Hamilton-Jacobi equations for nonholonomic dynamics

Michele Pavon

Dipartimento di Matematica Pura e Applicata,
Università di Padova, via Belzoni 7, 35131 Padova, Italy

(Dated: November 17, 2005)

We derive generalized Hamilton-Jacobi equations for dynamical systems subject to linear velocity constraints. As long as a solution of the generalized Hamilton-Jacobi equation exists, the action is actually minimized (not just extremized).
I. INTRODUCTION

Consider a mechanical system with configuration space $\mathbb{R}^n$. Let $L$ be the Lagrangian, and suppose that the system is subject to $k < n$ nonholonomic constraints of the form

$$\omega_i(x(t))^T \dot{x}(t) = 0, \quad i = 1, 2, \ldots, k, \quad t \in [t_0, t_1],$$

(I.1)

where the $\omega_i : \mathbb{R}^n \to \mathbb{R}^n$ are smooth functions and $T$ denotes transpose. Let $\Omega(x)$ be the $k \times n$ matrix whose $i^{th}$ row is $\omega_i(x)^T$. Then, an application of d’Alembert’s principle, together with the method of Lagrange multipliers, gives that the equations of motion for the system are

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \Omega^T \lambda,$$

(I.2)

where $\lambda$ is the $k$ dimensional Lagrange multiplier. Equation (I.2) together with (I.1) constitute a system of $n + k$ equations for the $n + k$ unknowns $x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_k$. The components of $\Omega^T \lambda$ can be physically interpreted as the components of the (polygenic) force which acts on the mechanical system in order to maintain the given non-holonomic conditions [9]. Notice that d’Alembert’s principle is not variational. A variational approach to dynamics of systems subject to linear velocity constraints was proposed in [14] (see also [1, Chap. 1, Sect.4]). A lucid critique of this “Vakonomic dynamics” (variational axiomatic kind dynamics) can be found in [19]. It is shown there that the vakonomic equations may lead to paradoxical behaviour. The relation between the vakonomic and holonomic approaches has also been discussed in [1, 4–6, 10, 16].

We show in this paper that the second, hydrodynamic form of Hamilton’s principle may be extended to nonholonomic systems. We concentrate on the optimization aspect which is largely ignored in the physics literature. As long as a solution of the generalized Hamilton-Jacobi equation exists, the action is minimized by a path satisfying the correct equations of motions (I.2). Our derivation relies on general nonlinear Lagrange functionals [11–13]. It would be quite feasible to derive the result, after a suitable transformation, using standard optimal control results such as [15, Section 5.2, Theorem 7] and [7, Chapter IV]. We find, however, that the approach based on Lagrange functionals is more transparent.

The paper is outlined as follows. In the next section, we recall the hydrodynamic form of the classical Hamilton principle as established in [18, Section II] (the latter developed from [8]). In Section III, we extend the latter result to systems subject to linear velocity
constraints. The paper concludes with a discussion section.

II. THE CLASSICAL HAMILTON PRINCIPLE

Consider a dynamical system with configuration space $\mathbb{R}^n$. Let

$$L(x, v) := \frac{1}{2}mv \cdot v - V(x)$$  \hspace{1cm} (II.3)

be the Lagrangian function, where $V(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is of class $C^1$. Extension of the results of this paper to general Lagrangian functions that are strictly convex with respect to $v$ appears straightforward. We prefer, however, to treat the simple case (II.3) in order to avoid obscuring ideas with technicalities. Let $X_0$ denote the class of all $C^1$ paths $x : [t_0, t_1] \to \mathbb{R}^n$ such that $x(t_0) = x_0$. Let $V$ denote the family of continuous functions $v : [t_0, t_1] \to \mathbb{R}^n$. For $(x, v) \in X_0 \times V$, we define the functional $J(x, v)$ by

$$J(x, v) = \int_{t_0}^{t_1} L(x(t), v(t)) \, dt - S_1(x(t_1)),$$  \hspace{1cm} (II.4)

where $S_1 : \mathbb{R}^n \to \mathbb{R}$ is continuous. Consider the following control problem:

$$\text{Minimize } \{ J(x, v) | (x, v) \in (X_0 \times V) \},$$  \hspace{1cm} (II.5)

subject to the constraint

$$\dot{x}(t) = v(t), \quad \forall t \in [t_0, t_1].$$  \hspace{1cm} (II.6)

**Remark II.1** It is apparent that this control problem is equivalent to minimizing the action functional $I(x) := J(x, \dot{x})$ over $X_0$.

To solve problem (II.5)-(II.6) we rely on the following elementary, albeit fundamental, result in the spirit of Lagrange. Let $Y$ be a nonempty set. Consider the minimization of $J : Y \to \bar{\mathbb{R}}$, where $\bar{\mathbb{R}}$ denotes the extended reals, over the nonempty subset $M$ of $Y$.

**Lemma II.2** (Lagrange Lemma) Let $\Lambda : Y \to \bar{\mathbb{R}}$ and let $y_0 \in M$ minimize $J + \Lambda$ over $Y$. Assume that $\Lambda(\cdot)$ is finite and constant over $M$. Then $y_0$ minimizes $J$ over $M$.

**Proof.** For any $y \in M$, we have $J(y_0) + \Lambda(y_0) \leq J(y) + \Lambda(y) = J(y) + \Lambda(y_0)$. Hence $J(y_0) \leq J(y)$. Q.E.D.
A functional $\Lambda$ which is constant and finite on $M$ is called Lagrange functional. For problem (II.5), let

$$M = \{ (x, v) \in X_0 \times V \mid \dot{x}(t) = v(t), \forall t \in [t_0, t_1] \}.$$  

We introduce a suitable class of nonlinear Lagrange functionals for our problem. Let $F : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}$ be of class $C^1$. Corresponding to such an $F$, we define the nonlinear functional $\Lambda^F$ on $X_0 \times V$

$$\Lambda^F(x, v) := F(t_1, x(t_1)) - F(t_0, x(t_0)) + \int_{t_0}^{t_1} \left[ -\frac{\partial F}{\partial t}(t, x(t)) - v(t) \cdot \nabla F(t, x(t)) \right] dt.$$  

When $(x, u) \in M$, by the chain rule, we have $\Lambda^F(x, u) = 0$. Thus, $\Lambda^F$ is indeed a Lagrange functional for our problem. The solution procedure is now outlined as follows. Consider the unconstrained minimization

$$\min_{(x,v) \in (X_0 \times V)} (J + \Lambda^F)(x,v). \quad (\text{II.7})$$

We perform two-stage optimization.  

**Step 1.** For each fixed $x \in X_0$, we try to compute an optimal control $v^*_x$ through pointwise minimization of the integrand of $J + \Lambda^F$. More explicitly, consider for each $x \in X_0$ and each $t \in [t_0, t_1]$ the finite-dimensional problem

$$\min_{v \in \mathbb{R}^n} \left\{ \frac{1}{2} m v \cdot v - V(x(t)) - \frac{\partial F}{\partial t}(t, x(t)) - v \cdot \nabla F(t, x(t)) \right\}. \quad (\text{II.8})$$

We get

$$v^*_x(t) = \frac{1}{m} \nabla F(t, x(t)). \quad (\text{II.9})$$

We notice that $v^*_x$ belongs to the class of admissible velocities $V$.  

**Step 2.** Consider now the minimization of the functional

$$\Gamma^F(x) = (J + \Lambda^F)(x, v^*_x)$$

on the space $X_0$. We have

$$\Gamma^F(x) = \quad -S_1(x(t_1)) + F(t_1, x(t_1)) - F(t_0, x(t_0)) + \int_{t_0}^{t_1} \left[ -\frac{\partial F}{\partial t}(t, x(t)) - \frac{1}{2m} \nabla F(t, x(t)) \cdot \nabla F(t, x(t)) - V(x(t)) \right] dt.$$  

If we can find $S$ such that $\Gamma^S(\cdot)$ is actually constant on $X_0$, then any pair $(x, v^*_x) \in X_0 \times V$ solves problem (II.7). Then, by Lemma II.2, if the pair $(x, v^*_x)$ satisfies

$$\dot{x}(t) = v^*_x(t) = \frac{1}{m} \nabla S(t, x(t)), \quad \forall t \in [t_0, t_1],$$

it also solves the original constrained problem (II.5)-(II.6).
Theorem II.3 ([18]) Let $S(t, x)$ be any $C^1$ solution on $[t_0, t_1] \times \mathbb{R}^n$ of the terminal value problem

\[
\frac{\partial S}{\partial t} + \frac{1}{2m} \nabla S \cdot \nabla S + V(x) = 0, \quad (II.10)
\]
\[
S(t_1, x) = S_1(x). \quad (II.11)
\]

Let $x^* \in \mathcal{X}_0$ be any solution on $[t_0, t_1]$ of

\[
\dot{x}(t) = \frac{1}{m} \nabla S(x(t), t). \quad (II.12)
\]

Then $(x^*, \frac{1}{m} \nabla S(x^*(t), t))$ solves problem (II.5)-(II.6).

Proof. If $S$ solves (II.10)-(II.11), we get $\Gamma_S(x) \equiv -S(x_0, t_0)$ on $\mathcal{X}_0$. Q.E.D.

Notice that when a $C^1$ solution $S(x, t)$ of (II.10)-(II.11) exists, then there are also solutions $x$ of the differential equation (II.12) satisfying $x(t_0) = x_0$, and therefore optimal pairs. In this case, the action functional is actually minimized, not just extremized. The difficulty lies, of course, with the terminal value problem (II.10)-(II.11) that, in general, only has a local in $t$ solution (namely, on some interval $[\bar{t}, t_1]$, $t_0 < \bar{t}$).

Remark II.4 Let us now assume that $S$ is of class $C^2$. Following [8], let us introduce the acceleration field $a(t, x)$ through a substantial time derivative

\[
a(t, x) := \left[\frac{\partial}{\partial t} + \frac{1}{m} \nabla S \cdot \nabla\right] \left(\frac{1}{m} \nabla S\right)(t, x) = \frac{1}{m} \nabla \left[\frac{\partial S}{\partial t} + \frac{1}{2m} \nabla S \cdot \nabla S\right](t, x).
\]

Then, (II.10) implies the local form of Newton’s law

\[
a(t, x) = -\frac{1}{m} \nabla V(x). \quad (II.13)
\]

III. NONHOLONOMIC DYNAMICAL SYSTEMS

Consider a system subject to linear velocity constraints of the form

\[
\Omega(x(t))\dot{x}(t) = 0, \quad t \in [t_0, t_1], \quad (III.14)
\]

where $\Omega : \mathbb{R}^n \to \mathbb{R}^{k \times n}, k < n$ is a continuous map. We assume that for each $x \in \mathbb{R}^n$, the rows of $\Omega$ are linearly independent. These constraints are called Pfaffian. A simple example is provided by a disk rolling on a plane without slipping. More complex nonholonomic systems
with Pfaffian constraints occur in many problems of robot motion planning and vehicular
dynamics, and have therefore been the subject of intensive study, see [2, 3, 17] and references
therein. Let
\[ \Omega(x(t))v(t) = 0, \quad t \in [t_0, t_1]. \]  
(III.15)
We now study the control problem (II.5)-(II.6)-(III.15), namely the same problem as in
the previous section when also constraint (III.15) is present. This problem is equivalent to
minimizing the action functional
\[ I(x) := J(x, \dot{x}) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t)) \, dt - S_1(x(t_1)), \]  
under the constraints (III.14). Reformulating the calculus of variations problem as a control
problem as before, we let
\[ M = \{ (x, v) \in X_0 \times V \mid \dot{x}(t) = v(t), \Omega(x(t))v(t) = 0, \forall t \in [t_0, t_1] \}. \]
Let \( F : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R} \) be of class \( C^1 \) and \( g : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^k \) be continuous. Corresponding
to such a pair, we define the nonlinear functional \( \Lambda^{F,g} \) on \( X_0 \times V \) by
\[ \Lambda^{F,g}(x, v) := F(t_1, x(t_1)) - F(t_0, x(t_0)) + \int_{t_0}^{t_1} \left[ -\frac{\partial F}{\partial t}(t, x(t)) - v(t) \cdot \nabla F(t, x(t)) + g(t, x(t))^T \Omega(x(t))v(t) \right] \, dt. \]
It is apparent that \( \Lambda^{F,g} \) is a Lagrange functional for the problem since it is identically zero
when (II.6) and (III.15) are satisfied. Following the same procedure as in the previous
section, we consider the un-constrained minimization of \( (J + \Lambda^{F,g})(x, v) \) over \( X_0 \times V \). For
\( x \in X_0 \) fixed, the pointwise minimization of the integrand of \( J + \Lambda^{F,g} \) at time \( t \) gives
\[ v_x^*(t) = \frac{1}{m} \left[ \nabla F(t, x(t)) - \Omega^T(x(t))g(t, x(t)) \right]. \]  
(III.17)
Notice that \( v_x^* \in V \). We consider next the minimization of the functional
\[ \Gamma^{F,g}(x) = (J + \Lambda^{F,g})(x, v_x^*) \]
on the space \( X_0 \). We have
\[ \Gamma^{F,g}(x) = -S_1(x(t_1)) + F(x(t_1), t_1) - F(x(t_0), t_0) \]
\[ + \int_{t_0}^{t_1} \left[ -\frac{\partial F}{\partial t}(t, x(t)) - \frac{1}{2m} \| \nabla F(t, x(t)) - \Omega^T(x(t))g(t, x(t)) \|^2 - V(x(t)) \right] \, dt, \]
where \( \| \cdot \| \) denotes the Euclidean norm in \( \mathbb{R}^n \). Let \( S(t, x) \) of class \( C^1 \) and \( \mu(t, x) \) continuous solve on \( \mathbb{R}^n \times [t_0, t_1] \) of the initial value problem

\[
\frac{\partial S}{\partial t} + \frac{1}{2m} \| \nabla S - \Omega^T \mu \|^2 + V(x) = 0, \\
S(x, t_0) = S_0(x).
\]

(III.18)

(III.19)

Then \( \Gamma^{S, \mu}(x) \equiv -S(x_0, t_0) \) on \( X_0 \). By Lemma (II.2), if \( x \in X_0 \) satisfies for all \( t \in [t_0, t_1] \)

\[
\dot{x}(t) = \frac{1}{m} \left[ \nabla S(t, x(t)) - \Omega^T(x(t))\mu(t, x(t)) \right], \\
\Omega(x(t)) \dot{x}(t) = 0,
\]

(III.20)

(III.21)

then it solves the problem together with the corresponding feedback velocity (III.17).

**Remark III.1** As in the unconstrained case, we now show that (III.18) implies the second principle of dynamics. Assume that \( S \) is of class \( C^2 \) and that \( \Omega \), and \( \mu \) are of class \( C^1 \). The acceleration field is again obtained through a substantial derivative of the velocity field

\[
a(t, x) := \left[ \frac{\partial}{\partial t} + \frac{1}{m} \left[ \nabla S - \Omega^T \mu \right] \cdot \nabla \right] \left( \frac{1}{m} \left[ \nabla S - \Omega^T \mu \right] \right)(t, x)
\]

\[
= \frac{1}{m} \left\{ \nabla \left[ \frac{\partial S}{\partial t} \right] + \frac{1}{2m} \| \nabla S - \Omega^T \mu \|^2 \right\} - \frac{\partial (\Omega^T \mu)}{\partial t} \right\}(t, x).
\]

Then, (III.18) yields

\[
a(t, x) = -\frac{1}{m} \nabla V(x) - \frac{1}{m} \Omega^T \frac{\partial \mu}{\partial t}.
\]

(III.22)

Define

\[
\lambda(t, x) := -\frac{\partial \mu}{\partial t}(t, x).
\]

For \( x \in X_0 \) satisfying (III.20) and of class \( C^2 \), let \( \lambda(t) := \lambda(t, x(t)) \). We then get

\[
m \ddot{x}(t) = -\nabla V(x(t)) + \Omega^T(x(t)) \lambda(t),
\]

(III.23)

namely equation (I.2). By differentiating (III.21), a simple calculation employing (III.23), shows that the Lagrange multipliers \( \lambda \) may be expressed as an instantaneous function of \( x \) and \( \dot{x} \), see e.g. [19, Section 2], [17, pp.269-270].

Next we show how the multiplier \( \mu \) can be also eliminated in our hydrodynamic context. If \( x \) satisfies (III.20)-(III.21), then, plugging (III.20) into (III.21), we get

\[
\Omega(x(t)) \frac{1}{m} \left[ \nabla S(t, x(t)) - \Omega^T(x(t))\mu(t, x(t)) \right] = 0
\]
Since $\Omega$ has full row rank, the latter is equivalent to
\[
\mu(t, x(t)) = (\Omega(x(t))\Omega^T(x(t)))^{-1}\Omega(x(t))\nabla S(t, x(t)) \tag{III.24}
\]
Plugging this into (III.20), we get
\[
\dot{x}(t) = \frac{1}{m} [(I - \pi(x(t))) \nabla S(t, x(t))], \tag{III.25}
\]
where $\pi(t, x)$ is defined by
\[
\pi(x) = \Omega^T(x)\Omega(x)\Omega^T(x)^{-1}\Omega(x). \tag{III.26}
\]
Observe that $\pi(x)^2 = \pi(x)$ and $\pi(x)^T = \pi(x)$. Thus, $\pi(x)$ is an orthogonal projection. In fact, $\pi(x)$ is the orthogonal projection onto $\text{Range}(\Omega^T(x))$.

**Remark III.2** Notice that (III.25) implies (III.21). Indeed,
\[
\Omega(x(t))\dot{x}(t) = \Omega(x(t))\frac{1}{m} [(I - \pi(x(t))) \nabla S(t, x(t))] = 0,
\]
since $I - \pi(x)$ projects onto the kernel of $\Omega(x)$.

Now we use the freedom we have in picking $S$ and $\mu$. Since (III.24) must be satisfied by an optimal solution, we impose that the pair $(S, \mu)$ satisfies *identically* on all of $[t_0, t_1] \times \mathbb{R}^n$
\[
\mu(t, x) = (\Omega(x)\Omega^T(x))^{-1}\Omega(x)\nabla S(t, x(t)). \tag{III.27}
\]
Hence, we can write $\Gamma^S(x)$ instead of $\Gamma^{S,\mu}(x)$, and equation (III.18) becomes
\[
\frac{\partial S}{\partial t} + \frac{1}{2m}||(I - \pi)\nabla S||^2 + V(x) = 0 \tag{III.28}
\]
We are now ready for our main result.

**Theorem III.3** For $x \in \mathbb{R}^n$, let $\sigma(x) = I - \pi(x)$ denote the orthogonal projection onto $\ker(\Omega(x))$. Let $S(t, x)$ be any $C^1$ solution on $[t_0, t_1] \times \mathbb{R}^n$ of
\[
\frac{\partial S}{\partial t} + \frac{1}{2m} \nabla S \cdot \sigma \nabla S + V(x) = 0, \quad S(t_1, x) = S_1(x). \tag{III.29}
\]
Then any $x \in X_0$ satisfying
\[
\dot{x}(t) = \frac{1}{m} \sigma(x(t)) \nabla S(t, x(t)) \tag{III.30}
\]
on $[t_0, t_1]$ solves together with

$$v(t) = \frac{1}{m} \sigma(x(t)) \nabla S(t, x(t))$$

problem (II.5)-(II.6)-(III.15) (equivalently, such an $x \in X_0$ minimizes (III.16) subject to (III.14)). If $S$ is of class $C^2$, and $x \in X_0$ satisfying (III.30) is also of class $C^2$, then $x$ satisfies equation (III.23)

$$m \ddot{x}(t) = -\nabla V(x(t)) + \Omega^T(x(t)) \lambda(t),$$

with $\lambda$ given by

$$\lambda(t) = -(\Omega(x(t))\Omega^T(x(t)))^{-1}\Omega(x(t))\nabla \frac{\partial S}{\partial t}(t, x(t)). \quad \text{(III.31)}$$

**Proof.** If $S$ solves (III.29), we get $\Gamma^S(x) = -S(x_0, t_0)$ for any $x \in X_0$. Thus any pair $(x, v) \in X_0 \times V$ solves the unconstrained minimization problem. Since $x$ satisfies (III.30), constraint (II.6) is fulfilled. Moreover, by Remark III.2, (III.15) is also satisfied. By Lemma II.2, the pair is optimal for the original constrained minimization. Finally, $x$ satisfies (III.23) with $\lambda$ as in (III.31) in view of Remark III.1 and (III.27). Q.E.D.

**IV. DISCUSSION**

We conclude the paper with a few simple remarks on how the minimization of the modified action leading to the vakonomic equations relates to the problem studied in the previous section. Assume that $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is of class $C^2$. Assume moreover that $S_1$ and $\Omega$ are of class $C^1$. Recall

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) \, dt - S_1(x(t_1))$$

and the constraint (III.14)

$$\Omega(x(t))\dot{x}(t) = 0, \quad t \in [t_0, t_1]. \quad \text{(IV.32)}$$

For $\chi$ taking values in $\mathbb{R}^k$ of class $C^1$, let

$$L(t, x, \dot{x}) := L(t, x, \dot{x}) + \chi^T(t)\Omega(x)\dot{x}.$$ 

Introduce also the modified action

$$\mathcal{I}(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) \, dt - S_1(x(t_1)) \quad \text{(IV.33)}$$
Taking as variations paths $y(t) \in C^1[t_0, t_1]$ such that $y(t_0) = 0$, and setting the first variation of $I$ equal to zero, we get the equations

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0, \quad (IV.34)
\]

\[
\frac{\partial L}{\partial \dot{x}}(t_1, x(t_1), \dot{x}(t_1)) = \frac{\partial S_1}{\partial x}(x(t_1)). \quad (IV.35)
\]

They read

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = -\Omega^T \chi - \chi \cdot \Omega \dot{x} + \frac{\partial}{\partial x} (\chi \cdot \Omega \dot{x}), \quad (IV.36)
\]

\[
\frac{\partial L}{\partial \dot{x}}(t_1, x(t_1), \dot{x}(t_1)) + \Omega^T(x(t_1))\chi(t_1) = \frac{\partial S_1}{\partial x}(x(t_1)), \quad (IV.37)
\]

where

\[
(\chi \cdot \Omega \dot{x})_i = \sum_{\alpha=1}^{k} \sum_{j=1}^{n} \partial \Omega_{\alpha i} \dot{x}_j, \quad \frac{\partial}{\partial x} (\chi \cdot \Omega \dot{x})_i = \sum_{\alpha=1}^{k} \sum_{j=1}^{n} \partial \Omega_{\alpha i} \dot{x}_j.
\]

Equations (IV.36) together with (IV.32) constitute the equations of the Vakonomic dynamics. Moreover, (IV.36)-(IV.37) are necessary conditions for the unconstrained minimization of the modified action (IV.33) over $X_0$. If $x$ does minimize (IV.33) over $X_0$ and it satisfies (IV.32), then, by Lemma II.2, it also minimizes $I(x)$ over $X_0$ subject to (IV.32).

Acknowledgments

I wish to thank Luis Bonilla and Franco Cardin for some useful conversations on the content of this paper and for pointing out some relevant bibliography to me.


