$ as $(U \oplus W, u)$, reflecting the notation of Proposition 5.6, we see that the following equations govern this trajectory:

$$W(t) = \nabla_{\gamma_0'(t)} U(t) + \tfrac{1}{2} T(U(t), \gamma_0'(t)),$$
$$\nabla_{\gamma_0'(t)}^2 U(t) + R(U(t), \gamma_0'(t))\gamma_0'(t) + \nabla_{\gamma_0'(t)}(T(U(t), \gamma_0'(t)))$$
$$= \nabla(\overline{Y}_0 + Y_{u_0})(U(t)) + (\nabla_{U(t)} \overline{Y}_1)(\gamma_0'(t)) + \overline{Y}_1(\nabla_{\chi'(t)} U(t))$$
$$+ \sum_{a=1}^{m} u^a(t) Y_a(\gamma_0(t)).$$

With the above as backdrop, we make the following definition, and in so doing, hope the reader will forgive our using the same notation as was used for control-affine systems.

Definition 5.7 (Linearization of ACCS)**.** *Let* $\Sigma = (\mathsf{Q}, \nabla, Y, \mathscr{Y})$ *be a forced affine connection control system where* Y *satisfies Assumption 5.5, and let* (γ_0, u_0) *be a controlled trajectory. The **linearization** of* Σ *about* (γ_0, u_0) *is given by* $\{A_\Sigma(\gamma_0, u_0), b_{\Sigma,1}(\gamma_0, u_0), \ldots, b_{\Sigma,m}(\gamma_0, u_0)\}$, *where*

(i) $A_\Sigma(\gamma_0, u_0)$ *is the differential operator in* TQ *along* γ_0 *defined by*

$$A_\Sigma(\gamma_0, u_0)(t) \cdot \xi(t) = R(\xi(t), \gamma_0'(t))\gamma_0'(t) + \nabla_{\gamma_0'(t)}(T(\xi(t), \gamma_0'(t)))$$
$$- \nabla_{\xi(t)}\overline{Y}_0(\gamma_0(t)) - \nabla_{\xi(t)}Y_{u_0}(t, \gamma_0(t)) + (\nabla_{\xi(t)}\overline{Y}_1)(\gamma_0'(t)) + \overline{Y}_1(\nabla_{\gamma_0'(t)}\xi(t)),$$

and

(ii) $b_{\Sigma,a}(\gamma_0, u_0)$, $a \in \{1, \ldots, m\}$, *are vector fields along* γ_0 *defined by*

$$b_{\Sigma,a}(\gamma_0, u_0)(t) = Y_a(\gamma_0(t)). \qquad \bullet$$

The equations governing the linearization are then

$$\nabla^2_{\gamma_0'(t)}\xi(t) + A_\Sigma(\gamma_0, u_0)(t) \cdot \xi(t) = \sum_{a=1}^{m} u^a(t)Y_a(\gamma_0(t)). \qquad (12)$$

In particular, a **controlled trajectory** for the linearization of Σ along (γ_0, u_0) is a pair (ξ, u), where $u \colon I \to \mathbb{R}^m$ is a locally integrable control, and where $\xi \colon I \to \mathsf{TQ}$ is the LAD curve along γ_0 satisfying (12).

Remarks 5.8. (1) *Note that the structure of the Ehresmann connection induced by* ∇ *allows us to use a differential operator along* γ_0 *rather than along* γ_0'.

(2) *If* ∇ *is torsion-free and if* $\overline{Y}_1 = 0$, *then* $A_\Sigma(\gamma_0, u_0)$ *is no longer a differential operator, but is actually a* $(1,1)$*-tensor field. In such a case, it is still possible to consider this as a differential operator, but one of "order zero."* $\qquad \bullet$

5.3. **Linearization of the unreduced equations along a relative equilibrium.** With the work done in the preceding two sections, it is easy to give the form of the linearization along a relative equilibrium. We let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ be a forced simple mechanical control system. In this and the next section, we make the following assumption about the external force F.

Assumption 5.9 (External force for linearization)**.** *Assume that the force* F *is time-independent and that* $F(v_q) = A^\flat(v_q - X(q))$ *for an* X*-invariant* $(0,2)$*-tensor field* A. $\qquad \bullet$

This assumption will allow the inclusion of Rayleigh dissipative forces. We suppose that X is an infinitesimal symmetry for Σ and that χ is a relative equilibrium. Then, according to Definition 5.7, a pair (ξ, u) is a controlled trajectory for the linearization of Σ along $(\chi, 0)$ if and only if

$$
\overset{\mathbb{G}}{\nabla}^2_{\chi'(t)}\xi(t) + R(\xi(t), \chi'(t))\chi'(t)
$$
$$
= -\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathrm{grad}V)(\chi(t)) - \overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^\sharp \circ A^\flat \circ X)(\chi(t))
$$
$$
+ (\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^\sharp \circ A^\flat))(\chi'(t)) + \mathbb{G}^\sharp \circ A^\flat(\overset{\mathbb{G}}{\nabla}_{\chi'(t)}\xi(t)) + \sum_{a=1}^{m} u^a(t)Y_a(\chi(t)). \quad (13)
$$

In order to facilitate making the connection between the preceding result and the reduced linearization given in the next section, we state the following characterization of the unreduced linearization.

Theorem 5.10 (Linearization before reduction). *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ be a simple mechanical control system satisfying Assumption 5.9, let X be a complete infinitesimal symmetry for Σ satisfying Assumption 3.1, and let $\chi\colon \mathbb{R} \to \mathsf{Q}$ be a regular relative equilibrium. For a vector field ξ along χ, let $x(t) = T_{\chi(t)}\pi_{\mathsf{B}}(\xi(t))$ and $\nu(t) = \dot{\zeta}(t)$, where $\mathrm{ver}(\xi(t)) = \zeta(t)X(\chi(t))$. Then the pair (ξ, u) is a controlled trajectory for the linearization of Σ along $(\chi, 0)$ if and only if*

$$
\mathrm{hlft}_{\chi(t)}(\ddot{x}(t)) + \dot{\nu}(t)X(\chi(t)) = -\mathbb{G}^\sharp \circ \mathrm{Hess}V_X^\sharp(\mathrm{hlft}_{\chi(t)}(x(t)))
$$
$$
- \frac{2\langle \boldsymbol{d}V(\chi(t)); \mathrm{hlft}_{\chi(t)}(\dot{x}(t))\rangle}{\|X\|^2_\mathbb{G}(\chi(t))}X(\chi(t))
$$
$$
+ \mathrm{hlft}_{\chi(t)}(C_X(\dot{x}(t))) + 2\nu(t)\mathrm{grad}V(\chi(t))
$$
$$
+ \mathbb{G}^\sharp \circ A^\flat(\mathrm{hlft}_{\chi(t)}(\dot{x}(t))) + \nu(t)\mathbb{G}^\sharp \circ A^\flat \circ X(\chi(t)), \quad (14)
$$

where $b_0 = \pi_{\mathsf{B}}(\chi(0))$.

Proof. As in Proposition 2.1(ii), $\overset{\mathbb{G}}{\nabla}_X X = -\frac{1}{2}\mathrm{grad}\|X\|^2_\mathbb{G}$. Therefore,

$$
\overset{\mathbb{G}}{\nabla}_X X + \mathrm{grad}V = \mathrm{grad}(V - \tfrac{1}{2}\|X\|^2_\mathbb{G}) = \mathrm{grad}V_X.
$$

By Proposition 2.5, for each $t \in \mathbb{R}$, $\mathrm{grad}V_X(\chi(t)) = 0$. Using this fact, it is straightforward (e.g., using coordinates) to show that

$$
\overset{\mathbb{G}}{\nabla}(\mathrm{grad}V_X)(\chi(t)) = \mathbb{G}^\sharp(\chi(t)) \circ \mathrm{Hess}V_X^\flat(\chi(t)).
$$

Furthermore, since V_X is X-invariant, for any $x \in \mathsf{T}_{b_0}\mathsf{B}$, we have

$$
\mathbb{G}^\sharp(\chi(t)) \circ \mathrm{Hess}V_X^\flat(\chi(t))(\mathrm{hlft}_{\chi(t)}(x)) = \mathrm{hlft}_{\chi(t)}(\mathbb{G}^\sharp_{\mathsf{B}}(b_0) \circ \mathrm{Hess}(V_X)^\sharp_{\mathsf{B}}(b_0)(x)).
$$

For a vertical tangent vector $v_{\chi(t)} \in \mathsf{V}_{\chi(t)}\mathsf{Q}$ we have

$$\overset{\mathbb{G}}{\nabla}(\mathrm{grad}V_X)(v_{\chi(t)}) = 0,$$

using the fact that $\mathrm{grad}V_X(\chi(t)) = 0$ and using X-invariance of V_X. Summarizing the preceding computations is the following formula for a vector field ξ along χ:

$$\overset{\mathbb{G}}{\nabla}(\overset{\mathbb{G}}{\nabla}_X X + \mathrm{grad}V)(\xi(t)) = \mathrm{hlft}_{\chi(t)}\big(\mathbb{G}_{\mathsf{B}}^{\sharp}(b_0) \circ \mathrm{Hess}(V_X)_{\mathsf{B}}^{\flat}(b_0)(T\pi_{\mathsf{B}}(\xi(t)))\big). \tag{15}$$

Now let ξ be a vector field along χ and let Ξ be a vector field extending ξ. Since $X(\chi(t)) = \chi'(t)$, we have, using the definition of the curvature tensor,

$$\overset{\mathbb{G}}{\nabla}^2_{\chi'(t)}\xi(t) + R(\xi(t), \chi'(t))\chi'(t)$$
$$= \overset{\mathbb{G}}{\nabla}_X \overset{\mathbb{G}}{\nabla}_X \Xi(\chi(t)) + \overset{\mathbb{G}}{\nabla}_\Xi \overset{\mathbb{G}}{\nabla}_X X(\chi(t)) - \overset{\mathbb{G}}{\nabla}_X \overset{\mathbb{G}}{\nabla}_\Xi X(\chi(t)) - \overset{\mathbb{G}}{\nabla}_{[\Xi,X]} X(\chi(t)).$$

A straightforward manipulation, using the fact that $\overset{\mathbb{G}}{\nabla}$ has zero torsion, gives

$$\overset{\mathbb{G}}{\nabla}_X \overset{\mathbb{G}}{\nabla}_X \Xi + \overset{\mathbb{G}}{\nabla}_\Xi \overset{\mathbb{G}}{\nabla}_X X - \overset{\mathbb{G}}{\nabla}_X \overset{\mathbb{G}}{\nabla}_\Xi X - \overset{\mathbb{G}}{\nabla}_{[\Xi,X]} X =$$
$$= \overset{\mathbb{G}}{\nabla}_\Xi \overset{\mathbb{G}}{\nabla}_X X + 2\overset{\mathbb{G}}{\nabla}_{[X,\Xi]} X + [X,[X,\Xi]]. \tag{16}$$

Around a point $\chi(t_0) \in \mathrm{image}(\chi)$, let $(\mathcal{U}, \phi)$ be a chart with coordinates $(q^1, \dots, q^n)$ having the following properties:

(1) $X = \frac{\partial}{\partial q^n}$;
(2) $((q^1, \dots, q^{n-1}), (q^n))$ are fiber bundle coordinates for $\pi_{\mathsf{B}} \colon \mathsf{Q} \to \mathsf{B}$;
(3) for any point $\chi(t) \in \mathcal{U}$, the basis $\{\frac{\partial}{\partial q^1}(\chi(t)), \dots, \frac{\partial}{\partial q^n}(\chi(t))\}$ for $\mathsf{T}_{\chi(t)}\mathsf{Q}$ is $\mathbb{G}$-orthogonal.

In these coordinates one readily determines that

$$[X,\Xi](\chi(t)) = \dot{\xi}^i(t)\frac{\partial}{\partial q^i}, \quad [X,[X,\Xi]](\chi(t)) = \ddot{\xi}^i(t)\frac{\partial}{\partial q^i} \tag{17}$$

for all values of t for which $\chi(t) \in \mathcal{U}$. In these coordinates it also holds that

$$\mathrm{hlft}_{\chi(t)}\frac{\partial}{\partial q^a}(b_0) = \frac{\partial}{\partial q^a}(\chi(t)), \qquad a \in \{1, \dots, n-1\}. \tag{18}$$

Therefore, if ξ is as above and if $x(t) = T\pi_{\mathsf{B}}(\xi(t))$, then we have

$$
\begin{aligned}
2\overset{\mathbb{G}}{\nabla}X([X,\Xi](\chi(t))) &= 2\big(\overset{\mathbb{G}}{\nabla}_{\mathrm{hlft}_{\chi(t)}(\dot{x}(t))}X(\chi(t))\big) + 2\overset{\mathbb{G}}{\nabla}X(\mathrm{ver}([X,\Xi])(\chi(t))) \\
&= -\,\mathrm{hlft}_{\chi(t)}(C_X(\dot{x}(t))) + 2\mathrm{ver}(\overset{\mathbb{G}}{\nabla}X(\mathrm{hlft}_{\chi(t)}(\dot{x}(t)))) \\
&\quad + 2\overset{\mathbb{G}}{\nabla}X(\mathrm{ver}([X,\Xi])(\chi(t))) \\
&= -\,\mathrm{hlft}_{\chi(t)}(C_X(\dot{x}(t))) + 2\overset{\mathbb{G}}{\nabla}X(\mathrm{ver}([X,\Xi])(\chi(t))) \\
&\quad + \frac{2X(\chi(t))}{\|X\|_{\mathbb{G}}^2(\chi(t))}\mathbb{G}(\overset{\mathbb{G}}{\nabla}X(\mathrm{hlft}_{\chi(t)}(\dot{x}(t))),X(\chi(t))) \\
&= -\,\mathrm{hlft}_{\chi(t)}(C_X(\dot{x}(t))) + 2\overset{\mathbb{G}}{\nabla}X(\mathrm{ver}([X,\Xi])(\chi(t))) \\
&\quad - \frac{2X(\chi(t))}{\|X\|_{\mathbb{G}}^2(\chi(t))}\mathbb{G}(\overset{\mathbb{G}}{\nabla}_X X(\chi(t)),\mathrm{hlft}_{\chi(t)}(\dot{x}(t))) \\
&= -\,\mathrm{hlft}_{\chi(t)}(C_X(\dot{x}(t))) + 2\overset{\mathbb{G}}{\nabla}X(\mathrm{ver}([X,\Xi])(\chi(t))) \\
&\quad + \frac{2X(\chi(t))}{\|X\|_{\mathbb{G}}^2(\chi(t))}\langle \boldsymbol{d}V(\chi(t));\mathrm{hlft}_{\chi(t)}(\dot{x}(t))\rangle, \qquad (19)
\end{aligned}
$$

using the fact that $\overset{\mathbb{G}}{\nabla}_X X = \mathrm{grad}V_X - \mathrm{grad}V$, and that $\boldsymbol{d}V_X(\chi(t)) = 0$ for all $t \in \mathbb{R}$. Also,

$$
\mathrm{hlft}_{\chi(t)}(\ddot{x}(t)) = \mathrm{hor}([X,[X,\Xi]](\chi(t))), \qquad t \in \mathbb{R}. \qquad (20)
$$

In the coordinates $(q^1, \ldots, q^n)$, one also computes

$$
\overset{\mathbb{G}}{\nabla}X = \overset{\mathbb{G}}{\Gamma}{}^i_{nj}\frac{\partial}{\partial q^i} \otimes \mathrm{d}q^j,
$$

from which we ascertain that

$$
2\overset{\mathbb{G}}{\nabla}X(\mathrm{ver}([X,\Xi])(\chi(t))) = 2\overset{\mathbb{G}}{\Gamma}{}^i_{nn}\dot{\xi}^n\frac{\partial}{\partial q^i},
$$

where no summation is intended over the index "n." One readily verifies that, in our coordinates,

$$
\overset{\mathbb{G}}{\Gamma}{}^i_{nn}\frac{\partial}{\partial q^i} = \overset{\mathbb{G}}{\nabla}_X X = \mathrm{grad}V_X - \mathrm{grad}V.
$$

Since $\boldsymbol{d}V_X(\chi(t)) = 0$, we have

$$
2\overset{\mathbb{G}}{\nabla}X(\mathrm{ver}([X,\Xi])(\chi(t))) = -2\mathrm{ver}([X,\Xi](\chi(t)))\mathrm{grad}V(\chi(t)). \qquad (21)
$$

We also clearly have, by definition of $\mathscr{L}^{X,\chi}$,

$$
\dot{\xi}^n(t)\frac{\partial}{\partial q^n} = \mathrm{ver}(\mathscr{L}^{X,\chi}(\xi(t))), \qquad (22)
$$

where no summation is intended over "n."

To simplify the terms involving the external force, we note that

$$
\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^{\sharp} \circ A^{\flat}(X(\chi(t)))) =
$$
$$
= (\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^{\sharp} \circ A^{\flat}))(X(\chi(t))) + \mathbb{G}^{\sharp} \circ A^{\flat}(\overset{\mathbb{G}}{\nabla}_{\xi(t)}X(\chi(t))).
$$

Thus, using the fact that $\overset{\mathbb{G}}{\nabla}$ is torsion-free, we have

$$
- \overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^{\sharp} \circ A^{\flat} \circ X)(\chi(t)) + (\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^{\sharp} \circ A^{\flat}))(\chi'(t))
$$
$$
+ \mathbb{G}^{\sharp} \circ A^{\flat}(\overset{\mathbb{G}}{\nabla}_{\chi'(t)}\xi(t)) = \mathbb{G}^{\sharp} \circ A^{\flat}([X, \Xi](\chi(t))).
$$

Using (17) and (18) we arrive at

$$
- \overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^{\sharp} \circ A^{\flat} \circ X)(\chi(t)) + (\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^{\sharp} \circ A^{\flat}))(\chi'(t))
$$
$$
+ \mathbb{G}^{\sharp} \circ A^{\flat}(\overset{\mathbb{G}}{\nabla}_{\chi'(t)}\xi(t)) = \mathbb{G}^{\sharp} \circ A^{\flat}(\text{hlft}_{\chi(t)}(\dot{x}(t))) + \dot{\xi}^{n}\mathbb{G}^{\sharp} \circ A^{\flat}(X(\chi(t))), \quad (23)
$$

where $x(t) = T_{\chi(t)}\pi_{\mathsf{B}}(\xi(t))$.

Finally, for a vector field ξ along χ, let $x(t) = T\pi_{\mathsf{B}}(\xi(t))$ and let $\nu(t)X(t) = \text{ver}(\mathscr{L}^{X,\chi}(\xi))$. In terms of our coordinates above, $\nu(t) = \dot{\xi}^{n}(t)$. One now combines equations (15), (16), (17), (19), (20), (21), (22) and (23) to get the result. $\qquad\square$

Remark 5.11. *The preceding theorem is not obvious; in particular, the equivalence of equations (13) and (14) is not transparent. Indeed, the relationship between the curvature tensor and the components of the system that appear in the theorem statement, C, $\text{Hess}(V_X)$, and $\text{grad}V$, is rather subtle. In this respect, the proof of the theorem bears study, if these relationships are to be understood.* $\qquad\bullet$

5.4. **Linearization of the reduced equations along a relative equilibrium.** We again consider a forced simple mechanical control system $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ satisfying Assumption 5.9, take X to be a complete infinitesimal symmetry for Σ satisfying Assumption 3.1, and let χ be a relative equilibrium. In this section we provide the form of the linearization along a relative equilibrium by linearizing, in the usual manner, the reduced

equations, which we reproduce here for convenience:

$$
\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{\eta'(t)}\eta'(t) = -\mathrm{grad}_{\mathsf{B}}\big(V_{X,v(t)}^{\mathrm{eff}}\big)_{\mathsf{B}}(\eta(t)) + v(t)C_X(\eta'(t))
$$

$$
+ T\pi_{\mathsf{B}} \circ \mathbb{G}^{\sharp} \circ A^{\flat}\big(\gamma'(t) - X(\gamma(t))\big) + Y_{\mathsf{B},u}(t,\eta(t)),
$$

$$
\dot{v}(t) = -\frac{v(t)\langle \boldsymbol{d}(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(\eta(t)); \eta'(t)\rangle}{(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(\eta(t))} + \frac{\langle A^{\flat}(\gamma'(t) - X(\gamma(t))); X(\gamma(t))\rangle}{(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(\eta(t))}
$$

$$
+ \frac{\mathbb{G}(Y_u(t,\gamma(t)), X(\gamma(t)))}{(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(\eta(t))}.
$$

$$(24)$$

Here (γ, u) is a controlled trajectory for Σ, $\eta = \pi_{\mathsf{B}} \circ \gamma$, and v is defined by $\mathrm{ver}(\gamma'(t)) = v(t)X(\gamma(t))$.

The reduced equations are straightforward to linearize, since we are merely linearizing about an equilibrium point. To compactly state the form of the linearization requires some notation. Define a $(1,1)$-tensor field A_{B} on B by

$$
A_{\mathsf{B}}(v_b) = T_q\pi_{\mathsf{B}} \circ \mathbb{G}^{\sharp}(q) \circ A^{\flat}(q) \circ \mathrm{hlft}_q(v_b),
$$

for $q \in \pi_{\mathsf{B}}^{-1}(b)$. This definition can be shown to be independent of the choice of $q \in \pi_{\mathsf{B}}^{-1}(b)$ by virtue of the X-invariance of A. Define a vector field a_{B} on B by

$$
a_{\mathsf{B}}(b) = T_q\pi_{\mathsf{B}} \circ \mathbb{G}^{\sharp}(q) \circ A^{\flat}(q)(X(q)),
$$

where $q \in \pi_{\mathsf{B}}^{-1}(b)$, and again this definition can be shown to be well-defined. Finally, define a one-form α_{B} on B by

$$
\langle \alpha_{\mathsf{B}}(b); v_b \rangle = \frac{\langle A^{\flat}(\mathrm{hlft}_q(v_b)); X(q)\rangle}{(\|X\|_{\mathbb{G}}^{2})_{\mathsf{B}}(b)},
$$

where $q \in \pi_{\mathsf{B}}^{-1}(b)$ and $v_b \in \mathsf{T}_b\mathsf{B}$. This definition, too, is independent of the choice of q.

We may now state the form of the linearization of the reduced equations. The proof of the following result is by fairly simple direct computation.

Proposition 5.12 (Linearization after reduction). *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ be a forced simple mechanical control system, let X be a complete infinitesimal symmetry of Σ satisfying Assumption 3.1, and let χ be a regular relative equilibrium with $b_0 = \pi_{\mathsf{B}} \circ \chi(0)$.*

The linearization of equations (24) *about* $(0_{b_0}, 1)$ *is the linear control system* $(\mathsf{T}_{b_0}\mathsf{B} \oplus \mathsf{T}_{b_0}\mathsf{B} \oplus \mathbb{R}, A_\Sigma(b_0), B_\Sigma(b_0))$, *where*

$$A_\Sigma(b_0) = \left[\begin{array}{cc} 0 & \mathrm{id}_{\mathsf{T}_{b_0}\mathsf{B}} \\ -\mathbb{G}_\mathsf{B}(b_0)^\sharp \circ \mathrm{Hess}(V_X)_\mathsf{B}(b_0)^\flat & C_X(b_0) + A_\mathsf{B}(b_0) \\ 0 & -2\frac{\boldsymbol{d}V_\mathsf{B}(b_0)}{(\|X\|_\mathbb{G}^2)_\mathsf{B}(b_0)} + \alpha_\mathsf{B}(b_0) \end{array} \right.$$
$$\left. \begin{array}{c} 0 \\ 2\mathrm{grad}_\mathsf{B}V_\mathsf{B}(b_0) + a_\mathsf{B}(b_0) \\ \frac{\langle A^\flat(X(q_0)); X(q_0)\rangle}{(\|X\|_\mathbb{G}^2)_\mathsf{B}(b_0)} \end{array} \right],$$

$$B_\Sigma(b_0) = \left[\begin{array}{c} 0 \\ B_{\Sigma,2}(b_0) \\ B_{\Sigma,3}(b_0) \end{array} \right],$$

where $B_{\Sigma,2}(b_0) \in L(\mathbb{R}^m; \mathsf{T}_{b_0}\mathsf{B})$ *is defined by*

$$B_{\Sigma,2}(b_0)(u) = \sum_{a=1}^m u^a Y_{\mathsf{B},a}(b_0),$$

and where $B_{\Sigma,3}(b_0) \in L(\mathbb{R}^m; \mathbb{R})$ *is defined by*

$$B_{\Sigma,3}(b_0)(u) = \sum_{a=1}^m u^a \frac{\mathbb{G}(Y_a(\chi(0)), X(\chi(0)))}{(\|X\|_\mathbb{G}^2)_\mathsf{B}(b_0)}.$$

The equations governing controlled trajectories for the linearization of the reduced system are

$$\ddot{x}(t) = -\mathbb{G}_\mathsf{B}(b_0)^\sharp \circ \mathrm{Hess}(V_X)_\mathsf{B}(b_0)^\flat(x(t)) + C_X(b_0)(\dot{x}(t))$$
$$+ 2\nu(t)\mathrm{grad}_\mathsf{B}V_\mathsf{B}(b_0) + A_\mathsf{B}(b_0)(\dot{x}(t)) + \nu(t)a_\mathsf{B}(b_0) + B_{\Sigma,2}(b_0) \cdot u(t),$$

$$\dot{\nu}(t) = -2\frac{\langle \boldsymbol{d}V_\mathsf{B}(b_0); \dot{x}(t)\rangle}{(\|X\|_\mathbb{G}^2)_\mathsf{B}(b_0)} + \alpha_\mathsf{B}(b_0)(\dot{x}(t))$$
$$+ \nu(t)\frac{\langle A^\flat(X(q_0)); X(q_0)\rangle}{(\|X\|_\mathbb{G}^2)_\mathsf{B}(b_0)} + B_{\Sigma,3}(b_0) \cdot u(t).$$

The following result gives the relationship between the reduced and the unreduced linearization.

Theorem 5.13 (Relationship between linearizations). *Let* $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F, \mathscr{F})$ *be a forced simple mechanical control system with* F *satisfying Assumption 5.9, let* X *be a complete infinitesimal symmetry of* Σ *satisfying Assumption 3.1, and let* χ *be a regular relative equilibrium with* $b_0 = \pi_\mathsf{B} \circ \chi(0)$.

For a curve $t \mapsto x(t) \in \mathsf{T}_{b_0}\mathsf{B}$, a vector field ξ along χ, a function $\nu \colon \mathbb{R} \to \mathbb{R}$, and a locally integrable control $t \mapsto u(t)$, the following statements are equivalent:

(i) $t \mapsto (x(t) \oplus \dot{x}(t) \oplus \nu(t), u(t))$ *is a controlled trajectory for the linearization of the equations* (24) *about* $(0_{b_0}, 1)$, *and in turn* $\mathrm{hor}(\xi(t)) = \mathrm{hlft}_{\chi(t)}(x(t))$, *and* $\nu(t) = \dot{\zeta}(t)$, *where* $\mathrm{ver}(\xi(t)) = \zeta(t)X(\chi(t))$;

(ii) (ξ, u) *is a controlled trajectory for the linearization of* Σ *about* $(\chi, 0)$, *and in turn* $x(t) = T\pi_{\mathsf{B}}(\xi(t))$, *and* $\nu(t) = \dot{\zeta}(t)$, *where* $\mathrm{ver}(\xi(t)) = \zeta(t)X(\chi(t))$.

Proof. This follows easily from Theorem 5.10 and Proposition 5.12. $\qquad\square$

The theorem is an important one, since it will allow us to switch freely between the reduced and unreduced linearizations. In some cases, it will be convenient to think of certain concepts in the unreduced setting, while computations are more easily performed in the reduced setting.

6. Linearized effective energies

In many existing results concerning stability of relative equilibria, a central role is played by Hessian of the energy. This is a consequence of the fact that definiteness of the Hessian, restricted to certain subspaces, can easily deliver stability results in various forms. In this section we study the Hessian of the effective energy for a relative equilibria. In particular, we consider the interplay of the various natural energies with the reduction process and with linearization. Specifically, we spell out the geometry relating the processes of linearization and reduction.

6.1. **Some geometry associated to an infinitesimal isometry.** The utility of the constructions in this section may not be immediately apparent, but will become clear in Proposition 6.6 below.

In this section we let $(\mathsf{Q}, \mathbb{G})$ be a Riemannian manifold with X an infinitesimal isometry satisfying Assumption 3.1. We denote by TTQ_X the restriction of the vector bundle $\pi_{\mathsf{TTQ}} \colon \mathsf{TTQ} \to \mathsf{TQ}$ to $\mathrm{image}(X)$. Thus TTQ_X is a vector bundle over $\mathrm{image}(X)$ whose fiber at $X(q)$ is $\mathsf{T}_{X(q)}\mathsf{TQ}$. We denote this fiber by $\mathsf{TTQ}_{X,X(q)}$. In like manner, HTQ_X and VTQ_X denote the restrictions of HTQ and TQ, respectively, to $\mathrm{image}(X)$. The Ehresmann connection on $\pi_{\mathsf{TQ}} \colon \mathsf{TQ} \to \mathsf{Q}$, defined by $\overset{\mathbb{G}}{\nabla}$ as in Section 4.2, gives a splitting of each fiber of TTQ_X as

$$\mathsf{TTQ}_{X,X(q)} = \mathsf{HTQ}_{X,X(q)} \oplus \mathsf{VTQ}_{X,X(q)}.$$

This gives a vector bundle isomorphism $\sigma_X \colon \mathsf{TTQ}_X \to \mathsf{HTQ}_X \oplus \mathsf{VTQ}_X$. Denote by $\Pi_\mathsf{B} \colon \mathsf{TQ} \to \mathsf{TQ}/\mathbb{R}$ the projection onto the set of X^T-orbits. We define $\phi_\mathsf{B} \colon \mathsf{TQ} \to \mathsf{TB} \times \mathbb{R}$ by

$$\phi_\mathsf{B}(w_q) = (T\pi_\mathsf{B}(w_q), \nu_X(w_q)),$$

where $\nu_X(w_q)$ is defined by $\mathrm{ver}(w_q) = \nu_X(w_q)X(q)$. Note that $\phi_\mathsf{B} \circ X(q) = (0_{\pi_\mathsf{B}(q)}, 1)$. Indeed, one can easily see that $\phi_{\mathsf{B},X} \triangleq \phi_\mathsf{B}|\mathrm{image}(X) \colon \mathrm{image}(X) \to Z(\mathsf{TB}) \times \{1\}$ is a surjective submersion. We next define a vector bundle map $\psi_\mathsf{B} \colon \mathsf{TTQ} \to \mathsf{T}(\mathsf{TB} \times \mathbb{R})$ over ϕ_B by $\psi_\mathsf{B} = T\phi_\mathsf{B}$. We denote $\psi_{\mathsf{B},X} = \psi_\mathsf{B}|\mathsf{TTQ}_X$, noting that this is a surjective vector bundle map from TTQ_X to the restricted vector bundle $\mathsf{T}(\mathsf{TB} \times \mathbb{R})|(Z(\mathsf{B}) \times \{1\})$. Next we wish to give a useful description of the vector bundle $\mathsf{T}(\mathsf{TB} \times \mathbb{R})|(Z(\mathsf{B}) \times \{1\})$. We think of $\mathsf{TB} \times \mathbb{R}$ as a vector bundle over $\mathsf{B} \times \mathbb{R}$, and we let $\mathbb{R}_{\mathsf{B} \times \mathbb{R}}$ be the trivial vector bundle $(\mathsf{B} \times \mathbb{R}) \times \mathbb{R}$ over $\mathsf{B} \times \mathbb{R}$. We then note that $\mathsf{T}_{0_b}\mathsf{TB} \simeq \mathsf{T}_b \oplus \mathsf{T}_b\mathsf{B}$, where the first component in the direct sum is tangent to $Z(\mathsf{TB})$ (i.e., is horizontal) and the second component is tangent to $\mathsf{T}_b\mathsf{B}$ (i.e., is vertical). Thus we have a natural identification

$$\mathsf{T}(\mathsf{TB} \times \mathbb{R})|(Z(\mathsf{B}) \times \{1\}) \simeq (\mathsf{TB} \times \mathbb{R}) \oplus (\mathsf{TB} \times \mathbb{R}) \oplus \mathbb{R}_{\mathsf{B} \times \mathbb{R}} \qquad (25)$$

of vector bundles over $Z(\mathsf{TB}) \times \{1\} \simeq \mathsf{B} \times \{1\}$. The fiber over $(b, 1)$ is isomorphic to $\mathsf{T}_b\mathsf{B} \oplus \mathsf{T}_b\mathsf{B} \oplus \mathbb{R}$. We shall implicitly use the identification (25) in the sequel. Next, we define a vector bundle map $\iota_\mathsf{B} \colon \mathsf{HTQ} \oplus \mathsf{VTQ} \to (\mathsf{TB} \times \mathbb{R}) \oplus (\mathsf{TB} \times \mathbb{R}) \oplus \mathbb{R}_{\mathsf{B} \times \mathbb{R}}$ by

$$\iota_\mathsf{B}(u_{v_q} \oplus w_{v_q}) = \left(T_q\pi_\mathsf{B}(u_{v_q}), T_q\pi_\mathsf{B}(w_{v_q} - \overset{\mathbb{G}}{\nabla}X(u_{v_q})), \nu_X(w_{v_q} - \overset{\mathbb{G}}{\nabla}X(u_{v_q}))\right).$$

We then let $\iota_{\mathsf{B},X}$ be the restriction of ι_B to $\mathsf{HTQ}_X \oplus \mathsf{VTQ}_X$.

The following result summarizes and ties together the above constructions.

Lemma 6.1. *The following statements hold:*

(i) *σ_X is a vector bundle isomorphism over $\mathrm{id}_{\mathrm{image}(X)}$ from TTQ_X to $\mathsf{HTQ}_X \oplus \mathsf{VTQ}_X$;*

(ii) *$\phi_{\mathsf{B},X}$ is a surjective submersion from $\mathrm{image}(X)$ to $\mathsf{B} \times \{1\}$;*

(iii) *$\psi_{\mathsf{B},X}$ is a surjective vector bundle map over $\phi_{\mathsf{B},X}$ from TTQ_X to $(\mathsf{TB} \times \mathbb{R}) \oplus (\mathsf{TB} \times \mathbb{R}) \oplus \mathbb{R}_{\mathsf{B} \times \mathbb{R}}$;*

(iv) *$\iota_{\mathsf{B},X}$ is a surjective vector bundle map over $\phi_{\mathsf{B},X}$ from $\mathsf{HTQ}_X \oplus \mathsf{VTQ}_X$ to $(\mathsf{TB} \times \mathbb{R}) \oplus (\mathsf{TB} \times \mathbb{R}) \oplus \mathbb{R}_{\mathsf{B} \times \mathbb{R}}$;*

(v) the following diagram commutes:

$$\mathsf{TTQ}_X \xrightarrow{\;\;\sigma_X\;\;} \mathsf{HTQ}_X \oplus \mathsf{VTQ}_X$$

$$\psi_{\mathsf{B},X} \searrow \qquad \swarrow \iota_{\mathsf{B},X}$$

$$(\mathsf{TB} \times \mathbb{R}) \oplus (\mathsf{TB} \times \mathbb{R}) \oplus \mathbb{R}_{\mathsf{B} \times \mathbb{R}}$$

Proof. This is most easily proved in an appropriate set of coordinates. Take coordinates $(q^1, \ldots, q^n)$ for Q with the following properties:

(1) $X = \frac{\partial}{\partial q^n}$;
(2) for times t for which $\chi(t)$ is in the chart domain, $\{\frac{\partial}{\partial q^1}(\chi(t)), \ldots, \frac{\partial}{\partial q^n}(\chi(t))\}$ is an orthogonal basis for $\mathsf{T}_{\chi(t)}\mathsf{Q}$.

This means that $(q^1, \ldots, q^{n-1})$ are coordinates for B. These also form, therefore, coordinates for $Z(\mathsf{TB})$ and thus also for $Z(\mathsf{TB}) \times \{1\}$. Since a typical point in image(X) has the form

$$((q^1, \ldots, q^n), (0, \ldots, 0, 1))$$

in natural coordinates for TQ, we can use $(q^1, \ldots, q^n)$ as coordinates for image(X). We denote natural coordinates for TTQ by $((\boldsymbol{q}, \boldsymbol{v}), (\boldsymbol{u}, \boldsymbol{w}))$. Then $(\boldsymbol{q}, \boldsymbol{u}, \boldsymbol{w})$ form a set of coordinates for TTQ_X.

The map ϕ_{B} from TQ to $\mathsf{TB} \times \mathbb{R}$ has the form

$$((q^1, \ldots, q^n), (v^1, \ldots, v^n)) \mapsto ((q^1, \ldots, q^{n-1}), (v^1, \ldots, v^{n-1}), v^n).$$

In the coordinates for image(X) and for $Z(\mathsf{TB}) \times \{1\}$, the map $\phi_{\mathsf{B},X}$ has the form

$$(q^1, \ldots, q^n) \mapsto (q^1, \ldots, q^{n-1}).$$

The coordinate form of ψ_{B} is then

$$(((q^1, \ldots, q^n), (v^1, \ldots, v^n)), ((u^1, \ldots, u^n), (w^1, \ldots, w^n)))$$
$$\mapsto (((q^1, \ldots, q^{n-1}), (v^1, \ldots, v^{n-1}), v^n), ((u^1, \ldots, u^{n-1}), (w^1, \ldots, w^{n-1}), w^n))$$

and the coordinate form for $\psi_{\mathsf{B},X}$ is given by

$$((q^1, \ldots, q^n), (u^1, \ldots, u^n), (w^1, \ldots, w^n))$$
$$\mapsto ((q^1, \ldots, q^{n-1}), (u^1, \ldots, u^{n-1}), (w^1, \ldots, w^{n-1}), w^n). \quad (26)$$

In coordinates, the map σ_X is given by

$$((q^1, \ldots, q^n), (u^1, \ldots, u^n), (w^1, \ldots, w^n))$$
$$\mapsto ((q^1, \ldots, q^n), (u^1, \ldots, u^n), (w^1 + \overset{\mathbb{G}}{\Gamma}{}^i_{nj} u^j, \ldots, w^n + \overset{\mathbb{G}}{\Gamma}{}^n_{nj} u^j)).$$

Finally, in our above coordinates, the form of the map $\iota_{\mathsf{B},X}$ is

$$((q^1, \ldots, q^n), (u^1, \ldots, u^n), (w^1, \ldots, w^n))$$
$$\mapsto ((q^1, \ldots, q^{n-1}), (u^1, \ldots, u^n), (w^1 - \overset{\mathbb{G}}{\Gamma}{}^1_{nj} u^j, \ldots, w^{n-1} - \overset{\mathbb{G}}{\Gamma}{}^{n-1}_{nj} w^j), w^n - \overset{\mathbb{G}}{\Gamma}{}^n_{nj} u^j).$$

All statements in the statement of the lemma follow directly from the preceding coordinate computations. $\qquad\square$

We shall see in Proposition 6.6 that $\iota_{\mathsf{B},X}$ relates two natural energies associated to a relative equilibrium.

6.2. **The effective energies and their linearizations.** We let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a simple mechanical system with X a complete infinitesimal symmetry for Σ satisfying Assumption 3.1. First recall from Lemma 2.4 that the effective energy for a forced simple mechanical system $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ with complete infinitesimal symmetry X is

$$E_X(v_q) = \tfrac{1}{2} \left\| v_q - X(q) \right\|_{\mathbb{G}}^2 + V_X(q),$$

where $V_X = V - \frac{1}{2} \|X\|_{\mathbb{G}}^2$ is the effective potential. The relative equilibria for Σ are then characterized by the critical points of E_X, as in Proposition 2.5. As we shall see in Section 7, the Hessian of the effective energy at such critical points is useful for determining the stability of the corresponding relative equilibrium. The following result characterizes this Hessian in terms of the splitting of the fibers of TTQ using the Ehresmann connection on $\pi_{\mathsf{TTQ}} \colon \mathsf{TTQ} \to \mathsf{TQ}$ associated with $\overset{\mathbb{G}}{\nabla}$.

Lemma 6.2. *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system and let X be a complete infinitesimal symmetry for Σ. Let v_q be a critical point for the effective energy and let $\mathsf{T}_q\mathsf{Q} \oplus \mathsf{T}_q\mathsf{Q}$ be the splitting of $\mathsf{T}_{v_q}\mathsf{TQ}$ associated with $\overset{\mathbb{G}}{\nabla}$, as described in Section 4.2. Then*

$$\mathrm{Hess} E_X(u_1 \oplus w_1, u_2 \oplus w_2) =$$
$$= \mathbb{G}(w_1 - \overset{\mathbb{G}}{\nabla}_{u_1} X(q), w_2 - \overset{\mathbb{G}}{\nabla}_{u_2} X(q)) + \mathrm{Hess} V_X(u_1, u_2).$$

Proof. This is a messy, but straightforward, proof in coordinates. $\qquad\square$

With this as background, we make the following definition, recalling the notation $u_{v_q} \oplus w_{v_q}$ to denote a point in $\mathsf{T}_{v_q}\mathsf{TQ}$ relative to the splitting defined by $\overset{\mathbb{G}}{\nabla}$.

Definition 6.3 (Linearized effective energy)**.** *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system, let X be a complete infinitesimal isometry*

*for Σ, and let χ be a relative equilibrium. The **linearized effective energy** is the function on $\mathsf{TTQ}|\mathrm{image}(\chi')$ defined by*

$$E_\chi(u_{v_q} \oplus w_{v_q}) = \tfrac{1}{2}\|w_{v_q} - \overset{\mathbb{G}}{\nabla}_{u_{v_q}} X\|_{\mathbb{G}}^2 + \tfrac{1}{2}\mathrm{Hess}V_X(u_{v_q}, u_{v_q}),$$

where $v_q = \chi'(0)$. •

Next we consider the linearized effective energy, but now for the reduced system. To do so, we assume that $\pi_\mathsf{B}\colon \mathsf{Q} \to \mathsf{B}$ is a surjective submersion, as in Assumption 3.1. The effective energy E_X is X-invariant and so drops to $\mathsf{TQ}/\mathbb{R} \simeq \mathsf{TB} \times \mathbb{R}$. We may further explicitly computed this "reduced effective energy," denoted by E_X^{red}, as

$$E_X^{\mathrm{red}}(w_b, v) = \tfrac{1}{2}\mathbb{G}_\mathsf{B}(w_b, w_b) + (V_X)_\mathsf{B}(b) + \tfrac{1}{2}(\|X\|_{\mathbb{G}}^2)_\mathsf{B}(b)(v-1)^2.$$

It then makes sense that the "reduced linearized effective energy" should be the Hessian of this function at a critical point, which corresponds, as we have seen, to a relative equilibrium. The following result records the form of the Hessian.

Lemma 6.4. *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system, let X be a complete infinitesimal symmetry for Σ satisfying Assumption 3.1, and let $(0_b, 1) \in \mathsf{TB} \times \mathbb{R}$ be a critical point for $(E_X)_{\mathsf{TB}\times\mathbb{R}}$. Then*

$$\mathrm{Hess}E_X^{\mathrm{red}}(0_b, 1)(u_1 \oplus v_1 \oplus \nu_1, u_2 \oplus v_2 \oplus \nu_2)$$
$$= \mathbb{G}_\mathsf{B}(b)(v_1, v_2) + \mathrm{Hess}(V_X)_\mathsf{B}(b)(u_1, u_2) + (\|X\|_{\mathbb{G}}^2)_\mathsf{B}(b)\nu_1\nu_2.$$

Proof. This is a straightforward computation. □

Based on this computation, let us make the following definition.

Definition 6.5 (Reduced linearized effective energy)**.** *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system, let X be a complete infinitesimal isometry for Σ satisfying Assumption 3.1, and let $b_0 = \pi_\mathsf{B}(\chi(0))$. The **reduced linearized effective energy** is the function on $\mathsf{T}_{b_0}\mathsf{B} \oplus \mathsf{T}_{b_0}\mathsf{B} \oplus \mathbb{R}$ defined by*

$$E_\chi^{\mathrm{red}}(u_{b_0}, v_{b_0}, \nu) = \tfrac{1}{2}\|v_{b_0}\|_{\mathbb{G}_\mathsf{B}}^2 + \tfrac{1}{2}\mathrm{Hess}(V_X)_\mathsf{B}(b_0)(u_{b_0}, u_{b_0}) + \tfrac{1}{2}(\|X\|_{\mathbb{G}}^2)_\mathsf{B}(b_0)\nu^2.$$
•

The preceding definition of the reduced linearized effective energy is obtained by "reducing" the effective energy, and then "linearizing" it. It should be possible to perform the operations in the opposite order to get to the same answer. To do this explicitly, we use the constructions of the preceding section. In particular, we use the vector bundle map $\iota_{\mathsf{B},X}$. As the following result indicates, one should think of this map as describing how

the process of linearization is reduced when using the Ehresmann connection on $\pi_{\mathsf{TTQ}} \colon \mathsf{TTQ} \to \mathsf{TQ}$ associated with $\overset{\mathbb{G}}{\nabla}$.

Proposition 6.6 (Relating the linearized effective energies). *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system with X a complete infinitesimal symmetry satisfying Assumption 3.1. If χ is a relative equilibrium, then $\iota_{\mathsf{B},X}(\chi'(t))^* E_\chi^{\mathrm{red}} = E_\chi$ for all $t \in \mathbb{R}$.*

Proof. Let $q = \chi(t)$ and let $b = \pi_{\mathsf{B}}(q)$. We compute

$$
\iota_{\mathsf{B},X}(\chi'(t))^* E_\chi^{\mathrm{red}}(u_{v_q} \oplus w_{v_q}) =
$$

$$
\begin{aligned}
&= \tfrac{1}{2}\|T_q \pi_{\mathsf{B}}(w_{v_q} - \overset{\mathbb{G}}{\nabla} X(u_{v_q}))\|_{\mathbb{G}_{\mathsf{B}}}^2 + \tfrac{1}{2}(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}\nu_X(w_{v_q} - \overset{\mathbb{G}}{\nabla} X(u_q))^2 + \\
&\quad + \tfrac{1}{2}\mathrm{Hess}(V_X)_{\mathsf{B}}(b)(T_q \pi_{\mathsf{B}}(u_b), T_q \pi_{\mathsf{B}}(u_b)) \\
&= \tfrac{1}{2}\|\mathrm{hor}(w_{v_q} - \overset{\mathbb{G}}{\nabla} X(u_{v_q}))\|_{\mathbb{G}}^2 + \|\mathrm{ver}(w_{v_q} - \overset{\mathbb{G}}{\nabla} X(u_{v_q}))\|_{\mathbb{G}}^2 + \\
&\quad + \tfrac{1}{2}\mathrm{Hess}V_X(q)(u_{v_q}, u_{v_q}),
\end{aligned}
$$

as desired. $\qquad\square$

7. Stability of relative equilibria

In this section we apply the results of the preceding sections to study the stability of a relative equilibrium. We deal with stability in two ways: linear and nonlinear. Because we use the linearized effective energy as a Lyapunov function in our linear stability analysis, if the conditions we give for linear stability are satisfied, then one can conclude nonlinear stability as well.

7.1. **Nonlinear stability definitions.** Let us establish the basic setup for all the following definitions and results. We consider a forced simple mechanical system $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ with a time-independent external force F and with the following governing equation:

$$
\overset{\mathbb{G}}{\nabla}_{\gamma'(t)} \gamma'(t) = -\mathrm{grad} V(\gamma(t)) + \mathbb{G}^\sharp(F(\gamma'(t))). \tag{27}
$$

We assume for simplicity that all solutions to this equation can be defined on all of $\mathbb{R}$.

Let us define the notions of stability for relative equilibria that we consider in this paper.

Definition 7.1 (Base and fiber stability). *Let X be a complete infinitesimal symmetry for Σ, let $\chi \colon \mathbb{R} \to \mathsf{Q}$ be a regular relative equilibrium, and let $b_0 = \pi_{\mathsf{B}}(\chi(0))$.*

(i) *The relative equilibrium χ is **base stable** if, for all neighborhoods $\mathcal{U}$ of 0_{b_0}, there exists a neighborhood $\mathcal{W}$ of $\chi'(0)$ such that each solution $\gamma\colon \mathbb{R} \to \mathsf{Q}$ of equation (27), with initial condition $\gamma'(0) \in \mathcal{W}$, satisfies $T\pi_{\mathsf{B}} \circ \gamma'(t) \in \mathcal{U}$ for all $t \in \mathbb{R}$.*

(ii) *The relative equilibrium χ is **locally asymptotically base stable** if it is base stable, and if there exists a neighborhood $\mathcal{V}$ of $\chi'(0)$ such that each solution $\gamma\colon \mathbb{R} \to \mathsf{Q}$ of equation (27), with initial condition $\gamma'(0) \in \mathcal{V}$, satisfies $\lim_{t\to+\infty} T\pi_{\mathsf{B}} \circ \gamma'(t) = 0_{b_0}$.*

(iii) *The relative equilibrium χ is **fiber stable** if, for all neighborhoods $\mathcal{U}_{\mathbb{R}}$ of 1, there exists a neighborhood $\mathcal{W}$ of $\chi'(0)$ such that each solution $\gamma\colon \mathbb{R} \to \mathsf{Q}$ of equation (27), with initial condition $\gamma'(0) \in \mathcal{W}$, satisfies $J_X(\gamma'(t)) \|X\|_{\mathbb{G}}^{-2}(\gamma(t)) \in \mathcal{U}_{\mathbb{R}}$ for all $t \in \mathbb{R}$.*

(iv) *The relative equilibrium χ is **locally asymptotically fiber stable** if it is fiber stable, and if there exists a neighborhood $\mathcal{V}$ of $\chi'(0)$ such that each solution $\gamma\colon \mathbb{R} \to \mathsf{Q}$ of equation (27), with initial condition $\gamma'(0) \in \mathcal{V}$, satisfies $\lim_{t\to+\infty} J_X(\gamma'(t)) \|X\|_{\mathbb{G}}^{-2}(\gamma(t)) = 1$.* $\bullet$

To describe base stability, consider an invariant tube around the regular relative equilibrium χ. Base stability means that trajectories starting in the tube remain in the tube at subsequent times; see Figure 3. Fiber stability means that trajectories starting with nearby initial velocities will maintain nearby fiber velocities at subsequent times. Even if a system is base and fiber stable, trajectories with nearby initial conditions may diverge.

What we call base stability of a relative equilibrium corresponds to the usual notion of stability for a relative equilibrium considered in the literature: stability of the equilibrium point in the reduced space. What we call fiber stability has to do, note, with only the velocity along the fiber. Thus fiber stability will generally not suffice to give other forms of stability along the relative equilibrium, such as the G-stability of Patrick [1992]. We also remark that the definitions of base and fiber stability are examples of what is sometimes called "partial stability" in more general contexts. With this sort of stability, one is only interested in the behavior of some of the states of the system. This is studied in the text [Vorotnikov 1998].

7.2. Linear stability definitions. We let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system where the force F satisfies Assumption 5.9, with X a complete infinitesimal symmetry for the system, and with $\chi\colon \mathbb{R} \to \mathsf{Q}$ a relative equilibrium. First we need a definition for linear stability of a relative equilibrium. The definition relies on the linearization along the relative equilibrium, which, from the developments of Section 5, satisfies an
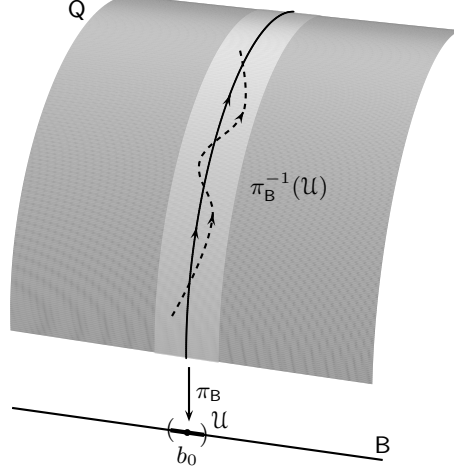
FIGURE 3. Base stability of relative equilibria

initial value problem of the form

$$
\overset{\mathbb{G}}{\nabla}^2_{\chi'(t)}\xi(t) + R(\xi(t), \chi'(t))\chi'(t)
$$

$$
= -\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathrm{grad}V)(\chi(t)) - \overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^\sharp \circ A^\flat \circ X)(\chi(t))
$$

$$
+ (\overset{\mathbb{G}}{\nabla}_{\xi(t)}(\mathbb{G}^\sharp \circ A^\flat))(\chi'(t)) + \mathbb{G}^\sharp \circ A^\flat(\overset{\mathbb{G}}{\nabla}_{\chi'(t)}\xi(t)),
$$

$$
\xi(0) = \xi_0, \ \mathscr{L}^{X,\chi}(\xi)(0) = v_{\xi,0}. \quad (28)
$$

Remarks 7.2.   (1) *Note that it is immaterial that we specify the initial condition at $t = 0$ due to X-invariance of the system.*

(2) *Also note that we can specify the initial derivative condition for ξ by specifying $\overset{\mathbb{G}}{\nabla}_{\chi'(0)}\xi(0)$. Since both $\mathscr{L}^{X,\chi}$ and $\overset{\mathbb{G}}{\nabla}_{\chi'}$ are differential operators in $\mathsf{T}\mathsf{Q}$ along χ, $\mathscr{L}^{X,\chi}(\xi)(t) - \overset{\mathbb{G}}{\nabla}_{\chi'(t)}\xi(t)$ depends only on $\xi(t)$. Thus specifying $\xi(0)$ and $\mathscr{L}^{X,\chi}(\xi)(0)$ is equivalent to specifying $\xi(0)$ and $\overset{\mathbb{G}}{\nabla}_{\chi'(0)}\xi(0)$. For our purposes, it is more convenient to specify the derivative initial condition in terms of $\mathscr{L}^{X,\chi}(\xi)(0)$.* ●

We may now state our stability definitions.

Definition 7.3 (Linear stability of relative equilibria). *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system with F satisfying Assumption 5.9, with X a complete infinitesimal symmetry for Σ, and with $\chi \colon \mathbb{R} \to \mathsf{Q}$ a*

relative equilibrium. For a vector field ξ along χ, let $\nu(t) = \dot{\zeta}(t)$, where $\mathrm{ver}(\xi(t)) = \zeta(t)X(\chi(t))$.

*(i) The relative equilibrium χ is **linearly base (resp. fiber) stable** if there exists $M > 0$ such that the solution $t \mapsto \xi(t)$ to the initial value problem (28) satisfies $\|\mathrm{hor}(\xi(t))\|_{\mathbb{G}} + \|\mathscr{L}^{X,\chi}(\mathrm{hor}(\xi))(t)\|_{\mathbb{G}} \leq M(\|\mathrm{hor}(\xi_0)\|_{\mathbb{G}} + \|v_{\xi,0}\|_{\mathbb{G}})$ (resp. $|\nu(t)| \leq M(\|\mathrm{hor}(\xi_0)\|_{\mathbb{G}} + \|v_{\xi,0}\|_{\mathbb{G}})$).*

*(ii) The relative equilibrium χ is **linearly asymptotically base (resp. fiber) stable** if each solution $t \mapsto \xi(t)$ to the initial value problem (28) satisfies $\lim_{t\to+\infty}(\|\mathrm{hor}(\xi(t))\|_{\mathbb{G}} + \|\mathscr{L}^{X,\chi}(\mathrm{hor}(\xi))(t)\|_{\mathbb{G}}) = 0$ (resp. $\lim_{t\to+\infty} \nu(t) = 0$).* •

Let us now give the relationship between these definitions of linear stability and the linear stability of the reduced system. To do this, let us write the equations governing the reduced linearization, following Proposition 5.12:

$$
\begin{aligned}
\ddot{x}(t) = & -\mathbb{G}_{\mathsf{B}}(b_0)^{\sharp} \circ \mathrm{Hess}(V_X)_{\mathsf{B}}(b_0)^{\flat}(x(t)) + C_X(b_0)(\dot{x}(t)) \\
& + 2\nu(t)\mathrm{grad}_{\mathsf{B}}V_{\mathsf{B}}(b_0) + F_{\mathsf{B}}(b_0)(\dot{x}(t)) + \nu(t)f_{\mathsf{B}}(b_0), \\
\dot{\nu}(t) = & -2\frac{\langle dV_{\mathsf{B}}(b_0); \dot{x}(t)\rangle}{(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(b_0)} + \alpha_{\mathsf{B}}(b_0)(\dot{x}(t)) + \nu(t)\frac{\langle A^{\flat}(X(q_0)); X(q_0)\rangle}{(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(b_0)},
\end{aligned}
\tag{29}
$$

where we adopt the notation for F_{B}, f_{B}, and α_{B} as given before the statement of Proposition 5.12. We also let $\bar{\mathbb{R}}_+$ denote the set of nonnegative real numbers.

Proposition 7.4 (Base characterization of linear stability). *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system with F satisfying Assumption 5.9, with X a complete infinitesimal symmetry for the system satisfying Assumption 3.1, and with $\chi \colon \mathbb{R} \to \mathsf{Q}$ a regular relative equilibrium. Let $b_0 = \pi_{\mathsf{B}}(\chi(0))$. The following statements hold:*

(i) χ is linearly base stable if and only if, for every solution $t \mapsto x(t) \oplus \dot{x}(t) \oplus \nu(t)$ of the equations (29), the function $\bar{\mathbb{R}}_+ \ni t \mapsto \|x(t)\|_{\mathbb{G}_{\mathsf{B}}(b_0)}$ is bounded;

(ii) χ is linearly asymptotically base stable if and only if, for every solution $t \mapsto x(t) \oplus \dot{x}(t) \oplus \nu(t)$ of the equations (29), $\lim_{t\to+\infty} \|x(t)\|_{\mathbb{G}_{\mathsf{B}}(b_0)} = 0$;

(iii) χ is linearly fiber stable if and only if, for every solution $t \mapsto x(t) \oplus \dot{x}(t) \oplus \nu(t)$ of the equations (29), the function $\bar{\mathbb{R}}_+ \ni t \mapsto |\nu(t)|$ is bounded;

(iv) χ is linearly asymptotically fiber stable if and only if, for every solution $t \mapsto x(t) \oplus \dot{x}(t) \oplus \nu(t)$ of the equations (29), $\lim_{t\to+\infty} |\nu(t)| = 0$.

Proof. This is a direct consequence of properties of time-independent linear differential equations. □

7.3. **Lyapunov stability analysis.** Next, we endeavor to provide stability criteria for relative equilibria. We start by introducing a few appropriate concepts.

Let us also introduce the appropriate notions of dissipative forces along a relative equilibrium.

Definition 7.5. *Let $X \in \Gamma^\infty(\mathsf{TQ})$. A time-independent force $F \colon \mathsf{TQ} \to \mathsf{T}^*\mathsf{Q}$ is **dissipative about X** if $\langle F(v_q); v_q - X(q) \rangle \leq 0$ for each $v_q \in \mathsf{TQ}$, and is **strictly dissipative about X** if it is dissipative about X, and if $\langle F(v_q); v_q - X(q) \rangle = 0$ only when $v_q = X(q)$.* •

We can now collect some basic statements about these concepts.

Lemma 7.6. *Consider the forced simple mechanical system $(\mathsf{Q}, \mathbb{G}, V, F)$ with infinitesimal symmetry X. If F is dissipative about X, and if $\gamma \colon I \to \mathsf{Q}$ is a trajectory of equation (27), then the function $t \mapsto E_X(\gamma'(t))$ is nonincreasing.*

Proof. Recall that the vector field on TQ describing the dynamics of $(\mathsf{Q}, \mathbb{G}, V, F)$ is $\overset{\mathbb{G}}{S} - \mathrm{vlft}(\mathrm{grad}V) + \mathrm{vlft}(\mathbb{G}^\sharp \circ F)$. Because E_X is a constant of motion for $(\mathsf{Q}, \mathbb{G}, V)$, we know that

$$\mathscr{L}_{\overset{\mathbb{G}}{S} - \mathrm{vlft}(\mathrm{grad}V)} E_X = 0.$$

Next, from Lemma 2.4 we compute

$$\mathscr{L}_{\mathrm{vlft}(\mathbb{G}^\sharp \circ F)} E_X(v_q) = \langle F(v_q); v_q - X(q) \rangle \leq 0. \qquad (30)$$

□

Finally, we can provide criteria for the stability of relative equilibria in terms of the character of V_X at critical points. These results parallel those presented for equilibrium configurations for equilibria of mechanical systems (see, for example, [Bullo and Lewis 2004, Theorem 6.45]). Recall that, under Assumption 3.1, $(V_X)_\mathsf{B}$ is the projection of V_X onto the base space B.

Theorem 7.7 (Stability of relative equilibria)**.** *Consider a forced simple mechanical system $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ and let X be a complete infinitesimal symmetry for Σ satisfying Assumption 3.1. Assume F is dissipative about X and assume $F(X(q)) = 0_q$ for each $q \in \mathsf{Q}$. Then the following statements hold.*

(i) A regular relative equilibrium $\chi \colon \mathbb{R} \to \mathsf{Q}$ is base and fiber stable if $(V_X)_\mathsf{B}$ is locally positive-definite about $\pi_\mathsf{B}(\chi(0))$.

> (ii) A regular relative equilibrium χ is locally asymptotically base stable and locally asymptotically fiber stable if $(V_X)_{\mathsf{B}}$ has an isolated local minimum at $\pi_{\mathsf{B}}(\chi(0))$, and if F is strictly dissipative about X.

Proof. To prove parts (i) and (ii), we use the regularity of the relative equilibrium χ to invoke Theorem 3.3. We recall the equivalent representation of the dynamics of Σ obtained in Theorem 3.3(ii). If $\gamma\colon \mathbb{R} \to \mathsf{Q}$ is a solution to (27), then the curve $\eta = \pi_{\mathsf{B}} \circ \gamma\colon \mathbb{R} \to \mathsf{B}$, and the curve $v\colon \mathbb{R} \to \mathbb{R}$ given by $v(t) = J_X(\gamma'(t)) \|X\|_{\mathbb{G}}^{-2}(\gamma(t))$, together satisfy

$$
\overset{\mathbb{G}_{\mathsf{B}}}{\nabla}_{\eta'(t)} \eta'(t) = -\operatorname{grad}_{\mathsf{B}}\left(V_{X,v(t)}^{\mathrm{eff}}\right)_{\mathsf{B}}(\eta(t)) + v(t)C_X(\eta'(t)) + T\pi_{\mathsf{B}}\left(\mathbb{G}^{\sharp} \circ F(\gamma'(t))\right),
$$

$$
\dot{v}(t) = -\frac{v(t)\langle \boldsymbol{d}(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t)); \eta'(t)\rangle}{(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t))} + \frac{1}{(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(\eta(t))}\langle F(\gamma'(t)); X(\gamma(t))\rangle.
$$

$$(31)$$

From the curves η and v, the curve $\gamma\colon \mathbb{R} \to \mathsf{Q}$ is computed to be $\gamma(t) = \Phi_{0,t}^{vX}(\mathrm{hlft}_{q_0}(\eta)(t))$.

Let $\Sigma_{\mathsf{TB} \times \mathbb{R}}$ be reduced dynamical system on $\mathsf{TB} \times \mathbb{R}$ defined by equations (31). Given this characterization, the curve $\chi\colon \mathbb{R} \to \mathsf{Q}$ is a relative equilibrium for Σ if and only if its projection $t \mapsto (\pi_{\mathsf{B}}(\chi(0)), 1)$ is a solution to equations (31); that is, an equilibrium point for $\Sigma_{\mathsf{TB} \times \mathbb{R}}$.

Now we note that E_X is a function invariant under the flow of X^T. Therefore, E_X can be projected onto the reduced space $\mathsf{TB} \times \mathbb{R}$. Given the decomposition $\mathsf{TQ} = \mathsf{HQ} \oplus \mathsf{VQ}$, we write $v_q \in \mathsf{TQ}$ as $v_q = \mathrm{hlft}_q(w_b) + \nu X(q)$ for $(w_b, \nu) \in \mathsf{TB} \times \mathbb{R}$. Accordingly

$$
\begin{aligned}
(E_X)_{\mathsf{TB} \times \mathbb{R}}(w_b, \nu) &= E_X(\mathrm{hlft}_q(w_b) + \nu X(q)) \\
&= (V_X)_{\mathsf{B}}(b) + \tfrac{1}{2}\mathbb{G}_{\mathsf{B}}(w_b, w_b) + \tfrac{1}{2}(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(b)(\nu - 1)^2.
\end{aligned}
$$

Because E_X is a nonincreasing function along the trajectories of Σ, its projection $(E_X)_{\mathsf{TB} \times \mathbb{R}}(w_b, \nu)$ is a nonincreasing function along the trajectories of $\Sigma_{\mathsf{TB} \times \mathbb{R}}$. If $(V_X)_{\mathsf{B}}$ is locally positive-definite about $\pi_{\mathsf{B}}(\chi(0))$, then

$$
(w_b, \nu) \mapsto (E_X)_{\mathsf{TB} \times \mathbb{R}}(w_b, \nu) - V_X(\chi(0))
$$

is a Lyapunov function for the dynamical system $\Sigma_{\mathsf{TB} \times \mathbb{R}}$ about the point $(0_{\pi_{\mathsf{B}}(\chi(0))}, 1) \in \mathsf{TB} \times \mathbb{R}$. This proves part (i), that is, the relative equilibrium $\chi\colon \mathbb{R} \to \mathsf{Q}$ is base and fiber stable.

To prove (ii), we invoke the LaSalle Invariance Principle for the dynamical system $\Sigma_{\mathsf{TB} \times \mathbb{R}}$. We now assume that F is strictly dissipative about X and recall that F is X-invariant. Since $\pi_{\mathsf{B}}(\chi(0)) \in \mathsf{B}$ is an isolated local minimum for $(V_X)_{\mathsf{B}}$, there exists a neighborhood $\mathcal{W} \subset \mathsf{B}$ of $\pi_{\mathsf{B}}(\chi(0))$ such that $(V_X)_{\mathsf{B}}(b) > (V_X)_{\mathsf{B}}(b_0)$ for all $b \in \mathcal{W} \setminus \{\pi_{\mathsf{B}}(\chi(0))\}$, and $\mathcal{W}$ contains no

critical point other than $\pi_\mathsf{B}(\chi(0))$. Next, we compute

$$
\begin{aligned}
A &= \left\{ (w_b, \nu) \in \mathsf{T}\mathcal{W} \times \mathbb{R} \mid\ \langle F(v_q); v_q \rangle = 0 \text{ for } v_q = \mathrm{hlft}_q(w_b) + \nu X(q) \right\} \\
&= \left\{ (w_b, \nu) \in \mathsf{T}\mathcal{W} \times \mathbb{R} \mid\ v_q = X(q) \text{ for } v_q = \mathrm{hlft}_q(w_b) + \nu X(q) \right\} \\
&= \left\{ (0_b, 1) \in \mathsf{T}\mathcal{W} \times \mathbb{R} \right\}.
\end{aligned}
$$

On A, the equations (31) reduce to $0 = -\mathrm{grad}_\mathsf{B}\big(V_X\big)_\mathsf{B}(\eta)$. Therefore, no trajectory of (31) takes values in A, other than the trivial $t \mapsto (\pi_\mathsf{B}(\chi(0)), 1)$. Under these conditions, the LaSalle Invariance Principle guarantees that $(\pi_\mathsf{B}(\chi(0)), 1)$ is locally asymptotically stable.   □   □

Theorem 7.7(i) can be stated directly in terms of the effective potential function rather than in terms of its projection. The helpful equivalence is stated as follows.

Lemma 7.8. *Let $X \in \Gamma^\infty(\mathsf{T}\mathsf{Q})$ satisfy Assumption 3.1. Let $\psi \in C^\infty(\mathsf{Q})$ be X-invariant and let $\psi(q_0) = 0$ for $q_0 \in \mathsf{Q}$. Then the following statements are true:*

    (i) *$\pi_\mathsf{B}(q_0)$ is a critical point for ψ_B if and only if q_0 is a critical point for ψ;*

    (ii) *ψ_B is locally positive-definite about $\pi_\mathsf{B}(q_0)$ if and only if there exists a neighborhood $\mathcal{W}$ of q_0 with the properties that*

        (a) *$q \in \mathcal{W}$ implies $\psi(q) \geq 0$, and*

        (b) *$\psi(q) = 0$ only if q belongs to the integral curve of X through q_0;*

    (iii) *if q_0 is a critical point for ψ, then $\mathrm{Hess}\psi_\mathsf{B}(\pi_\mathsf{B}(q_0))$ is positive-definite if and only if $\mathrm{Hess}\psi(q_0)$ is positive-definite on any complement to $\mathrm{span}_\mathbb{R}\{X(q_0)\}$ in $\mathsf{T}_{q_0}\mathsf{Q}$.*

Proof. There exists a chart $(\mathcal{U}, \phi)$ with coordinates $(q^1, \ldots, q^n)$, with $\phi(q_0) = \mathbf{0}$, and $X = \frac{\partial}{\partial q^1}$. Because ψ is X-invariant, the coordinate representation of ψ does not depend on q^1 and therefore its Taylor expansion is

$$
\psi(q^1, \ldots, q^n) =
$$
$$
= \sum_{i=2}^n (\boldsymbol{d}\psi(q_0))_i q^i + \tfrac{1}{2} \sum_{i,j=2}^n (\mathrm{Hess}\psi(q_0))_{ij} q^i q^j + O\big(\|(q^1, \ldots, q^n)\|_{\mathbb{R}^n}^3\big).
$$

The integral curve of X through q_0 is an open subset of the axis $\left\{(q^1, 0, \ldots, 0) \mid q^1 \in \mathbb{R}\right\}$, and by construction $\psi_\mathsf{B}(q^2, \ldots, q^n) = \psi(0, q^2, \ldots, q^n)$. Statements (i) and (ii) are consequences of these facts. Part (iii) follows by noting that $\mathrm{Hess}\psi(q_0)$ is positive-definite on any complement to $\mathrm{span}_\mathbb{R}\{(1, 0, \ldots, 0)\}$ if and only if the $(n-1)$ eigenvalues of the $(n-1) \times (n-1)$ matrix with components $(\mathrm{Hess}\psi(q_0))_{ij}$, $i, j \in \{2, \ldots, n\}$, are strictly positive.   □

Remarks 7.9. (1) Effective versus amended potential functions: *Proposition 2.5 and Theorem 7.7*(i) *can be stated in terms of the amended potential function instead of the effective potential function (see [Bullo and Lewis 2004, Exercise 6.20]). For part* (ii), *it is convenient to rely on the effective potential function rather than the amended potential function.*

(2) Similarities with equilibrium configurations: *The results of Theorem 7.7 are similar to those characterizing the existence and stability of equilibrium configurations. Roughly speaking, the effective energy plays the same role in analyzing a relative equilibria as the energy function plays in analyzing equilibrium configurations (cf. [Bullo and Lewis 2004, Theorem 6.45]). In the study of relative equilibria, the effective potential function plays a role parallel to the one played by the potential function in the study of equilibrium configurations. Furthermore, the second term in Lemma 2.4 characterizing the effective energy is a positive-definite function in the "velocity error"* $(v_q - X(q))$.

(3) Comparison with the Energy-Momentum Method: *The stability criterion in Theorem 7.7 is only sufficient: it turns out that, because of the Noether Conservation Law, it is not necessary to require positive-definitness of* $\mathrm{Hess} V_X$ *on every complement to* $\mathrm{span}_{\mathbb{R}}\{X(q_0)\}$ *in* $\mathsf{T}_{q_0}\mathsf{Q}$. *The Energy-Momentum Method described in [Marsden 1992, Marsden and Ratiu 1999, Simo, Lewis, and Marsden 1991] provides a sharper, more detailed analysis by taking this into account.* •

7.4. **Linear stability analysis.** Now let us give some natural sufficient conditions that rely only on the linearization. These conditions should be thought of as the linear analogue to the stability results of the preceding section. As such, they use the notions of linearized effective energies from Section 6.2. In this regard, the next result is the main result in this section.

Theorem 7.10 (Linear stability of relative equilibria)**.** *Let* $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ *be a forced simple mechanical system, with* X *a complete infinitesimal symmetry for* Σ *satisfying Assumption 3.1, and with* $\chi \colon \mathbb{R} \to \mathsf{Q}$ *a regular relative equilibrium. Suppose that* $F(v_q) = -R^{\flat}_{\mathrm{diss}}(v_q - X(q))$, *where* R_{diss} *is a symmetric positive-semidefinite* $(0,2)$-*tensor field. For* $b_0 = \pi_{\mathsf{B}} \circ \chi(0)$, *the following statements hold:*

(i) χ *is linearly base and fiber stable if* $\mathrm{Hess}(V_X)_{\mathsf{B}}(b_0)$ *is positive-definite;*

(ii) χ *is linearly asymptotically base stable and linearly asymptotically fiber stable if* $\mathrm{Hess}(V_X)_{\mathsf{B}}(b_0)$ *is positive-definite and if* R_{diss} *is positive-definite.*

Proof. Note that combined linear (asymptotic) base and fiber stability of χ is equivalent to the linear (asymptotic) stability of the equilibrium point b_0 for the reduced system on $\mathsf{TB} \times \mathbb{R}$. Therefore, in the proof, we shall consider the stability in the reduced space, using the reduced linearized effective energy, E_χ^{ref}, as a candidate Lyapunov function. First note that, under the hypothesis that $\mathrm{Hess}(V_X)_{\mathsf{B}}(b_0)$ is positive-definite, it follows that E_χ^{ref} is positive-definite about $0_b \oplus 0_b \oplus 0$. Next, a straightforward computation, the details of which we omit, shows that

$$\frac{\mathrm{d}E_\chi^{\mathrm{red}}}{\mathrm{d}t}(x(t) \oplus \dot{x}(t) \oplus \nu(t)) = \mathbb{G}_{\mathsf{B}}(\dot{x}(t), F_{\mathsf{B}}(b_0)(\dot{x}(t))) + \nu(t)\mathbb{G}_{\mathsf{B}}(\dot{x}(t), f_{\mathsf{B}}(b_0))$$
$$+ \nu(t)(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(b_0)\alpha_{\mathsf{B}}(b_0)(\dot{x}(t)) + \nu(t)^2\langle A^\flat(X(q_0)); X(q_0)\rangle,$$

along a solution $t \mapsto x(t) \oplus \dot{x}(t) \oplus \nu(t)$ to equations (29), where $q_0 \in \pi_{\mathsf{B}}^{-1}(b_0)$, and where A is as in Assumption 5.9. One now can easily show that

$$\mathbb{G}_{\mathsf{B}}(\dot{x}(t), F_{\mathsf{B}}(b_0)(\dot{x}(t))) = A(\mathrm{hlft}_{q_0}(\dot{x}(t)), \mathrm{hlft}_{q_0}(\dot{x}(t))),$$
$$\nu(t)\mathbb{G}_{\mathsf{B}}(\dot{x}(t), f_{\mathsf{B}}(b_0)) = A(\nu(t)X(q_0), \mathrm{hlft}_{q_0}(\dot{x}(t))),$$
$$\nu(t)(\|X\|_{\mathbb{G}}^2)_{\mathsf{B}}(b_0)\alpha_{\mathsf{B}}(b_0)(\dot{x}(t)) = A(\mathrm{hlft}_{q_0}(\dot{x}(t)), \nu(t)X(q_0)).$$

These computations allow us to conclude that

$$\frac{\mathrm{d}E_\chi^{\mathrm{red}}}{\mathrm{d}t}(x(t) \oplus \dot{x}(t) \oplus \nu(t)) =$$
$$= A(\mathrm{hlft}_{q_0}(\dot{x}(t)) + \nu(t)X(q_0), \mathrm{hlft}_{q_0}(\dot{x}(t)) + \nu(t)X(q_0)).$$

In part (i), $A = -R_{\mathrm{diss}}$ is negative-semidefinite, and in part (ii), $A = -R_{\mathrm{diss}}$ is negative-definite, and the result then follows directly. $\square$

Alternatively, one can check the hypotheses of the theorem using the linearized effective energy. The following result contains the results of this transcription.

Corollary 7.11 (Unreduced linear stability of relative equilibria). *Let $\Sigma = (\mathsf{Q}, \mathbb{G}, V, F)$ be a forced simple mechanical system, with X a complete infinitesimal symmetry for Σ satisfying Assumption 3.1, and with $\chi\colon \mathbb{R} \to \mathsf{Q}$ a regular relative equilibrium. Suppose that $F(v_q) = -R_{\mathrm{diss}}^\flat(v_q - X(q))$, where R_{diss} is a Rayleigh dissipation function. The following statements hold:*

(i) χ is linearly base and fiber stable if $\mathrm{Hess}V_X(\chi(t))$ is positive-definite on any (and so every) complement to $\mathrm{span}_{\mathbb{R}}\{X(\chi(t))\}$ for some (and so for all) $t \in \mathbb{R}$;

(ii) χ is linearly asymptotically base stable and linearly asymptotically fiber stable if $\mathrm{Hess}V_X(\chi(t))$ is positive-definite on any (and so

every) complement to $\mathrm{span}_{\mathbb{R}}\{X(\chi(t))\}$ *for some (and so for all)* $t \in \mathbb{R}$, *and if* R_{diss} *is positive-definite.*

Remarks 7.12. (1) *Note that since positive-definiteness of the Hessian of* $(V_X)_{\mathsf{B}}$ *at* b_0 *implies that* b_0 *is an isolated local minimum for* $(V_X)_{\mathsf{B}}$, *the satisfaction of the hypotheses of Theorem 7.10 implies the satisfaction of the hypotheses of Theorem 7.7.*

(2) *The presence of gyroscopic forces in the reduced linearization makes it difficult to draw the sharpest possible conclusions regarding linear stability. This is to be contrasted with the linear stability of equilibrium points, where it is possible to give much sharper stability conditions in the presence of only dissipative forces (see, for example, [Bullo and Lewis 2004, Theorem 6.42]).*

(3) *In part* (ii) *of Theorem 7.10 we require that* R_{diss} *be positive-definite. As is the case with stability of equilibria for mechanical systems, this hypothesis is stronger than required.* •

8. Open problems and future directions

We conclude with a discussion of some of the more obvious research directions suggested by our approach.

(1) First and foremost, it would be important to analyze the setting of reduction by a general Lie group action. In such a setup, one might consider a principal fiber bundle $\pi\colon \mathsf{Q} \to \mathsf{B}$ with structure group being a Lie group G. One would then be interested in the Riemannian geometry of the principal bundle. For example, reduction should be describable in terms of ideas such as the reduced Riemannian metric on B, the mechanical connection and its curvature, the locked inertia tensor, the second fundamental form of the fibers, and the amended/effective potential. This picture is only partially understood at present (but see some initial results in [Cortés and Martínez 2003]), and the application of this picture to stability theory is described in [Marsden 1992].

(2) In a related direction, with regards to the stability analysis for relative equilibria, it would be interesting to describe the Energy-Momentum method of [Simo, Lewis, and Marsden 1991, Simo, Posbergh, and Marsden 1990] in Riemannian geometric language.

(3) Generalizing from the Riemannian context, one could study the geometry of a G-invariant affine connection on a principal fiber bundle $\pi\colon \mathsf{Q} \to \mathsf{B}$. From the point of view of mechanics, this is interesting since the dynamics of systems with nonholonomic constraints can be described using an affine connection that is generally

different from the affine connection. This is an idea that goes back to Synge [1928].

(4) It would also be interesting to characterize the linearization of a mechanical control system along more general trajectories than relative equilibria. An interesting example of such trajectories are those arising in the theory of kinematic reduction for affine connection control systems [Bullo and Lewis 2003, Bullo and Lynch 2001].

(5) From a control theoretical viewpoint, very little is known about the linear and nonlinear controllability properties of mechanical control systems along relative equilibria. This is due to the complications arising from the gyroscopic forces present in the reduced equations. The geometry of the controllability properties of linear mechanical systems with gyroscopic forces are not well understood presently, but some results are contained in the paper of Hughes and Skelton [1980].

References

Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, second edition, Addison Wesley, Reading, MA, ISBN 0-8053-0102-X.

Arnol'd, V. I. [1966] *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits*, Université de Grenoble. Annales de l'Institut Fourier, **16**(1), 319–361.

Bloch, A. M. [2003] *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Applied Mathematics*, Springer-Verlag, New York–Heidelberg–Berlin, ISBN 0-387095535-6.

Bullo, F. [2000] *Stabilization of relative equilibria for underactuated systems on Riemannian manifolds*, Automatica. The Journal of IFAC. The International Federation of Automatic Control, **36**(12), 1819–1834.

Bullo, F. and Lewis, A. D. [2003] *Low-order controllability and kinematic reductions for affine connection control systems*, SIAM Journal on Control and Optimization, to appear.

— [2004] *Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Systems*, number 49 in Texts in Applied Mathematics, Springer-Verlag, New York–Heidelberg–Berlin, ISBN 0-387-22195-6.

— [2005] *Supplementary chapters for* Geometric Control of Mechanical Systems *[Bullo and Lewis 2004]*.
  URL: `http://penelope.mast.queensu.ca/smcs/`

Bullo, F. and Lynch, K. M. [2001] *Kinematic controllability and decoupled trajectory planning for underactuated mechanical systems*, Institute

of Electrical and Electronics Engineers. Transactions on Robotics and Automation, **17**(4), 402–412.

Cendra, H., Marsden, J. E., Pekarsky, S., and Ratiu, T. S. [2003] *Variational principles for Lie–Poisson and Hamilton–Poincaré equations*, Moscow Mathematical Journal, **3**(3), 833–867.

Cendra, H., Marsden, J. E., and Ratiu, T. S. [2001] *Lagrangian reduction by stages*, Memoirs of the American Mathematical Society, **152**(722).

Cortés, J. and Martínez, S. [2003] *Motion control algorithms for simple mechanical systems with symmetry*, Acta Applicandae Mathematicae. An International Journal on Applying Mathematics and Mathematical Applications, **76**(3), 221–264.

Crampin, M. and Pirani, F. A. E. [1986] *Applicable Differential Geometry*, number 59 in London Mathematical Society Lecture Note Series, Cambridge University Press, New York/Port Chester/Melbourne/Sydney, ISBN 0-521-23190-6.

Guillemin, V. and Sternberg, S. [1984] *Symplectic Techniques in Physics*, Cambridge University Press, New York/Port Chester/Melbourne/Sydney, ISBN 0-521-38990-9.

Hughes, P. C. and Skelton, R. E. [1980] *Controllability and observability of linear matrix-second-order systems*, American Society of Mechanical Engineers. Transactions of the ASME. Series E. Journal of Applied Mechanics, **47**(2), 415–420.

Jalnapurkur, S. M. and Marsden, J. E. [2000] *Stabilization of relative equilibria*, Institute of Electrical and Electronics Engineers. Transactions on Automatic Control, **45**(8), 1483–1491.

— [2001] *Stabilization of relative equilibria. II*, Regular & Chaotic Dynamics. Regulyarnaya & Khaoticheskaya Dinamika, **3**(3), 161–179.

Kobayashi, S. and Nomizu, K. [1963] *Foundations of Differential Geometry*, Volumes I and II, number 15 in Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers, New York.

Lewis, A. D. and Tyner, D. R. [2003] *Jacobian linearisation in a geometric setting*, in *Proceedings of the 42nd IEEE Conference on Decision and Control*, pages 6084–6089, Maui, HI.

Marsden, J. E. [1992] *Lectures on Mechanics*, number 174 in London Mathematical Society Lecture Note Series, Cambridge University Press, New York/Port Chester/Melbourne/Sydney, ISBN 0-521-42844-0.

Marsden, J. E. and Ratiu, T. S. [1999] *Introduction to Mechanics and Symmetry*, second edition, number 17 in Texts in Applied Mathematics, Springer-Verlag, New York–Heidelberg–Berlin, ISBN 0-387-98643-X.

Marsden, J. E. and Weinstein, A. [1974] *Reduction of symplectic manifolds with symmetry*, Reports on Mathematical Physics, **5**(1), 121–130.

Meyer, K. R. [1973] *Symmetries and integrals in mechanics*, in *Dynamical Systems*, M. Peixoto, editor, pages 259–273, Academic Press, New York.

Ortega, J.-P. and Ratiu, T. S. [2004] *Momentum Maps and Hamiltonian Reduction*, number 222 in Progress in Mathematics, Birkhäuser, Boston/Basel/Stuttgart, ISBN 0-8176-4307-9.

Patrick, G. W. [1992] *Relative equilibria in Hamiltonian systems: the dynamic interpretation of nonlinear stability on a reduced phase space*, Journal of Geometry and Physics, **9**(2), 111–119.

Sasaki, S. [1958] *On the differential geometry of tangent bundles of Riemannian manifolds*, The Tôhoku Mathematical Journal. Second Series, **10**, 338–354.

— [1962] *On the differential geometry of tangent bundles of Riemannian manifolds II*, The Tôhoku Mathematical Journal. Second Series, **14**, 146–155.

Simo, J. C., Lewis, D. R., and Marsden, J. E. [1991] *Stability of relative equilibria I: The reduced energy momentum method*, Archive for Rational Mechanics and Analysis, **115**(1), 15–59.

Simo, J. C., Posbergh, T. A., and Marsden, J. E. [1990] *Stability of coupled rigid bodies and geometrically exact rods—block diagonalisation and the energy momentum method*, Physics Reports. A Review Section of Physics Letters, **193**(6), 279–360.

Sussmann, H. J. [1997] *An introduction to the coordinate-free maximum principle*, in *Geometry of Feedback and Optimal Control*, B. Jakubczyk and W. Respondek, editors, pages 463–557, Dekker Marcel Dekker, New York.

Synge, J. L. [1928] *Geodesics in nonholonomic geometry*, Mathematische Annalen, **99**, 738–751.

Vorotnikov, V. I. [1998] *Partial Stability and Control*, Birkhäuser, Boston/Basel/Stuttgart, ISBN 0-8176-3917-9.

Yano, K. and Ishihara, S. [1973] *Tangent and Cotangent Bundles*, number 16 in Pure and Applied Mathematics, Dekker Marcel Dekker, New York, ISBN 0-8247-6071-9.

Francesco Bullo
Associate Professor
Mechanical & Environmental Eng.
Univ. of California at Santa Barbara
Engineering II Bldg.
Santa Barbara, CA 93106-5070, U.S.A.
bullo@engineering.ucsb.edu

Andrew D. Lewis
Associate Professor
Dep. of Math. and Statistics
Queen's University
Kingston, ON K7L 3N6, Canada
andrew@mast.queensu.ca