Variational integrators and nonholonomic integrators

Andrei Tsyganov

Yanqi Lake Beijing Institute of Mathematical Sciences and Applications,

2024.

It is well known that discrete Hamilton's variational principle is one of the structure-preserving numerical integrators, providing a long-time numerically stable scheme for conservative Lagrangian systems.

This principle is based on the fact that the discrete Euler–Lagrange equations preserve a discrete symplectic structure, called the discrete Lagrangian two-form, along the discrete Lagrangian map, which is the discrete analogue of the flow of the Lagrangian vector field in the continuous setting.

On the other hand, when nonholonomic constraints are present, as in nonholonomic mechanical systems, the associated equations of motion can be derived using the Lagrange–d'Alembert principle.

In the discrete setting, discrete Hamilton's principle is replaced by the discrete Lagrange–d'Alembert principle, from which the discrete Lagrange–d'Alembert equations can be obtained.

These algorithms, which generalize variational integrators for unconstrained Lagrangian systems, exhibit geometric properties similar to those of continuous nonholonomic systems.

From a slightly different perspective, we can consider nonholonomic systems that admit reversing symmetries and developed integrators for such systems that preserve an analogous reversing symmetry.

In these cases, the numerical integrator, referred to as a ”nonholonomic integrator,” no longer preserves the discrete symplectic structure.

In fact, it is not yet fully understood what structure, if any, is preserved by such a nonholonomic integrator.

Nevertheless, it remains a long-term numerically stable scheme for conservative nonholonomic mechanical systems.

Hamilton's variational principle

Consider a mechanical system with a Lagrangian

$$L : TQ \to \mathbb{R},$$

where TQ is the tangent bundle of an n-dimensional configuration manifold Q with local coordinates q^i, $i = 1, ..., n$ for $q \in Q$.

Consider a path space

$$\mathcal{C}(Q) = \{q : I = [0, T] \to Q \mid q \text{ is a } C^2 \text{ curve on } Q \text{ such that } q(0) = q_1 \text{ and } q(T) = q_2\},$$

where $I = [0, T] \subset \mathbb{R}^+$ is the space of time.

A point q in the manifold $\mathcal{C}(Q)$ is a curve on Q, namely, $q = q(t)$.

The deformation of $q = q(t) \in \mathcal{C}(Q)$ is given by $q(t, \epsilon) = q_\epsilon(t)$ such that $q_0(t) = q(t, 0) = q(t)$.

Then, the variation of the curve $q(t)$ is defined by

$$\delta q(t) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} q_\epsilon(t),$$

which is the tangent vector to a curve $q(t)$.

Let

$$\tau_Q : TQ \to Q; (q, \delta q) \mapsto q$$

be the canonical projection and we get $\tau_Q \circ \delta q = q$.

The restrictions $q_\epsilon(0) = q_1$ and $q_\epsilon(T) = q_2$ lead to $\delta q(0) = 0$ and $\delta q(T) = 0$ respectively.

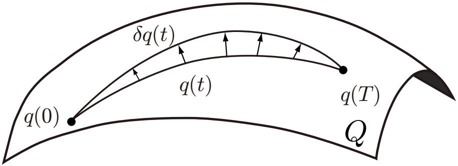

Figure: Variations $\delta q(t)$ of a curve $q(t)$.

Hamilton's variational principle.

Define the action functional $A : \mathcal{C}(Q) \to \mathbb{R}$ by

$$A(q) = \int_0^T L\left(q(t), \dot{q}(t)\right) dt,$$

where $\dot{q}(t)$ denotes the time derivative of $q(t)$.

If a curve $q = q(t) \in \mathcal{C}(Q)$ is a critical point of $A : \mathcal{C}(Q) \to \mathbb{R}$,

$$\delta A(q) = 0, \qquad \delta q(0) = \delta q(T) = 0,$$

the direct computation in local coordinates yields

$$\begin{aligned}\delta A(q) &= dA(q) \cdot \delta q = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} A(q_\epsilon(t)) \\ &= \int_0^T \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right) \delta q^i dt + \frac{\partial L}{\partial \dot{q}^i}\, \delta q^i \bigg|_0^T \\ &= \int_0^T \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right) \delta q^i dt = 0,\end{aligned}$$

for all δq^i.

In the above, the Einstein summation convention is employed; that is, a repeated index implies summation over that index. We shall use this convention throughout the paper unless stated otherwise.

Thus we get the Euler–Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0, \quad i = 1, \ldots, n.$$

These Euler–Lagrange equations with degenerate Lagrangian can have no solution at all. Here is a trivial example:

$$L = \dot{q} + q, \qquad q \in \mathbb{R}.$$

It is fair to say that the first author who considered the Hamiltonian aspects of Lagrangian dynamical systems with degenerate Lagrangian was Hamilton himself. The subject of his investigations was first geometrical optics and then the dynamics of mechanical systems in a tensional force field. partial force field.

The optical properties of a medium are determined by the refractive index n which is equal to the inverse of the speed of light. In general it is a function of the point x and the direction of the velocity of light particles $v = \dot{x}$

$$n = f(\left(x, \frac{\dot{x}}{|\dot{x}|}\right) .$$

The propagation time of light along the beam (optical path length) is determined by the integral

$$A = \int_{t_1}^{t_2} L(x, \dot{x}) dt , \qquad L(x, \dot{x}) = |\dot{x}|\, f .$$

This integral is often referred to as the Fermat action.

According to Fermat's principle, light propagates along the path with the shortest time duration.

In other words, the path $t \to x(t)$ is a stationary point of the functional A with fixed ends and thus satisfies the Lagrangian equation with Lagrangian L.

Since the Lagrangian L is a homogeneous function of the first degree on velocities

$$\det \left| \frac{\partial^2 L}{\partial v^2} \right| = 0$$

and we have singular Lagrangian.

This difficulty was overcome by Hamilton in the following way.

We introduce a closed set of points $v \in \mathbb{R}^3$ satisfying the equation $L(x, v) = 1$. This surface is called the indicatrix at the point x.

Then we introduce a figuratrix, the set of momenta $y \in \mathbb{R}^3$ defined by the following relations

$$y = \frac{\partial L}{\partial v}, \qquad L(x, v) = 1 .$$

If the indicatrix is a convex surface, then the figuratrix has the same property.

In this important case, there exists a single function $H(x, y)$ that is positively homogeneous on y

$$H(x, \lambda y) = \lambda H(x, y)$$

for all $\lambda > 0$ and equal to 1 for all momenta lying on the figuratrix. Transformation

$$v = \frac{\partial H}{\partial y}, \qquad H(x, y) = 1$$

translates the figuratrix into an indicatrix.

Thus, the functions L and H (as well as as well as indicatrix and figuratrix) are dual to each other.

As shown Hamilton in 1824, the path $t \to x(t)$ is a light ray if and only if there is a 'conjugate' function $t \to y(t)$ such that they together satisfy the canonical equations

$$\dot{x} = \frac{\partial H}{\partial y}, \qquad \dot{y} = -\frac{\partial H}{\partial x}.$$

Let us comeback to non-singular case.

Suppose the Lagrangian $L : TQ \to \mathbb{R}$ is hyperregular, namely, for every point $\dot{q} \in T_q Q$,

$$\det \left[\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right] \neq 0.$$

From the Euler–Lagrange equations, we get

$$\ddot{q}^j = \left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right)^{-1} \left(\frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^k} \dot{q}^k \right), \quad j = 1, \ldots, n.$$

In fact the above equations ensure that there exists a second-order vector field, called the Lagrangian vector field, denoted by

$$X_L : TQ \to \ddot{Q} \subset TTQ.$$

Here $\ddot{Q}$ is the second-order submanifold defined by

$$\ddot{Q} := \left\{ w \in TTQ \,\middle|\, T\tau_Q(w) = \tau_{TQ}(w) \right\},$$

and

$$T\tau_Q : TTQ \to TQ; \qquad (q, \dot{q}, \delta q, \delta \dot{q}) \mapsto (q, \delta q)$$

and

$$\tau_{TQ} : TTQ \to TQ; \qquad (q, \dot{q}, \delta q, \delta \dot{q}) \mapsto (q, \dot{q}).$$

Hence, $T\tau_Q(w) = \tau_{TQ}(w)$ yields

$$\delta q = \dot{q}$$

and therefore an element w in the second-order submanifold $\ddot{Q}$ has the components $(q, \dot{q}, \ddot{q})$.

Say X, Y are two vector bundles over Q and $f : X \to Y$ is a smooth fiber-preserving map (not necessarily fiberwise linear).

Then, we can define the object $\mathbb{F}f : X \to Hom(X, Y)$, by setting for each $x \in Q$ and $v \in X_x$,

$$(\mathbb{F}f)(v) := D(f|_{X_x})(v) \in Hom(X_x, Y_x) = Hom(X, Y)_x$$

One can show this is a smooth fiber-bundle morphism or Fiber Derivative.

This is really the appropriate terminology because we're restricting f to the fiber to get the mapping

$$f|_{X_x} : X_x \to Y_x$$

between vector spaces, and we are taking the usual derivative of such an object.

Now, to such a morphism $f : X \to Y$, we can define another morphism $E_f : X \to Y$ as

$$E_f(v) := (\mathbb{F}f(v))(v) - f(v)$$

The reason for the symbol E is that it's kind of like the "energy mapping associated to f".

Suppose that the fiber derivative $\mathbb{F}f : X \to Hom(X, Y)$ is a fiber-bundle isomorphism (for which it is necessary that Y have one-dimensional vector spaces as its fibers, so that X_x and $Hom(X, Y)_x$ have the same vector space dimension).

In this case, we can consider the mapping

$$\lambda_f = E_f \circ (\mathbb{F}f)^{-1} : Hom(X, Y) \to Y.$$

Classically, this mapping λ_f is called the Legendre-transform of f. So, given the function f, we can consider its energy E_f, and then change variables (compose with $\mathbb{F}f)^{-1}$.

As a special case, suppose $L : TQ \to \mathbb{R}$ is a smooth function (the Lagrangian, which we can trivially think of as a fiber-bundle map $TQ \to Q \times \mathbb{R}$, $v \to (x, f(v))$, hence everything above can be applied).

Then the fiber derivative is $\mathbb{F}L : TQ \to T^*Q$; in terms of bundle coordinates, it is

$$(x^1, \dots, x^n, v^1, \dots, v^n) \to (x^1, \dots, x^n; \frac{\partial L}{\partial v^1}(x, v), \dots, \frac{\partial L}{\partial v^n}(x, v))$$

Now, the energy function is

$$E = E_L : TQ \to \mathbb{R},$$

which by unwinding the definitions, can be written in coordinates as

$$(x, v) \to v_i \frac{\partial L}{\partial v_i}(x, v) - L(x, v)\,.$$

If we make the assumption that the fiber derivative

$$\mathbb{F}L : TQ \to T^*Q$$

is a diffeomorphism (typically called a hyperregular Lagrangian), then we can consider the function

$$H = E \circ (\mathbb{F}L)^{-1} : T^*Q \to \mathbb{R},$$

and this is what we call the Hamiltonian function associated to the Lagrangian L.

It is this function H (defined on a completely different space) that is usually referred to as "the Legendre transform of L" in coordinates, people often write

$$H = \dot{q}_i p_i - L.$$

So, there's two things to distinguish: the first is the fiber derivative, the second is the Legendre transformation (which is the composition of the "energy" by the inverse of the fiber-derivative). Often though, people may use "Legendre transform" to mean both these things.

Legendre transform

The Legendre transform $\mathbb{F}L : TQ \to T^*Q$ is a map defined by

$$\mathbb{F}L(v) \cdot w = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} L(v + \epsilon\, w).$$

It is apartial case of the fiber derivative of L at $v \in T_qQ$ along $w \in T_qQ$), which is given in local coordinates by

$$\mathbb{F}L(q^i, v^i) = \left(q^i, p_i = \frac{\partial L}{\partial v^i}\right).$$

When L is hyperregular, the Legendre transform $\mathbb{F}L : TQ \to T^*Q$ is globally diffeomorphic.

Marsden, J. E. and T. S. Ratiu. Introduction to Mechanics and Symmetry. 2nd ed. Texts Appl. Math. 17, New York: Springer-Verlag, 1999.

G. D. Birkhoff, Dynamical systems, Amer. Math. Soc. Colloq. Publ., 9, Amer. Math. Soc., New York, 1927, viii+295 pp.

L. C. Young, Lectures on the calculus of variations and optimal control theory, W. B. Saunders Co., Philadelphia–London–Toronto, ON, 1969, xi+331 pp.

G. A. Bliss, Lectures on the calculus of variations, Univ. of Chicago Press, Chicago, IL, 1946, ix+296 pp.

H. Hertz, Die Principien der Mechanik in neuem Zusammenhange dargestellt, Gesammelte Werke, III, J. A. Barth, Leipzig, 1894, xxix+312 pp.