The linearisation of a simple mechanical control system*

Francesco Bullo† Andrew D. Lewis‡

27/06/2002

Abstract

A geometric interpretation is given for the linearisation of a mechanical control system with a kinetic minus potential energy Lagrangian.

Keywords. Mechanical systems, linearisation, controllability

AMS Subject Classifications. 70Q05, 93B18, 93B05

1. Introduction

In an earlier paper [Lewis and Murray 1997], one of the authors had investigated the character of the control Lie algebra for simple mechanical systems, i.e., those with kinetic energy minus potential energy Lagrangians, and with forces whose directions vary only as functions of configuration. A special case that was *not* considered in this early work was that when the linearisation of the system was actually controllable. In this paper, this special case is considered, and as a consequence, the usual controllability condition for linear systems is given an interpretation involving the geometry inherent in simple mechanical control systems. Other work on controllability of the linearisation of mechanical systems includes [Hughes and Skelton 1980] and [Laub and Arnold 1985]. In both of these papers, the emphasis was on the algebraic, rather than the geometric structure of the linearisation.

2. Results

In this section the problem is setup, and the main result is stated and proved. In the paper, all data is assumed to be of class C^∞, unless otherwise stated. The following notation will be convenient. The dual of a $\mathbb{R}$-vector space V is denoted V^*. If V and W are $\mathbb{R}$-vector spaces, $L(V;W)$ is the set of $\mathbb{R}$-linear maps from V to W. If B is a symmetric $(0,2)$-tensor on a vector space V, $B^\flat \in L(V;V^*)$ is defined by $B^\flat(v_1)\cdot v_2 = B(v_1,v_2)$. If B is nondegenerate then $B^\flat$ is invertible, and its inverse is denoted $B^\sharp \in L(V^*;V)$. Given a linear map $A \in L(V;V)$ on a vector space V and a subspace $U \subset V$, let $\langle A, U\rangle$ denote the smallest A-invariant subspace containing U. The identity map on a set S is denoted id_S.

*Preprint

†Research Assistant Professor, Coordinated Science Laboratory, University of Illinois, Urbana-Champaign, Urbana-Champaign, IL 61801, U.S.A.
Email: `bullo@uiuc.edu`, URL: `http://motion.csl.uiuc.edu/~bullo/`
Research supported in part by NSF grants IIS-0118146 and CMS-0100162.
‡Assistant Professor, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
Email: `andrew@mast.queensu.ca`, URL: `http://www.mast.queensu.ca/~andrew/`
Research supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

2.1. Problem setup. We consider a simple mechanical control system $\Sigma = (Q, g, V, \mathscr{F} = \{F^1,\dots,F^m\})$ characterised by

1. a configuration manifold Q,
2. a Riemannian metric g which defines the kinetic energy via $\mathrm{KE}(v_q) = \frac{1}{2}g(v_q,v_q)$,
3. a potential function V defining the potential energy via $\mathrm{PE}(q) = V(q)$, and
4. a collection of one-forms $\mathscr{F}$ on Q specifying the control forces for the system.

As in [Lewis and Murray 1997], this defines the governing second-order equations as

$$\nabla_{\gamma'(t)}\gamma'(t) = -\operatorname{grad} V(\gamma(t)) + \sum_{a=1}^m u^a(t) Y_a(\gamma(t)), \tag{2.1}$$

where ∇ is the Levi-Civita connection for g, where $\operatorname{grad} V = g^\sharp(\boldsymbol{d}V)$ is the gradient of V ($\boldsymbol{d}V$ is the usual differential of V), and where $Y_a = g^\sharp(F^a)$, $a \in \{1,\dots,m\}$, are the generators for the inputs, now written as vector fields rather than one-forms. The equations (2.1) are second-order differential equations on Q, when written in coordinates. These may also be thought of as first-order equations on the tangent bundle TQ, and upon doing this one arrives at the control-affine system

$$\Upsilon'(t) = Z(\Upsilon(t)) - \mathrm{vlft}(\operatorname{grad} V)(\Upsilon(t)) + \sum_{a=1}^m u^a(t)\,\mathrm{vlft}(Y_a)(\Upsilon(t)), \tag{2.2}$$

where Z is the geodesic spray for ∇, and where vlft is the vertical lift map, sending vector fields on Q to vector fields on TQ by $\mathrm{vlft}(X)(v_q) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}(v_q + tX(q))$. We refer to [Lewis and Murray 1997] for details on the setting up of these equations, and to [Kobayashi and Nomizu 1963] for background on affine differential geometry.

Note that (2.2) is a control-affine system on $M = TQ$ with drift vector field $f_0 = Z - \mathrm{vlft}(\operatorname{grad} V)$ and with control vector fields $f_a = \mathrm{vlft}(Y_a)$, $a \in \{1,\dots,m\}$. A point $v_{q_0} \in TQ$ is an equilibrium point for the drift vector field if and only if $v_{q_0} = 0_{q_0}$, the zero vector in the tangent space $T_{q_0}Q$, and $\boldsymbol{d}V(q_0) = 0$ (i.e., q_0 is a critical point for V). Since $\boldsymbol{d}V(q_0) = 0$ at an equilibrium point q_0, the Hessian, $\mathrm{Hess}(V)(q_0)$, is well-defined as a symmetric $(0,2)$-tensor on $T_{q_0}Q$. In coordinates, it is represented by the matrix of second partial derivatives of V with respect to the coordinates. We will consider the linearisation of (2.2) about an equilibrium point 0_{q_0}, and ascertain when the linearisation is controllable. To represent the linearisation, we note that the tangent space $T_{0_{q_0}}TQ$ at the equilibrium point admits a canonical decomposition as $T_{0_{q_0}}TQ \simeq T_{q_0}Q \oplus T_{q_0}Q$. The first component in this decomposition is the tangent space of the zero section of TQ, thought of as a submanifold passing through 0_{q_0} and being naturally diffeomorphic to Q. The second component in the decomposition is the tangent space to the fibre $T_{q_0}Q$, again thought of as a submanifold of TQ passing through 0_{q_0}. These submanifolds are clearly transverse, and so the decomposition is a direct sum, as asserted. It will therefore be convenient to write vectors in $T_{0_{q_0}}TQ$ and linear transformations of $T_{0_{q_0}}TQ$ in terms of this decomposition. We shall write these objects as

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad\text{and}\quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

respectively, where $v_1, v_2 \in T_{q_0}Q$ and $A_{ij} \in L(T_{q_0}Q; T_{q_0}Q)$, $i, j \in \{1, 2\}$. With this notation, the following result may be proved by a direct computation in coordinates.

2.1 LEMMA: *The linearisation of (2.2) about an equilibrium point 0_{q_0} is the system*

$$\dot{x}(t) = A_\Sigma(q_0)(x(t)) + B_\Sigma(q_0)(u(t)),$$

where $A_\Sigma(q_0) \in L(T_{0_{q_0}}TQ; T_{0_{q_0}}TQ)$ and $B_\Sigma(q_0) \in L(\mathbb{R}^m; T_{0_{q_0}}TQ)$ are given by

$$A_\Sigma(q_0) = \begin{bmatrix} 0 & \mathrm{id}_{T_{q_0}Q} \\ -K_\Sigma(q_0) & 0 \end{bmatrix}, \quad B_\Sigma(q_0)(u) = \sum_{a=1}^m u^a \begin{bmatrix} 0 \\ Y_a(q_0) \end{bmatrix},$$

with $K_\Sigma(q_0) = g(q_0)^\sharp \circ \mathrm{Hess}V(q_0)^\flat$.

We denote the linear map from $\mathbb{R}^m$ to $T_{q_0}Q$ given by $u \mapsto \sum_{a=1}^m u^a Y_a(q_0)$ by $Y_\Sigma(q_0)$.

2.2. Main result. With the preceding notation in place, we may now formulate our main result. Since the main contribution is the geometric interpretation, we first need to setup the language with which to provide this interpretation. We refer to [Kobayashi and Nomizu 1963] for the necessary background on affine differential geometry. Given a distribution $\mathcal{D}$ on Q, smooth but not necessarily of constant rank, $\Gamma^\infty(\mathcal{D})$ denotes the collection of vector fields taking values in $\mathcal{D}$. In particular, $\Gamma^\infty(TQ)$ denotes the collection of vector fields on Q. If $\mathcal{D}$ is a distribution on Q, and if $X_0 \in \Gamma^\infty(TQ)$, then X_0 is ***$\mathcal{D}$-parallel*** if $\nabla_X X_0 \in \Gamma^\infty(\mathcal{D})$ for each $X \in \Gamma^\infty(\mathcal{D})$. For an affine connection ∇ on Q the ***symmetric product*** assigns to each pair of vector fields X and Y on Q the vector field $\langle X : Y\rangle = \nabla_X Y + \nabla_Y X$. For $X \in \Gamma^\infty(TQ)$ let ∇X be the $(1,1)$-tensor field on Q defined by $\nabla X(Y) = \nabla_Y X$ for $Y \in \Gamma^\infty(TQ)$. This tensor field is the ***covariant differential*** of X.

If a vector field X vanishes at $q_0 \in Q$, then the map $Y \mapsto \langle X : Y\rangle (q_0)$ from the set of vector fields on Q to $T_{q_0}Q$ depends only on the value of Y at q_0, and not on the derivatives of Y. Therefore we may define a map $\mathrm{sym}_X : T_{q_0}Q \to T_{q_0}Q$ by $\mathrm{sym}_X(v) = \langle X : V\rangle (q_0)$, where V is any vector field extending $v \in T_{q_0}Q$.

We now state our main result. In the statement of the result, $\mathcal{Y}$ is the distribution generated by the family of vector fields $\mathscr{Y}$.

2.2 PROPOSITION: *Let $\Sigma = (Q, g, V, \mathscr{F})$ be a simple mechanical control system with $\dim(Q) = n$, let q_0 be a critical point for V, and consider the linearisation of Σ at q_0 as given in Lemma 2.1. Let $\mathscr{Y} = \{g^\sharp(F^1), \dots, g^\sharp(F^m)\}$. The following statements are equivalent:*

(i) the linearisation of Σ at q_0 is controllable;

(ii) the linear map from $\oplus_{i=1}^n \mathbb{R}^m$ to $T_{q_0}Q$ represented by the block matrix

$$\left[\ Y_\Sigma(q_0) \ \middle|\ K_\Sigma(q_0) \circ Y_\Sigma(q_0) \ \middle|\ \cdots \ \middle|\ K_\Sigma^{n-1}(q_0) \circ Y_\Sigma(q_0)\ \right],$$

has full rank;

(iii) $\langle \mathrm{sym}_{\mathrm{grad}\,V}, \mathcal{Y}_{q_0}\rangle = T_{q_0}Q$;

(iv) $\langle \nabla \mathrm{grad}\,V, \mathcal{Y}_{q_0}\rangle = T_{q_0}Q$;

(v) if $\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y})$ is the smallest distribution containing $\mathcal{Y}$ for which $\mathrm{grad}\,V$ is $\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y})$-parallel, then $\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y})_{q_0} = T_{q_0}Q$.

Proof: (i) $\iff$ (ii) Since the linearisation is controllable, the linear map represented by the block matrix

$$\left[\ B_\Sigma(q_0) \ \middle|\ A_\Sigma(q_0) \circ B_\Sigma(q_0) \ \middle|\ \cdots \ \middle|\ A_\Sigma^{2n-1}(q_0) \circ B_\Sigma(q_0)\ \right] \tag{2.3}$$

must have full rank, where $A_\Sigma(q_0)$ and $B_\Sigma(q_0)$ are as in Lemma 2.1. A direct computation shows that (2.3), when written using the decomposition $T_{0_{q_0}}TQ \simeq T_{q_0}Q \oplus T_{q_0}Q$, is given by

$$\begin{bmatrix} 0 & Y_\Sigma & 0 & -K_\Sigma \circ Y_\Sigma & \cdots & 0 & (-1)^{n-1}K_\Sigma^{n-1} \circ Y_\Sigma \\ Y_\Sigma & 0 & -K_\Sigma \circ Y_\Sigma & 0 & \cdots & (-1)^{n-1}K_\Sigma^{n-1} \circ Y_\Sigma & 0 \end{bmatrix},$$

where we abbreviate $K_\Sigma(q_0)$ and $Y_\Sigma(q_0)$ with K_Σ and Y_Σ, respectively. This part of the lemma now follows directly from this representation of (2.3).

(ii) $\iff$ (iii) Let us obtain an expression for the linear map $\mathrm{sym}_{\mathrm{grad}\,V}$. Let $X \in \Gamma^\infty(TQ)$ and let $(q^1, \dots, q^n)$ be coordinates around q_0. In these coordinates we compute

$$\begin{aligned} \langle \mathrm{grad}\,V : X\rangle &= \Big(\frac{\partial (\mathrm{grad}\,V)^k}{\partial q^i} X^i + \frac{\partial X^k}{\partial q^i} (\mathrm{grad}\,V)^i + 2\Gamma^k_{ij} (\mathrm{grad}\,V)^i X^j\Big) \frac{\partial}{\partial q^k} \\ &= \Big(\frac{\partial g^{kj}}{\partial q^i} \frac{\partial V}{\partial q^j} X^i + g^{kj} \frac{\partial^2 V}{\partial q^i \partial q^j} X^i + \frac{\partial X^k}{\partial q^i} g^{ij} \frac{\partial V}{\partial q^j} + 2\Gamma^k_{ij} g^{i\ell} \frac{\partial V}{\partial q^\ell} X^j\Big) \frac{\partial}{\partial q^k}. \end{aligned}$$

where Γ^k_{ij}, $i, j, k \in \{1, \dots, n\}$ are the Christoffel symbols for the Levi-Civita connection, and where we have used the expression $\mathrm{grad}\,V^i = g^{ij}\frac{\partial V}{\partial q^j}$. Evaluating at q_0, noting that $\boldsymbol{d}V(q_0) = 0$, gives $\mathrm{sym}_{\mathrm{grad}\,V}(X(q_0)) = g(q_0)^\sharp \circ \mathrm{Hess}V(q_0)^\flat(X(q_0))$. That is, $\mathrm{sym}_{\mathrm{grad}\,V} = K_\Sigma(q_0)$. Now $\langle \mathrm{sym}_{\mathrm{grad}\,V}, \mathcal{Y}_{q_0}\rangle = T_{q_0}Q$ if and only if the vectors

$$\mathrm{sym}^{k-1}_{\mathrm{grad}\,V}(Y_a(q_0)), \qquad a \in \{1, \dots, m\},\ k \in \mathbb{N},$$

generate $T_{q_0}Q$. In the usual way, given the Cayley-Hamilton theorem, for $k \geq n$ the linear map $\mathrm{sym}^{k-1}_{\mathrm{grad}\,V}$ is a linear combination of $\mathrm{id}_{T_{q_0}Q}, \mathrm{sym}_{\mathrm{grad}\,V}, \dots, \mathrm{sym}^{n-1}_{\mathrm{grad}\,V}$, and thus this part of the result follows.

(iii) $\iff$ (iv) Since $\mathrm{grad}\,V(q_0) = 0_{q_0}$ one readily verifies from the coordinate computation of the preceding part of the proof that $\langle \mathrm{grad}\,V : X\rangle (q_0) = \nabla_X \mathrm{grad}\,V(q_0)$ for any $X \in \Gamma^\infty(TQ)$. Therefore, $\mathrm{sym}^k_{\mathrm{grad}\,V} = \nabla \mathrm{grad}\,V^k(q_0)$, and this part of the result follows.

(iv) $\implies$ (v) Since $\mathrm{grad}\,V$ is $\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y})$-parallel and since $\mathscr{Y} \subset \Gamma^\infty(\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y}))$, it follows that $\nabla_{Y_a} \mathrm{grad}\,V \in \Gamma^\infty(\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y}))$, $a \in \{1, \dots, m\}$. That is, $\nabla \mathrm{grad}\,V(Y_a) \in \Gamma^\infty(\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y}))$. This argument may be repeated to conclude that $\nabla_{\nabla_{Y_a}\mathrm{grad}\,V} \mathrm{grad}\,V = \nabla \mathrm{grad}\,V^2(Y_a) \in \Gamma^\infty(\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y}))$, $a \in \{1, \dots, m\}$. Carrying on, we deduce inductively that $\nabla \mathrm{grad}\,V^k(Y_a) \in \Gamma^\infty(\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y}))$, $k \in \mathbb{N}$, $a \in \{1, \dots, m\}$. Assuming (iv) we immediately see that $\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y})_{q_0} = T_{q_0}Q$.

(v) $\implies$ (iv) By definition, $\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y})$ is generated by the vector fields

$$\nabla_{Y_a} \mathrm{grad}\,V,\ \nabla_{\nabla_{Y_a}\mathrm{grad}\,V} \mathrm{grad}\,V, \dots \qquad a \in \{1, \dots, m\}.$$

Working backwards through the argument of the preceding part of the proof we see that $\mathcal{P}(\mathrm{grad}\,V, \mathscr{Y})_{q_0} = T_{q_0}Q$ implies that

$$\langle \nabla \mathrm{grad}\,V, \mathcal{Y}_{q_0}\rangle = T_{q_0}Q.$$

This part of the result now follows. ■

2.3 Remarks: 1. The equivalence of (i) and (ii) is one of the results proved by Hughes and Skelton [1980]. They also consider some special cases of gyroscopic and damping terms in the linearisation.

2. The vector field grad V may be replaced an arbitrary vector field F for which q_0 is an equilibrium point. In this case the result holds provided one replaces $K_\Sigma(q_0)$ with the linearisation of F at q_0.

3. When $V = 0$ we see that $\mathcal{P}(\mathrm{grad}\, V, \mathscr{Y})$ is simply the distribution generated by the vector fields $\mathscr{Y}$. In this case we see that the linearisation is controllable if and only if $\mathrm{span}_{\mathbb{R}}\{Y_1(q_0), \ldots, Y_m(q_0)\} = T_{q_0}Q$, i.e., if and only if the system is fully actuated.

4. When the linearisation is not controllable, e.g., when $V = 0$, one must consider the effects of brackets other than $\mathrm{ad}^k_{f_0} f_a$ in the accessibility algebra. For simple mechanical control systems, an algorithm for doing this, using symmetric products and Lie brackets of the vector fields $\{Y_1, \ldots, Y_m, \mathrm{grad}\, V\}$, is presented by Lewis and Murray [1997]. However, there is presently no geometric interpretation of the resulting distribution, such as we provide here for the linearly controllable case. □

References

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.

Kobayashi, S. and Nomizu, K. [1963] *Foundations of Differential Geometry,* Volume I, number 15 in Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers, New York, ISBN 0-470-49647-9.

Laub, A. J. and Arnold, W. F. [1985] *Controllability and observability criteria for multivariable linear second-order models*, Institute of Electrical and Electronics Engineers. Transactions on Automatic Control, **29**(2), 163–165.

Lewis, A. D. and Murray, R. M. [1997] *Controllability of simple mechanical control systems*, SIAM Journal on Control and Optimization, **35**(3), 766–790.