Invariant derivation of the Euler-Lagrange equation

1988 J. Phys. A: Math. Gen. 21 L1013

(http://iopscience.iop.org/0305-4470/21/21/003)

J. Phys. A: Math. Gen. 21 (1988) L1013-L1017. Printed in the UK

LETTER TO THE EDITOR

Invariant derivation of the Euler–Lagrange equation

James M Nester
Department of Physics, National Central University, Chung-Li, Taiwan 32054

Received 3 August 1988

Abstract. The tangent bundle geometry is used to obtain a coordinate-free derivation of the Euler-Lagrange equation.

The Lagrangian formulation of classical mechanics is the most fundamental approach to dynamics. Nevertheless, the usual practice is to transform to the Hamiltonian form. An underlying mathematical reason is that phase space T^*Q (the cotangent bundle of configuration space Q) is *canonically* a symplectic manifold [1] whereas TQ (the tangent bundle) is not. It is this symplectic structure on phases space which gives rise to the elegant simplicity of the Hamiltonian formalism.

Although not as well known, there is also a rich geometric structure on TQ which has been studied, in particular, by Klein [2, 3] (see also Godbillion [4]). Using this structure, the Euler-Lagrange equations may be given an invariant geometric formulation directly in terms of the (pre)-symplectic geometry determined by a Lagrangian function [5] without any reference to the symplectic structure on T^*Q. This is sometimes necessary since not all Lagrangians L lead to a nice Legendre transformation (fibre derivative FL) from TQ to T^*Q. In that case there is no Hamiltonian form [5, 6].

The geometry obtained from a *degenerate* Lagrangian on TQ is *pre*-symplectic without any natural symplectic extension. The Dirac constraint algorithm [7, 8] (just as on the cotangent bundle [9-12]) can be invariantly formulated directly on TQ in terms of this presymplectic geometry [5]. A central role is played by the second-order vector field condition (i.e. $\dot{q} = v$) [13-15].

While there are geometric coordinate-free *studies* of these Lagrangian equations we know of no coordinate-free *derivation* of the Euler-Lagrange equation. Our aim is to present such an invariant derivation of the Euler-Lagrange equation from the usual starting point in physics: extremising the action determined by a Lagrangian function.

Some results, in particular the Euler-Lagrange equation, are most easily obtained by first using a coordinate system and then demonstrating suitable covariant transformation laws. Modern work, however, has shown the value of coordinate-free geometric *formulations*. Likewise there is also a value in geometrically invariant *derivations*. Not only do they directly show that the result is coordinate independent they also serve to clarify certain assumptions.

First we establish some notation and recall the basic definitions necessary for the geometric structures we will use on the tangent bundle. For more details see [2-4]. For any manifold M let τ_M denote the projection from the tangent bundle TM. A

0305-4470/88/211013+05$02.50

differentiable map $f: M \to Q$ induces the tangent map $Tf: TM \to TQ$. In particular, from $\tau_Q: TQ \to Q$ we obtain $T\tau_Q: T(TQ) \to TQ$ such that the diagram

$$\begin{array}{ccc} T(TQ) & \xrightarrow{T\tau_Q} & TQ \\ {\scriptstyle \tau_{TQ}}\downarrow & & \downarrow{\scriptstyle \tau_Q} \\ TQ & \xrightarrow[\tau_Q]{} & Q \end{array}$$

commutes. This diagram is the foundation for several structures. The vertical subbundle $V(TQ)$ of the second tangent bundle $T(TQ)$ is defined by $V(TQ) := \ker T\tau_Q$. The vertical lift $\xi_y: T_qQ \to V_y(TQ)$ is defined by

$$\xi_y(w) := \frac{\mathrm{d}}{\mathrm{d}\lambda}(y + \lambda w)$$

where $q = \tau_Q(y) = \tau_Q(w)$. From ξ we can construct the almost tangent structure

$$J_y := \xi_y \circ T\tau_Q : T_y(TQ) \to T_y(TQ)$$

which has the properties $\operatorname{Im} J = \ker J = V(TQ)$, hence $J^2 = 0$, and the Liouville canonical vector field V on TQ, which is defined by

$$V_y := \xi_y(y) \qquad \text{for } y \in TQ.$$

A curve $C:[a, b] \to Q$ prolongs to sections C' of TQ and C'' of $T(TQ)$. The vector field $X := C''$ on $T(TQ)$ is special in that it is second order. A second-order vector field is characterised by the property $T\tau_Q X = \tau_{TQ} X$. A direct application of the above definitions leads to the alternate characterisation

$$JX = V \tag{1}$$

which is more convenient for our purposes.

We recall that a linear endomorphism $A: TM \to TM$ induces a derivation (with grading rank 0) the interior product $i_A: \wedge^p M \to \wedge^p M$ on differential forms on M defined by

$$(i_A\beta)(X_1, \ldots, X_p) := \sum_{k=1}^{p} \beta(X_1, \ldots, AX_k, \ldots, X_p)$$

where $X_j \in TM$, with $i_A f := 0$ for any function. Further derivations may be obtained from the graded commutators with the basic derivations: the exterior derivative d and the interior product with a vector field i_X of grading rank $+1$ and -1 respectively.

On TQ this construction naturally leads to the vertical derivative

$$d_J := [i_J, d] = i_J d - d i_J.$$

It is easy to verify that

$$[d, d_J] = dd_J + d_J d = 0$$

$$[d_J, i_V] = d_J i_V + i_V d_J = i_J$$

$$[i_X, i_J] = i_X i_J - i_J i_X = i_{JX}$$

and that $d_J^2 = 0$. The remaining basic graded commutator $[i_Z, d] = i_Z d + d i_Z$ is just the Lie derivative $\mathscr{L}_Z$.

For completeness we include the purely tangent space definition of the fibre derivative $Fg: TQ \to T^*Q$ of $g: TQ \to \mathbb{R}$:

$$Fg(y) := dg(y) \circ \xi_y : T_qQ \to \mathbb{R}.$$

Remark. The fibre derivative may be used to relate the structures on TQ and T^*Q, in particular $d_Jg = Fg^*\theta$ where θ is the canonical 1-form on T^*Q. In general $dd_Jg = Fg^*\omega$ is only presymplectic whereas $\omega = d\theta$ is the natural symplectic structure on the cotangent bundle.

To each path $C:[a, b] \to Q$ a Lagrangian function $L: TQ \to \mathbb{R}$ associates an action

$$S[C] := \int_a^b L \circ C' \, \mathrm{d}t. \tag{2}$$

We wish to show the following proposition.

Proposition. For paths with fixed endpoints the action (2) is stationary for the path C iff the Euler-Lagrange equation

$$\mathscr{E} := i_X dd_J L + dE_L = 0 \tag{3}$$

is satisfied, where $X := C''$ is the Lagrangian vector field on TQ and $E_L := i_V dL - L$ is the energy.

Proof. Consider a one-parameter set of paths $C_\lambda(t)$ in Q with fixed endpoints. They determine two vector fields on TQ, the Lagrangian vector field $X := C''$ which satisfies the second-order equation condition (1) and the deviation vector field $Z := (\partial/\partial\lambda)C'_\lambda$ which is characterised by JZ vanishing at the endpoints and the properties

$$[X, Z] = 0 = [Z, V - JX]. \tag{4}$$

We wish to find the necessary and sufficient conditions for the action S to be stationary:

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\lambda}\right|_0 = \int_a^b \langle dL|Z\rangle \, \mathrm{d}t = 0.$$

(i) For all Z

$$\begin{aligned}\langle dL|Z\rangle &= Z\langle dL|V\rangle - \langle dE_L|Z\rangle = Z\langle dL|V\rangle - \langle \mathscr{E}|Z\rangle + \langle dd_JL|X, Z\rangle \\ &= Z\langle dL|V\rangle - \langle \mathscr{E}|Z\rangle + X\langle d_JL|Z\rangle - Z\langle d_JL|X\rangle - \langle d_JL|[X, Z]\rangle \\ &= -\langle \mathscr{E}|Z\rangle + X\langle dL|JZ\rangle + Z\langle dL|V - JX\rangle - \langle dL|J[X, Z]\rangle \\ &= -\langle \mathscr{E}|Z\rangle + X\langle dL|JZ\rangle + \langle \mathscr{L}_Z dL|V - JX\rangle + \langle dL|[Z, V - JX]\rangle - \langle dL|J[X, Z]\rangle.\end{aligned} \tag{5}$$

Hence the vanishing of $\mathscr{E}$ is sufficient for $\mathrm{d}S/\mathrm{d}\lambda = 0$, since under the conditions (1) and (4) imposed on Z and X all of the terms vanish except

$$X\langle dL|JZ\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle dL|JZ\rangle$$

which integrates to an evaluation at the endpoints where JZ vanishes.

(ii) The above calculation does *not* show that $\mathscr{E} = 0$ is necessary since Z is not arbitrary. In particular we have the restrictions (4). To obtain an arbitrary vector field we add JW for any W to Z, then

$$\langle \mathscr{E}|Z\rangle = \langle \mathscr{E}|Z + JW\rangle - \langle \mathscr{E}|JW\rangle = \langle \mathscr{E}|Z + JW\rangle - \langle i_J\mathscr{E}|W\rangle. \tag{6}$$

Lemma.

$$i_J\mathscr{E} = i_{V-JX}dd_JL.$$

Proof.

$$i_Ji_Xdd_JL = [i_J, i_X]dd_JL + i_Xi_Jdd_JL$$
$$= -i_{JX}dd_JL + i_Xd_J^2L + i_Xdi_Jd_JL = -i_{JX}dd_JL$$

and

$$i_JdE_L = d_JE_L = d_Ji_VdL - d_JL$$
$$= [d_J, i_V]dL - d_JL - i_Vd_JdL$$
$$= -i_Vd_JdL = i_Vdd_JL.$$

Consequently (6) becomes

$$\langle \mathscr{E}|Z\rangle = \langle \mathscr{E}|Z+JW\rangle - \langle i_{V-JX}dd_JL|W\rangle$$
$$= \langle \mathscr{E}|Z+JW\rangle + \langle i_Wdd_JL|V-JX\rangle.$$

Using this result (5) may be written in the form

$$\langle dL|Z\rangle = -\langle \mathscr{E}|Z+JW\rangle + \langle i_Wdd_JL|V-JX\rangle + \langle \mathscr{L}_ZdL|V-JX\rangle - \langle dL|J[X, Z]\rangle$$
$$+ \langle dL|[Z, V-JX]\rangle + X\langle dL|JZ\rangle$$

for all W, Z. With the restrictions that $V-JX=0$, $[X, Z]=0$, $[Z, V-JX]=0$ and JZ vanishes at the endpoints, we have

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\lambda}\right|_0 = \int_a^b \langle dL|Z\rangle\,\mathrm{d}t = -\int_a^b \langle \mathscr{E}|Z+JW\rangle\,\mathrm{d}t.$$

Although Z is not completely arbitrary, $Z+JW$ is. Consequently, $\mathrm{d}S/\mathrm{d}\lambda|_0 = 0$ with $V-JX=0$ implies $\mathscr{E}=0$ and conversely $\mathscr{E}=0=V-JX$ implies $\mathrm{d}S/\mathrm{d}\lambda|_0=0$.

In general the second-order condition plays an essential and independent role. Although the Euler–Lagrange equation $\mathscr{E}=0$, via the lemma, does impose a restriction:

$$i_J\mathscr{E} = i_{V-JX}dd_JL = 0$$

it is not always strong enough to assure that $V-JX$ vanishes, since dd_JL is only presymplectic if L is degenerate. Consequently, to guarantee that the action S be stationary we must supplement $\mathscr{E}=0$ in general with $V-JX=0$. Together these conditions are necessary and sufficient.

The form of the Euler–Lagrange equation used here corresponds to Hamilton's equations on the cotangent bundle. This form of the equation can be transcribed into

$$\mathscr{E} := i_Xdd_JL + dE_L = \mathscr{L}_Xd_JL - di_Xd_JL + d(i_VdL - L) = \mathscr{L}_Xd_JL - dL - di_{V-JX}dL$$

which, along with the second-order condition $V-JX$, is equivalent to

$$\mathscr{E}' := \mathscr{L}_Xd_JL - dL.$$

This latter form is less convenient geometrically (it depends differentially on the Lagrangian vector field) but is more recognisable to physicists.

This idea was first considered while at the University of Saskatchewan. Discussions with Mark Gotay were most helpful. This present work has been supported by the National Science Council of the Republic of China under contract NSC77-0298-M008-20.

References

[1] Abraham R and Marsden J 1978 *Foundations of Mechanics* (New York: Benjamin) 2nd edn
[2] Klein J 1962 *Ann. Inst. Fourier* **12** 1
[3] Klein J 1974 *Symp. Math.* **14** 181
[4] Godbillion C 1969 *Géométrie Différentiel et Mécanique Analytique* (Paris: Hermann)
[5] Gotay M and Nester J M 1979 *Ann. Inst. H Poincaré* A **30** 129
[6] Batlle C, Gomis J, Pons J M and Roman-Roy N 1987 *J. Phys. A: Math. Gen.* **20** 5113
[7] Dirac P A M 1964 *Lectures on Quantum Mechanics* (New York: Academic)
[8] Hanson A, Regge T and Teitelboim C 1976 *Accad. Naz. dei Lincei* **22** 1976
[9] Gotay M, Nester J M and Hinds G 1978 *J. Math. Phys.* **19** 2388
[10] Gotay M and Nester J M 1979 *Proc. 7th Int. Colloq. on Group Theoretical Methods in Physics* (*Lecture Notes in Physics* **94**) (Berlin: Springer) p 272
[11] Sundermeyer K 1982 *Constrained Dynamics* (*Lecture Notes in Physics* **169**) (Berlin: Springer)
[12] Batlle C, Gomis J, Pons J M and Roman-Roy N 1986 *J. Math. Phys.* **27** 2953
[13] Künzle H P 1969 *Ann. Inst. H Poincaré* A **11** 393
[14] Gotay M and Nester J M 1980 *Ann. Inst. H Poincaré* A **32** 1
[15] Skinner R and Rusk R 1983 *J. Math. Phys.* **24** 2589