Variational and nonholonomic mechanics

Andrew D. Lewis

Department of Mathematics and Statistics
Queen's University, Kingston, ON, Canada

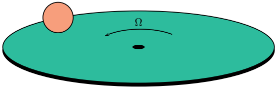
In honour of Manuel de León's 70th birthday
11/12/2023

The problem

Problem (Imprecise version)

When taking a variational approach to the equations of motion in mechanics, do you apply the constraints before or after specifying the variations?

My initial foray:[1]



- Studied two ways of modelling the system with nonholonomic constraints
- Did some friction modelling
- Compared numerics and data collected from a sophisticated experiment apparatus[2]

---

[1] L/Murray, *Int. J. Nonlinear Mech.*, **30**(6), 793–815, 1995

[2] A stereo turntable, a ping-pong ball, and a VHS camera

Other, better, work and a return to the problem

- Some other contributions in the intervening years:
  1. Cardin/Favretti, *J. Geom. Phys.*, **18**(4), 295–325, 1996
  2. Favretti, *J. Dyn. Diff. Eq.*, **10**(4), 511-536, 1998
  3. Zampieri, *J. Diff. Eq.*, **163**(2), 335-347, 2000
  4. de León/Marrero/Martin de Diego, *J. Geom. Phys.*, **35**(2–3), 126–144
  5. Kupka/Oliva, *J. Diff. Eq.*, **169**(1), 169–189, 2001
  6. Cortés/de León/Martín de Diego/Martínez, *SIAM J. Control Optim.*, **41**(5), 1389–1412, 2002
  7. Fernandez/Bloch, *J. Phys. A*, **41**(3), no. 344005, 2008
  8. Borisov/Mamaev/Bizyaev, *Russian Math. Surveys*, **72**(5), 783–840, 2017
  9. Terra, *São Paulo J. Math. Sci.*, **12**(1), 136-145, 2018
  10. Jóźwikowski/Respondek, *J. Geom. Mech.*, **11**(1), 77-122, 2019
- After 25 years. . . a revisitation.[3]

[3] L, *J. Geom. Mech.*, **12**(2), 165–308, 2020

Two problems

- Data:
  1. configuration manifold Q of regularity $r \in \{\infty, \omega\}$;
  2. kinetic energy modelled by a C^r-Riemannian metric $\mathbb{G}$;
  3. potential energy modelled by a potential function $V \in \mathrm{C}^r(\mathsf{Q})$;
  4. nonholonomic constraints modelled by a C^r-distribution $\mathsf{D} \subseteq \mathsf{TQ}$.
- Action for a Lagrangian $L\colon \mathsf{TQ} \to \mathbb{R}$ is

$$A_L(\gamma) = \int_{t_0}^{t_1} L \circ \gamma'(t)\,\mathrm{d}t,$$

for $\gamma \in \mathrm{H}^1([t_0, t_1]; \mathsf{Q}; q_0, q_1)$.

Problem (Nonholonomic (N))

Find $\gamma \in \mathrm{H}^1([t_0, t_1]; \mathsf{Q}; \mathsf{D}; q_0, q_1)$ *such that*

$$\langle \mathrm{d}(A_{\mathbb{G}} - A_V); \delta \rangle = 0, \quad \delta \in \mathrm{H}^1([t_0, t_1]; \gamma^*\mathsf{D}; q_0, q_1).$$

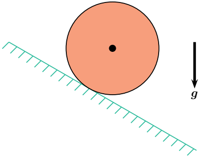
Problem (Variational (V))

Find $\gamma \in \mathrm{H}^1([t_0, t_1]; \mathsf{Q}; \mathsf{D}; q_0, q_1)$ *such that*

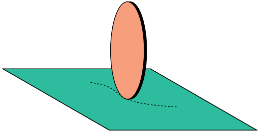
$$\langle \mathrm{d}(A_{\mathbb{G},\mathsf{D}} - A_{V,\mathsf{D}}); \delta\sigma(0) \rangle = 0,$$

$$\sigma\colon (-\epsilon, \epsilon) \to \mathrm{H}^1([t_0, t_1]; \mathsf{Q}; \mathsf{D}; q_0, q_1).$$

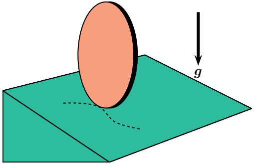
Three examples



All solutions to Problems (N) and (V) give the same physical motions.



For every solution to Problem (N), *there exists* a solution to Problem (V) giving the same physical motion.



For *almost no* solution to Problem (N) does there exist a solution to Problem (V) that gives the same physical motion.[a]

[a] Lemos, *Acta Mech.*, **233**, 47-56, 2022

Problem Statement

Problem (Vague version)

Characterise the set of initial conditions for which the solution to Problems (N) and (V) gives the same physical motion.

- To make this vague statement more precise, one needs to understand more about the solutions to Problems (N) and (V).

Some connections and tensors

Given a Riemannian manifold $(\mathsf{Q}, \mathbb{G})$ and a distribution D:

1. ***Levi-Civita connection:*** $\nabla^{\mathbb{G}}$
2. $\mathbb{G}$-orthogonal projections P_{D} and $P_{\mathsf{D}^\perp}$
3. ***constrained connection:*** ∇^{D} (project $\nabla^{\mathbb{G}}$ onto D)
4. ***Fröbenius curvature:*** $F_{\mathsf{D}}(X, Y) = P_{\mathsf{D}^\perp}([X, Y]) \quad (X, Y \in \Gamma^r(\mathsf{D}))$
5. ***geodesic curvature:*** $G_{\mathsf{D}}(X, Y) = P_{\mathsf{D}^\perp}(\nabla^{\mathbb{G}}_X Y + \nabla^{\mathbb{G}}_Y X) \quad (X, Y \in \Gamma^r(\mathsf{D}))$
6. ***second fundamental form:*** $S_D(X, Y) = -(\nabla^{\mathbb{G}}_X P_{\mathsf{D}^\perp})(Y)$
   $(X \in \Gamma^r(\mathsf{TM}),\ Y \in \Gamma^r(\mathsf{D}))$

Solutions to Problem (N)

Theorem

For $\gamma \in \mathsf{H}^1([t_0, t_1]; \mathsf{Q}; \mathsf{D}; q_0, q_1)$*, the following statements are equivalent:*

1. γ *is a solution to Problem (N);*
2. $\gamma \in \mathsf{H}^2([t_0, t_1]; \mathsf{Q})$ *and there exists* $\lambda \in \mathsf{L}^2([t_0, t_1]; \gamma^*\mathsf{D}^\perp)$ *such that*

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

3. $\gamma \in \mathsf{H}^2([t_0, t_1]; \mathsf{Q})$ *and satisfies*

$$\nabla^{\mathsf{D}}_{\gamma'}\gamma' + P_{\mathsf{D}} \circ \operatorname{grad} V \circ \gamma = 0.$$

Solutions to Problem (V)

Theorem

For $\gamma \in \mathsf{H}^1([t_0,t_1];\mathsf{Q};\mathsf{D};q_0,q_1)$*, the following statements are equivalent:*

1. γ *is a solution to Problem (V);*
2. *at least one of the following holds:*
   a. *some interesting condition for singular extremals that I will ignore, sacrificing correctness for expediency;*
   b. $\gamma \in \mathsf{H}^2([t_0,t_1];\mathsf{Q})$ *and there exists* $\lambda \in \mathsf{H}^1([t_0,t_1];\gamma^*\mathsf{D}^\perp)$ *such that*

$$\nabla^{\mathbb{G}}_{\gamma'}\gamma' + \operatorname{grad} V \circ \gamma - \nabla^{\mathbb{G}}_{\gamma'}\lambda - S^*_{\mathsf{D}}(\gamma')(\lambda) = 0;$$ [a]

3. *at least one of the following holds:*
   a. *some other interesting condition for singular extremals that I will again ignore;*
   b. $\gamma \in \mathsf{H}^2([t_0,t_1];\mathsf{Q})$ *and there exists* $\lambda \in \mathsf{H}^1([t_0,t_1];\gamma^*\mathsf{D}^\perp)$ *such that*

$$\nabla^{\mathsf{D}}_{\gamma'}\gamma' + P_{\mathsf{D}} \circ \operatorname{grad} V \circ \gamma = F^*_{\mathsf{D}}(\gamma')(\lambda),$$

$$\nabla^{\mathsf{D}^\perp}_{\gamma'}\lambda = \frac{1}{2}G_{\mathsf{D}}(\gamma',\gamma') + P_{\mathsf{D}^\perp} \circ \operatorname{grad} V \circ \gamma + \frac{1}{2}G^\star_{\mathsf{D}^\perp}(\gamma')(\lambda) + \frac{1}{2}F^\star_{\mathsf{D}^\perp}(\gamma')(\lambda).$$

---

[a] Kupka/Oliva, *J. Diff. Equations*, **169**(1), 169–189, 2001

The crucial observation

Compare

$$\nabla^{\mathsf{D}}_{\gamma'}\gamma' + P_{\mathsf{D}} \circ \operatorname{grad} V \circ \gamma = 0$$

with

$$\begin{aligned} &\nabla^{\mathsf{D}}_{\gamma'}\gamma' + P_{\mathsf{D}} \circ \operatorname{grad} V \circ \gamma = F^*_{\mathsf{D}}(\gamma')(\lambda), \\ &\nabla^{\mathsf{D}^\perp}_{\gamma'}\lambda = \frac{1}{2}G_{\mathsf{D}}(\gamma',\gamma') + P_{\mathsf{D}^\perp} \circ \operatorname{grad} V \circ \gamma + \frac{1}{2}G^\star_{\mathsf{D}^\perp}(\gamma')(\lambda) + \frac{1}{2}F^\star_{\mathsf{D}^\perp}(\gamma')(\lambda). \end{aligned} \tag{1}$$

Problem

Given a physical motion $t \mapsto \gamma(t)$ satisfying Problem (N), find all (if any) initial conditions for λ so that the resulting solution to (1) *is such that $F^*_{\mathsf{D}}(\gamma')(\lambda) = 0$.*

Symbolic abstraction

If we think of γ as given, the equation (1) for λ has the form of an affine differential equation,

$$\underbrace{\nabla^{\mathsf{D}^\perp}_{\gamma'}\lambda}_{\dot{\lambda}(t)} = \underbrace{\frac{1}{2}G^\star_{\mathsf{D}^\perp}(\gamma')(\lambda) + \frac{1}{2}F^\star_{\mathsf{D}^\perp}(\gamma')(\lambda)}_{A(t)(\lambda(t))} + \underbrace{\frac{1}{2}G_{\mathsf{D}}(\gamma',\gamma') + P_{\mathsf{D}^\perp}\circ\operatorname{grad} V\circ\gamma}_{b(t)},$$

and the condition satisfied by λ is a “satisfies a linear equation” condition,

$$\underbrace{F^*_{\mathsf{D}}(\gamma')(\lambda) = 0}_{B(t)(\lambda(t))=0}.$$

Problem (In symbols)

Find all solutions of the affine differential equation

$$\dot{\lambda}(t) = A(t)(\lambda(t)) + b(t)$$

satisfying $B(t)(\lambda(t)) = 0.$

Complete abstraction

We have the following data:

1. a C^r-vector bundle $\pi: \mathsf{E} \to \mathsf{M}$ (abstracting the pull-back bundle $\pi_{\mathsf{D}}^* \mathsf{D}^\perp \to \mathsf{D}$);
2. a C^r-cogeneralised subbundle $\mathsf{F} \subseteq \mathsf{E}$ (abstracting $\ker(B)$);
3. a C^r-affine vector field X on E (abstracting $\dot{\lambda} = A \circ \lambda + b$) over a C^r-vector field X_0 on M (abstracting the equations governing Problem (N)).

Representation of an affine vector field X on E over a vector field X_0 on M.

- Assuming a linear connection ∇ on E, we can write

$$X = X_0^{\mathsf{h}} + A^{\mathsf{e}} + b^{\mathsf{v}}$$

  for $A \in \Gamma^r(\mathsf{E} \otimes \mathsf{E}^*)$ and $b \in \Gamma^r(\mathsf{E})$ where $\cdot^{\mathsf{h}}$ is horizontal lift, $\cdot^{\mathsf{e}}$ is "vertical evaluation," and $\cdot^{\mathsf{v}}$ is vertical lift.
- Then, for a curve $\Upsilon: I \to \mathsf{E}$, the following are equivalent:
  1. Υ is an integral curve for X;
  2. $\nabla_{\gamma'}\Upsilon = A \circ \Upsilon + b \circ \gamma$, where $\gamma = \pi \circ \Upsilon$.

A general problem

Given:

1. an affine vector field X on a vector bundle $\pi \colon \mathsf{E} \to \mathsf{M}$ over a vector field X_0 on M and
2. a C^r-cogeneralised subbundle $\mathsf{F} \subseteq \mathsf{E}$.

Problem (General Geometric Problem)

Find all initial conditions $e \in \mathsf{E}$ such that the associated integral curve Υ of X through e satisfies $\mathrm{image}(\Upsilon) \subseteq \mathsf{F}$.

Dualise using elementary linear algebra

Proposition (Linear equations)

Let V *be a finite-dimensional* $\mathbb{R}$*-vector space. There is a 1–1 correspondence between the sets of solutions of linear equations*

$$A(v) + b = 0, \qquad A \in \mathsf{End}_{\mathbb{R}}(\mathsf{V}),\ b \in \mathsf{V},$$

and subspaces $\Delta \subseteq \mathsf{V}^* \oplus \mathbb{R}$ *with positive codimension. Moreover,*

1. *the set of solutions to the linear equation is nonempty if and only if* $(0,1) \in \Delta$ *and*
2. *the set of solutions is* $\{v \in \mathsf{V} \mid (v,1) \in \Lambda(\Delta)\}$ *with* Λ *being the annihilator.*

Making the dual point of view geometric

- For a vector bundle $\pi\colon \mathsf{E} \to \mathsf{M}$, we think of subbundles of positive codimension of $\mathsf{E}^* \oplus \mathbb{R}_{\mathsf{M}}$ as being bundles of linear equations, and call them ***defining subbundles***.
- Let $\Delta \subseteq \mathsf{E}^* \oplus \mathbb{R}_{\mathsf{M}}$ be a defining subbundle. Call the set of solutions

$$\mathsf{A}(\Delta) = \{e \in \mathsf{E} \mid \langle \lambda; e\rangle + a = 0,\ (\lambda, a) \in \Delta_{\pi(e)}\}$$

  an ***affine subbundle variety***.
- For an affine subbundle variety $\mathsf{A}(\Delta)$, we have

$$\mathsf{S}(\mathsf{A}) = \{x \in \mathsf{M} \mid \mathsf{A} \cap \mathsf{E}_x \neq \varnothing\},$$

  called the ***base variety***.

Making the dual point of view geometric (cont'd)

- The connection ∇ on E and the flat connection on $\mathbb{R}_{\mathsf{M}}$ induce a connection $\widehat{\nabla}$ on $\mathsf{E} \oplus \mathbb{R}_{\mathsf{M}}$.
- The affine vector field $X = X_0^{\mathsf{h}} + A^{\mathsf{e}} + b^{\mathsf{v}}$ induces a *linear* vector field

$$\widehat{X} = X_0^{\mathsf{h}} + (A, b)^{\mathsf{e}}$$

in $\mathsf{E} \oplus \mathbb{R}_{\mathsf{M}}$ and the dual vector field $\widehat{X}^*$ in $\mathsf{E}^* \oplus \mathbb{R}_{\mathsf{M}}$.

Proposition

Let Δ be a defining subbundle and let $\mathsf{A}(\Delta)$ be the associated affine subbundle variety. TFAE (morally):

1. $\mathsf{A}(\Delta)$ *is invariant under* X*;*
2.
   a. $\mathsf{S}(\mathsf{A}(\Delta))$ *is invariant under* X_0 *and*
   b. $\Delta \cap (\pi^* \times \mathrm{pr}_1)^{-1}(\mathsf{S}(\mathsf{A}(\Delta)))$ *is invariant under* $\widehat{X}^*$*.*

Results

Theorem

Assume X_0 is complete.
The set of initial conditions A *of the General Geometric Problem is the largest affine subbundle variety contained in* F *and invariant under* X*.*

There's more... infinitesimal conditions satisfied by "A"... a PDE one can analyse à la Spencer... connections to sub-Riemannian geometry... application to the original mechanics problem...