Unspecified Journal
Volume 00, Number 0, Pages 000–000
S ????-????(XX)0000-0

HAMILTON-JACOBI ON THE SYMPLECTIC GROUP

IVAR EKELAND

Dedicated to Jean Mawhin and his students

Abstract. The classical Hamilton-Jacobi-Bellman theory in the calculus of variations, which is associated with the Bolza problem, is extended to other kinds of boundary-value problems, such as periodicity. By using the dual action principle of Clarke and earlier results by the author, one can establish the analogue of HJB on the symplectic group and show that it has a solution

1. Introduction

It is great pleasure to do mathematics, and I have enjoyed it for more than fifty years now. Part of that pleasure comes from meeting other mathematicians, and there are very few, if any, that I have enjoyed more than Jean Mawhin. His sense of humour of course is part of the enjoyment, but so are his mathematics, always deeply rooted in classical problems, and bringing to them modern methods and deceptively simple solutions. He belongs to a long and distinguished Belgian tradition of the calculus of variations, starting with de la Vallée-Poussin and de Donder [2], continued by Jean himself and his students, such as Michel Willem. I wish to point out that their teaching is no less remarkable than their research. The treatise of de la Vallée-Poussin, in its second edition [6], was the first one to introduce the Lebesgue integral, and Jean's treatise [4] , is no less revolutionary and a pleasure to read. I cannot resist the opportunity of commending Michel's expository talent as well, in [8], [7], [9]: short books, which contain an amazing amount of well-digested material.

All these people have been a great inspiration to me, and I dedicate this work to them. I will try to fit into the same tradition by describing a research program which starts in the classical calculus of variations and ends in the symplectic group. I have long asked myself whether the Hamilton-Jacobi equation, nowadays known as Hamilton-Jacobi-Bellman because of the latter's important contribution, and shortened to HJB, can be extended to other settings, for instance to periodic boundary conditions. It turns out that they can. I will describe how to proceed, and leave the detailed study of the equation to further studies, preferably by younger people.

Received by the editors March 10, 2016.

2. The classical situation

Let me first summarize the classical theory (see [1] or [7]). Consider the classical Bolza problem in the one-dimensional calculus of variations:

$$\inf \int_0^T f\left(q, \frac{dq}{dt}\right) dt \tag{2.1}$$

$$q(0) = q_0, \quad q(T) = q_1, \quad \frac{dq}{dt} \in L^1 \tag{2.2}$$

where $T > 0$, q_0 and $q_1 \in R^n$ are prescribed. Suppose $f(q, \chi)$ is continuously differentiable, convex wrt χ and coercive, meaning that we have $f(q, \chi) \geq \Phi(|\chi|)$ where Φ is bounded from below and $\Phi(t)\, t^{-1} \to \infty$ when $t \to \infty$. Then it can be shown that the minimizer q exists, and that $\left|\frac{dq}{dt}\right| \in L^\infty$, so that it satisfies the Euler-Lagrange equation:

$$\frac{d}{dt} \frac{\partial f}{\partial \chi} = \frac{\partial f}{\partial q}$$

Since the function f has been assumed to be convex wrt the second variable, it has a Legendre transform:

$$H(p, q) = \max \{p\chi - f(q, \chi)\} \tag{2.3}$$

and it is well-known that the Euler-Lagrange equation (2.1), which is a second-order equation in R^n, can be rewritten as a Hamiltonian system, which is a first-order equation in R^{2n} :

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \tag{2.4}$$

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} \tag{2.5}$$

If now we fix q_0 and introduce the so-called value function $V(q_1, T)$ associated with the optimization problem (2.1) (2.2), namely:

$$V(q_1, T) := \inf \left\{ \int_0^T f\left(q, \frac{dq}{dt}\right) dt \mid q(0) = q_0, \quad q(T) = q_1 \right\} \tag{2.6}$$

we find that it satisfies a first-order PDE on $R^n \times [0,\ T]$:

$$\frac{\partial V}{\partial T} + H\left(q_1, \frac{\partial V}{\partial q_1}\right) = 0 \tag{2.7}$$

This is the HJB equation. We have approached it through the value function, but the same equation can also be obtained by trying to find a change of variable which would simplify the problem (i.e. generating functions), and this is how it appears in the work of Hamilton and Jacobi.

Note that q_0 and q_1 play symmetric roles, so that a similar equation exists for q_0.

3. Other boundary conditions

From now on we shall simplify notations by writing $x = (p, q)$ and:

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$$

It follows that the system (2.4), (2.5) can be rewritten compactly as $\frac{dx}{dt} = JH'(x)$.

Recall that a matrix M is symplectic if $M^*JM = J$. The group of symplectic matrices in $R^n \times R^n$ will be denoted by $\mathrm{Sp}(n)$. It has dimension $n(2n+1)$. It is a compact Lie group, and the tangent space at I_{2n} is given by:

$$T_{I_{2n}}\mathrm{Sp}(n) = \{m \mid m^*J + Jm = 0\} \tag{3.1}$$

In other words, Jm is symmetric.

The Bolza problem is a natural one when f is coercive, for instance when $f(q,\chi) = \frac{1}{2}q^2 + \frac{1}{2}\chi^2$. Note that in that case, by formula (2.3), we have $H(p,q) = \frac{1}{2}p^2 - \frac{1}{2}q^2$, which is neither convex nor coercive. For convex and coercive Hamiltonians, such as $H(p,q) = \frac{1}{2}p^2 + \frac{1}{2}q^2$, the natural one is to look for periodic solutions, that is, to investigate the problem:

$$\frac{dx}{dt} = JH'(x)$$
$$x(T) = x(0)$$

This problem can be imbedded in a family of problems indexed by $M \in \mathrm{Sp}(n)$

$$\frac{dx}{dt} = JH'(x) \tag{3.2}$$

$$x(T) = Mx(0) \tag{3.3}$$

Here M will play the role devoted to (q_0, q_1) in the Bolza problem: the value function will be $V(M,T)$ instead of $V(q_0,q_1,T)$. Note that we cannot define it by using the least action principle:

$$V(M,T) = \inf\left\{\int_0^T \left[\frac{1}{2}\left(J\frac{dx}{dt}, x\right) + H(x)\right] dt \mid x(T) = Mx(0)\right\}$$

because the right-hand side takes the value $+\infty$ (and would take the value $-\infty$ if we tried to minimize). In fact, a solution to problem (3.2), (3.3) is neither a minimizer nor a maximizer, but a critical point of the right-hand side. So we have to define the value in another way, and for this reason we make some additional assumptions on H

Definition 3.1. Suppose $H : R^{2n} \to R$ is convex, with $H(0) = 0$. It is called subquadratic near infinity if $H(x)|x|^{-2} \to 0$ when $|x| \to \infty$, and subquadratic near 0 if $H(x)|x|^{-2} \to \infty$ when $|x| \to 0$.

We use the results in [3] (see Chapter II.4, notably Proposition 6, Chapter III.3, notably Corollary 6; see also [5])

Theorem 3.2. *Suppose H is convex and subquadratic near 0, and consider the problem:*

$$\inf \int_0^T \left[\left(J\frac{dy}{dt}, y\right) + H^*\left(-J\frac{dy}{dt}\right)\right] dt \tag{3.4}$$

$$y(T) = My(0) \tag{3.5}$$

This problem has a solution $y(t)$ for any $M \in \mathrm{Sp}(n)$, and there is a constant $y_0 \in R^{2n}$ such that $y(t) + y_0$ solves problem (3.2) (3.3). If in addition H is subquadratic near 0, this solution is not constant.

To understand the theorem, note that $y(t)=0$ is always a solution. Note also that, for any $y(t)$, adding a constant y_0 changes the value of the integral by

$$\int_0^T \left(J\frac{dy}{dt}, y_0\right) dt = \left(J\left(y(T)-y(0)\right), y_0\right) = -\left(y(0), \left(M^* - I_{2n}\right) y_0\right)$$

If M has 1 as an eigenvalue, and if y_0 is an eigenvector, the right-hand side vanishes. So the solution $y(t)$ of (3.4) (3.5) is defined modulo a 1-eigenvector y_0 of M, and the latter can be chosen so that $y(t)+y_0$ solves (3.2) (3.3).

We are now in a position to define the function $V(M,T)$ in a proper way:

$$V(M,T) = \inf\left\{\int_0^T \left[\frac{1}{2}\left(J\frac{dy}{dt}, y\right) + H^*\left(-J\frac{dy}{dt}\right)\right] dt \mid y(T) = My(0)\right\} \tag{3.6}$$

The right-hand side is the dual action functional. Note that setting $y=0$ gives the value 0 to the integral, so that $V(M,T)<0$ for $T>0$. The function $V : \mathrm{Sp}(n)\times[0,\ T]\to R$ is well-defined and does not vanish. Let us show that it satisfies a PDE system of the first order.

For the sake of simplicity, assume that M does not have the eigenvalue 1. Set

$$u = \frac{dx}{dt}$$
$$x(t) = x(0) + \int_0^t u(s)\, ds$$

The boundary condition $x(T)=Mx(0)$ becomes:

$$x(0) + \int_0^T u dt = Mx(0) \tag{3.7}$$

$$x(0) = \left(M - I_{2n}\right)^{-1} \int_0^T u dt \tag{3.8}$$

Set $\Pi u(t) := \int_0^t u(s)\, ds$. We have:

$$x(t) = x(0) + \int_0^t u(s)\, ds = \left(M - I_{2n}\right)^{-1} \Pi u(T) + \Pi u(t) \tag{3.9}$$

Writing this into the right-hand side of (3.6), we get:

$$V(M,T) = \inf_u \int_0^T \left[\frac{1}{2}\left(Ju, \left(M-I_{2n}\right)^{-1}\Pi u(T) + \Pi u(t)\right) + H^*(-Ju)\right] dt \tag{3.10}$$

Let us now compute the partial derivatives wrt M and to T.

If $m \in T_{I_{2n}}\mathrm{Sp}(n)$, we have $mM \in T_M\mathrm{Sp}(n)$. Hence, for every m satisfying (3.1), we have, by the envelope theorem:

$$\begin{aligned}
\frac{\partial V}{\partial M}(M,T)\,mM &= -\int_0^T \frac{1}{2}\Big(Ju, (M-I_{2n})^{-1} mM (M-I_{2n})^{-1}\Pi u(T)\Big)\,dt \\
&= -\int_0^T \frac{1}{2}\Big(Ju, (M-I_{2n})^{-1} mMx(0)\Big)\,dt \\
&= -\frac{1}{2}\left(\int_0^T Judt, (M-I_{2n})^{-1} mx(T)\right) \\
&= -\frac{1}{2}\Big(J(M-I_{2n})x(0), (M-I_{2n})^{-1} mx(T)\Big) \\
&= -\frac{1}{2}\Big((M^*-I_{2n})^{-1} J(M-I_{2n})x(0), mx(T)\Big)
\end{aligned}$$

where $u(t)$ is a minimizer in (3.10) and $x(t)$ is given by formula (3.9).

Lemma 3.3. *If M is symplectic, we have:*

$$(M^*-I_{2n})^{-1} J(M-I_{2n}) = -JM$$

Proof. Multiply both sides by M^*-I_{2n}. We get:

$$JM - J = -M^*JM + JM$$

which is true since $M^*JM = J$ □

Finally, we find:

$$\frac{\partial V}{\partial M}(M,T)\,mM = \frac{1}{2}(JMx(0), mx(T)) = \frac{1}{2}(Jx(T), mx(T)) \tag{3.11}$$

To find the partial derivative wrt T, we rewrite the right-hand side of (3.10) as follows:

$$\int_0^1 \left[\frac{1}{2}\frac{1}{T}\left(J\frac{dx}{dt}, x\right) + H^*\left(-J\frac{1}{T}\frac{dx}{dt}\right)\right]Tdt$$

The envelope theorem then yields:

$$\frac{\partial V}{\partial T}(M,T) = \int_0^1 \left[H^*\left(-J\frac{1}{T}\frac{dx}{dt}\right) + \left(\nabla H^*\left(-J\frac{1}{T}\frac{dx}{dt}\right), J\frac{1}{T^2}\frac{dx}{dt}\right)T\right]dt$$

By the Fenchel identity, $H(x) = (\nabla H^*(y), y) - H^*(y)$ for $x = \nabla H^*(y)$, so the integrand is just:

$$-H\left(\nabla H^*\left(-J\frac{1}{T}\frac{dx}{dt}\right)\right)$$

Inverting the equation $\frac{dx}{dt} = JH'(x)$, we have $x = \nabla H^*\left(-J\frac{dx}{dt}\right)$. Finally, we get

$$\frac{\partial V}{\partial T}(M,T) = -\int_0^T H(x)\frac{dt}{T}$$

Bearing in mind that $H(x(t))$ is constant along trajectories of the Hamiltonian system, we get:

$$\frac{\partial V}{\partial T}(M,T) = -H(x(0)) = -H(x(T)) \tag{3.12}$$

4. HJB

We have found the partial derivatives of $V(M,T)$ at any point $(M,T) \in \mathrm{Sp}(n) \times R_+$ such that M does not have the eigenvalue 1. If $u(t)$ is the corresponding minimizer in (3.1), and $x(t)$ is the corresponding solution of (3.2) (3.3) given by formula (3.8), so that $\frac{dx}{dt} = u$, we have:

$$\frac{\partial V}{\partial M}(M,T)\, mM = \frac{1}{2}\left(Jx(T), mx(T)\right) \quad \forall m \in T_{I_{2n}}\mathrm{Sp}(n) \tag{4.1}$$

$$\frac{\partial V}{\partial T}(M,T) = -H(x(T)) \tag{4.2}$$

This is the HJB equation we are seeking. Indeed, we can invert the first equation to express $x(T)$ in terms of $\frac{\partial V}{\partial M}$, say $x(T) = \varphi\left(\frac{\partial V}{\partial M}\right)$ and write the result in the second, getting $\frac{\partial V}{\partial T} = -H \circ \varphi\left(\frac{\partial V}{\partial M}\right)$. Note that equation (4.1) is in reality a system of $2n^2 + n$ equation (one for each $m \in T_{I_{2n}}\mathrm{Sp}(n)$) in $2n$ variables, so that it is overdetermined. However, by the preceding analysis, we have shown that formula (3.10) gives a solution. Let us summarize:

Theorem 4.1. *Suppose $H : R^{2n} \to R$ is convex, $H(0) = 0$, and subquadratic near 0 and infinity. Then the function $V : \mathrm{Sp}(n) \times [0,\ T] \to R$ defined by formula (3.10) is negative for $T > 0$. If M does not have 1 as an eigenvalue, and V is differentiable at (M,T), the HJB relations (4.1) and (4.2)*

Note that there are two terms in the HJB system: the first one, (4.1), does not depend on H, which appears only in the second, (4.2).

Let us give an example. Take $n = 1$, so that M is symplectic if and only if it preserves volume:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Sp}(2) \Longleftrightarrow ad - bc = 1$$
$$m = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in T_{I_{2n}}\mathrm{Sp}(2) \Longleftrightarrow \alpha + \delta = 0$$

Then, for $\xi = (\xi_1, \xi_2)$:

$$\frac{1}{2}(J\xi, m\xi) = \gamma\xi_1^2 - \beta\xi_2^2$$

Let us now take local coordinates in $\mathrm{Sp}(2)$. If $c \neq 0$, for instance, we can take a, b, c and set $d = (ab-1)c^{-1}$. Then $V(M,T)$ becomes $V(a,b,c,T)$ and:

$$\begin{aligned}
\frac{\partial V}{\partial M} mM &= \frac{\partial V}{\partial a}(\alpha a + \beta c) + \frac{\partial V}{\partial b}(\alpha b + \beta d) + \frac{\partial V}{\partial c}(\gamma a + \delta c) \\
&= \alpha\left(a\frac{\partial V}{\partial a} + b\frac{\partial V}{\partial b}\right) + \beta\left(c\frac{\partial V}{\partial a} + d\frac{\partial V}{\partial b}\right) + \gamma\left(a\frac{\partial V}{\partial c}\right) + \delta\left(c\frac{\partial V}{\partial c}\right) \\
&= \alpha\left(a\frac{\partial V}{\partial a} + b\frac{\partial V}{\partial b} - c\frac{\partial V}{\partial c}\right) + \beta\left(c\frac{\partial V}{\partial a} + \frac{ab-1}{c}\frac{\partial V}{\partial b}\right) + \gamma\left(a\frac{\partial V}{\partial c}\right)
\end{aligned}$$

Equation (4.1) becomes:

$$a\frac{\partial V}{\partial a}+b\frac{\partial V}{\partial b}-c\frac{\partial V}{\partial c}=0 \tag{4.3}$$

$$c\frac{\partial V}{\partial a}+\frac{ab-1}{c}\frac{\partial V}{\partial b}=-x_2\left(T\right)^2 \tag{4.4}$$

$$a\frac{\partial V}{\partial c}=x_1\left(T\right)^2 \tag{4.5}$$

We can derive $x_1\left(T\right)$ and $x_2\left(T\right)$ from the last two equations, and plug into second HJB relation (4.2), getting:

$$\frac{\partial V}{\partial T}=-H\left(\pm\sqrt{a\frac{\partial V}{\partial c}},\pm\sqrt{\frac{1-ab}{c}\frac{\partial V}{\partial b}-c\frac{\partial V}{\partial a}}\right)$$

We still have to satisfy equation (4.3). Finally, we get an overdetermined system for $V\left(a,b,c,T\right)$:

$$\begin{aligned}0&=a\frac{\partial V}{\partial a}+b\frac{\partial V}{\partial b}-c\frac{\partial V}{\partial c}\\ \frac{\partial V}{\partial T}&=-H\left(\pm\sqrt{a\frac{\partial V}{\partial c}},\pm\sqrt{\frac{1-ab}{c}\frac{\partial V}{\partial b}-c\frac{\partial V}{\partial a}}\right)\end{aligned}$$

The sign indeterminacy in the second equation arises also in the Bolza problem. For instance, it is found in the classical eikonal equation.

If one takes $H\left(x\right)=\left|x\right|^{\alpha}$ with $0<\alpha<2$, which is convex and subquadratic near 0 and infinity, the system becomes:

$$\begin{aligned}0&=a\frac{\partial V}{\partial a}+b\frac{\partial V}{\partial b}-c\frac{\partial V}{\partial c}\\ 0&=\frac{\partial V}{\partial T}+\left|a\frac{\partial V}{\partial c}+\frac{1-ab}{c}\frac{\partial V}{\partial b}-c\frac{\partial V}{\partial a}\right|^{\frac{\alpha}{2}}\\ a\frac{\partial V}{\partial c}&>0,\quad \frac{1-ab}{c}\frac{\partial V}{\partial b}-c\frac{\partial V}{\partial a}>0\end{aligned}$$

and the problem (3.2) (3.3) can be solved explicitly, yielding a solution to this system

5. Conclusion

This aim of this paper is to open up a problem. There are many questions to be answered:

(1) We have investigated only points where M does not have the eigenvalue 1 and $V\left(M,T\right)$ is differentiable. What happens at other points ? Is it true that the value function $V\left(M,T\right)$ provides a viscosity solution of the system (4.1) (4.2) over $\mathrm{Sp}(n)\times R_+$?
(2) What is the geometry of the solution ? What is the meaning of the indeterminacy which arises when solving (4.1), and which appears as $\pm$ in the example ? Does it mean that the graph of the value function can be extended to a sheet which covers $\mathrm{Sp}(n)\times R_+$ several times, in the manner of a Riemann surface ?

(3) What happens when the Hamiltonian H is no longer subquadratic ? If it is superquadratic, for instance, the dual action principle still holds, and can be used to prove the existence of a solution to the problem (3.2) (3.3), but the minimum on the right-hand side of (3.6) is not attained. The solution is a saddle-point, and defines a critical value rather than a minimum. However, the system (4.1) (4.2) is still valid. Does it have a solution, and is it provided by the analogue of formula (3.10), where one seeks a critical value of the right-hand side ?

(4) Finally, what about general Hamiltonians H, assuming simply that $H(0) = 0$ and that $H(x) \to \infty$ when $|x| \to \infty$, so that energy surfaces $H(x) = h$ are bounded, and the boundary-value problem (3.2) (3.3) is reasonable ?

References

[1] Vladimir I. Arnol'd, *Méthodes mathématiques de la mécanique classique*, Éditions Mir, Moscou, 1974

[2] Théophile de Donder, Théorie invariantive du calcul des variations, Gauthier-Villars, Paris, 1935

[3] Ivar Ekeland, *Convexity methods in Hamiltonian mechanics*, Springe Verlag, 1990

[4] Jean Mawhin, Analyse: fondements, techniques, évolution, Decitre, 1992

[5] Jean Mawhin et Michel Willem, *Critical point theory and Hamiltonian systems*, Springer Verlag, 1989

[6] Charles-Jean de la Vallée-Poussin, *Cours d 'analyse infinitésimale*, Gauthier-Villars, Paris, second edition (1907)

[7] Michel Willem, *Analyse convexe et optimisation*, Editions CIACO, 1989

[8] Michel Willem, *Analyse harmonique réelle*, Hermann, Paris, 1995

[9] Michel Willem, *Principes d 'analyse fonctionnelle*, Cassini, Paris, 1998

CEREMADE, Université Paris-Dauphine
E-mail address: `ekeland@ceremade.dauphine.fr`
URL: `http://www.ceremade.dauphine.fr/ ~ekeland`