Introduction to the Calculus of Variations
Lecture Notes
Lecture 5

Swarnendu Sil

Spring 2021, IISc

Classical Methods

Hamilton-Jacobi equations

Hamilton-Jacobi equations

Now we are going to show that in some cases, a solution to the Hamilton's equations, which are $2N$ first order ODEs

$$\begin{cases} \dot{u}(t) = H_v(t, u(t), v(t)), \\ \dot{v}(t) = -H_u(t, u(t), v(t)). \end{cases} \tag{\textbf{H}}$$

can be furnished by finding a complete integral of a first order PDE. These PDE is called the **Hamilton-Jacobi equation**

$$S_t + H(t, u, S_u) = 0. \tag{\textbf{HJE}}$$

First we begin by showing that if $S = S(t, u)$ satisfies (**HJE**) and $u(t)$ satisfies

$$\dot{u} = H_v(t, u, S_u),$$

then $v = S_u$ satisfies the other equation of (**H**). Also, an m-parameter family of solutions to (**HJE**) yields m first integrals of (**H**).

Theorem 1 (Hamilton-Jacobi equation)**.** *Let $H \in C^1\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$, $H = H(t, u, v)$. Suppose there exists $S \in C^2\left([a,b] \times \mathbb{R}^N\right)$, $S = S(t, u)$, a solution of the* ***Hamilton-Jacobi equation***

$$S_t + H(t, u, S_u) = 0 \qquad \text{for all } (t, u) \in [a,b] \times \mathbb{R}^N. \tag{1}$$

Assume also that there exists $u \in C^1\left([a,b]; \mathbb{R}^N\right)$, a solution of

$$\dot{u}(t) = H_v(t, u, S_u) \qquad \text{for all } t \in [a,b]. \tag{2}$$

Set $v(t) = S_u(t, u(t))$. Then (u, v) is a solution of the Hamilton's equation.

Moreover if $S \in C^2\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^m\right)$ is an m-parameter family of solutions to the Hamilton-Jacobi equation (1)*, then*

$$\frac{\partial S}{\partial \alpha_i} \quad \text{is a first integral of Hamilton's equations for each } 1 \leq i \leq m.$$

Proof. Fix $1 \leq i \leq N$. Differentiating the Hamilton-Jacobi equation w.r.t. u_i, we get

$$S_{u_i t} + H_{u_i} + \left\langle H_v, \frac{\partial}{\partial u_i} S_u \right\rangle = 0.$$

Note that this is justified since $S \in C^2$ and $H \in C^1$. Now since $v(t) = S_u(t, u(t))$, differentiating we obtain

$$\dot{v}_i(t) = S_{tu_i} + \left\langle \frac{\partial}{\partial u} S_{u_i}, \dot{u} \right\rangle.$$

Once again, this is justified since $S \in C^2$ and $H \in C^1$. Also, since $S \in C^2$, for any $1 \leq i \leq N$, we have

$$\begin{aligned} S_{u_i t} &= S_{tu_i}, \\ \frac{\partial}{\partial u_i} S_u &= \frac{\partial}{\partial u} S_{u_i}. \end{aligned}$$

Thus, using (2) and the symmetry of the scalar product, we deduce

$$\dot{v}_i(t) = -H_{u_i}.$$

For the last part, differentiating the Hamilton-Jacobi equation w.r.t. α_i, we get

$$S_{\alpha_i t} + \left\langle H_v, \frac{\partial}{\partial \alpha_i} S_u \right\rangle = 0.$$

So, once again using the fact that the second derivatives of S commute, we have

$$\frac{d}{dt}\left(\frac{\partial S}{\partial \alpha_i}\right) = S_{\alpha_i t} + \left\langle \dot{u}, \frac{\partial}{\partial u}\left[\frac{\partial S}{\partial \alpha_i}\right] \right\rangle = \left\langle \dot{u} - H_v, \frac{\partial}{\partial u}\left[\frac{\partial S}{\partial \alpha_i}\right] \right\rangle = 0.$$

This proves $\frac{\partial S}{\partial \alpha_i}$ is a first integral and completes the proof. □

Now we present another theorem whose conclusion is stronger, in the sense that it furnishes us with a general solution of the Hamilton's equation, but it needs us to find a complete integral of the Hamilton-Jacobi equation instead of a single solution.

Theorem 2 (Jacobi's theorem)**.** *Let $S \in C^2\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$, written as $S = S(t, u_1, \ldots, u_N, \alpha_1, \ldots, \alpha_N)$ be a* complete integral *of the Hamilton-Jacobi equation, i.e. a general solution of* (**HJE**) *depending on N parameters $\alpha_1, \ldots, \alpha_N$. Let*

$$\det\left(\frac{\partial^2 S(t,u,\alpha)}{\partial \alpha \partial u}\right) \neq 0 \quad \text{for every } (t,u,\alpha) \in [a,b] \times \mathbb{R}^N \times \mathbb{R}^N$$

and let $\beta_1, \ldots, \beta_N$ be N arbitrary constants. Then the N parameter family of $\mathbb{R}^N$-valued functions $u(t) = u(t, \alpha_1, \ldots, \alpha_N, \beta_1, \ldots, \beta_N)$ defined by the relations

$$\frac{\partial}{\partial \alpha_i} S(t, u_1, \ldots, u_N, \alpha_1, \ldots, \alpha_N) = \beta_i \qquad \text{for } 1 \leq i \leq N,$$

together with the N parameter family of $\mathbb{R}^N$-valued functions

$$v_i = \frac{\partial}{\partial u_i} S\left(t, u_1, \ldots, u_N, \alpha_1, \ldots, \alpha_N\right) \qquad \textit{for } 1 \leq i \leq N$$

constitute a general solution *of the Hamilton's equations* (**H**).

Proof. Note that since

$$\det\left(\frac{\partial^2 S}{\partial \alpha \partial u}\right) \neq 0 \quad \text{ for every } (t, u, \alpha) \in [a, b] \times \mathbb{R}^N \times \mathbb{R}^N,$$

it is indeed possible, using the implicit function [1] theorem, to determine u as a function of t, α and β from the relations

$$\frac{\partial}{\partial \alpha_i} S\left(t, u_1, \ldots, u_N, \alpha_1, \ldots, \alpha_N\right) = \beta_i \qquad \text{for } 1 \leq i \leq N. \tag{3}$$

Once we have determined u, we can define v via the equations

$$v_i = \frac{\partial}{\partial u_i} S\left(t, u_1, \ldots, u_N, \alpha_1, \ldots, \alpha_N\right) \qquad \text{for } 1 \leq i \leq N. \tag{4}$$

So all we need to show is that the pair (u, v) so constructed satisfy (**H**). Differntiating(3) w.r.t. t, we obtain

$$0 = \frac{d}{dt}\left(\frac{\partial}{\partial \alpha_i} S\right) = \left\langle \dot{u} - H_v, \frac{\partial}{\partial u}\left[\frac{\partial S}{\partial \alpha_i}\right]\right\rangle \text{ for } 1 \leq i \leq N.$$

This implies $\dot{u} = H_v$. Now, differentiating (4), we obtain as before

$$\dot{v}_i\left(t\right) = S_{tu_i} + \left\langle \frac{\partial}{\partial u} S_{u_i}, \dot{u}\right\rangle = S_{tu_i} + \left\langle \frac{\partial}{\partial u} S_{u_i}, H_v\right\rangle \quad \text{ for } 1 \leq i \leq N.$$

Again as before, we differentiate (**HJE**) to deduce

$$S_{u_i t} + H_{u_i} + \left\langle H_v, \frac{\partial}{\partial u_i} S_u\right\rangle = 0 \quad \text{ for } 1 \leq i \leq N.$$

These clearly imply

$$\dot{v} = -H_u.$$

This completes the proof. □

[1] To be even more explicit, check the hypotheses and apply the implicit function theorem on the vector valued function $F : [a, b] \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$, defined componentwise by the equations

$$F_i\left(t, u, \alpha, \beta\right) := \frac{\partial}{\partial \alpha_i} S\left(t, u_1, \ldots, u_N, \alpha_1, \ldots, \alpha_N\right) - \beta_i \qquad \text{for } 1 \leq i \leq N.$$

Geometric content of the HJE

The strange looking function S might appear to drop out of nowhere, but it actually has a geometric meaning. Let $A = (t_0, x_0), B = (t_1, x_1)$ be two points in $[a, b] \times \mathbb{R}^N$ such that there is a unique integral curve of (**H**) that passes through those two points. Then the value of the integral

$$\int_{t_0}^{t_1} f(t, u(t), \dot{u}(t)) \, \mathrm{d}t,$$

where u is the unique integral curve joining A and B, clearly depends upon the endpoints A and B and is usually known as the **geodesic distance** between A and B. As the prototypical example, this reduces to **the usual distance when the Lagrangian is arc length**. Now if we fix the point A, clearly the above integral can then be expressed as a function of the coordinates of point B, which is precisely our function S.

On the other hand, if the you are familiar with the method of characteristics technique for solving first order PDEs, then you would instantly recognize that (**H**) is nothing but the characteristic system for the nonlinear first order PDE (**HJE**).

There are other important geometric aspect of the Hamilton-Jacobi equation. Very briefly, if you think of a disturbance propagating in a possibly anisotropic, nonhomogeneous medium, there are essentially two viewpoints by which such a phenomena can be described. One is the so-called **'ray viewpoint'**, in which we follow the trajectories of disturbance propagation. The other is the so-called **'wave front viewpoint'**, where we describe the time evolution of the boundary of the disturbed region, called **wave fronts**. The analogy of the terminology with optics (and wave propagation in general) is quite deliberate and intended. Indeed, the equivalence of the two descriptions is known in physics as the Huygen's principle. Here what we proved is exactly a rigorous mathematical formulation of the Huygen's principle. **The wave front approach leads to the Hamilton-Jacobi PDE, which describes the evolution of the wave front (given by the zero set of the function S)**. On the other hand, **the ray approach leads to the Hamilton's equations**. This has also deep connections with the **theory of optimal control**, where the **Pontrjagin Maximum principle** is an analogue (and a generalization) of the Hamilton's equations and the 'ray viewpoint', whereas the **Hamilton-Jacobi-Bellman equations or HJB equations** are the analogue (and a generalization) of the Hamilton-Jacobi equation and the wave-front viewpoint. However, discussing these ideas precisely would need a much more detailed study of envelopes and convex duality than we have done so far. Interested reader can see [2] and [1] for much more on these topics.

Second Variation

So far in this chapter we were concerned with **any critical point**. Now we want to investigate necessary and sufficient conditions for a given critical point to be a minimizer of the functional. We begin with a simple result which calculates the second variation and gives a necessary criterion a given critical point to be a minimizer in terms of the second variation.

Theorem 3 (Second Variation)**.** *Let $f = f(t, u, \xi) \in C^3\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$, $\alpha, \beta \in \mathbb{R}^N$ be given and $X = \left\{u \in C^1\left([a,b]\,; \mathbb{R}^N\right) : u(a) = \alpha,\ u(b) = \beta\right\}$. Consider the problem*

$$(P) \qquad \inf_{u \in X}\left\{ I(u) = \int_a^b f(t, u(t), \dot{u}(t))\ \mathrm{d}t \right\} = m.$$

If $\bar{u} \in X \cap C^2\left([a,b]\,; \mathbb{R}^N\right)$ is a minimizer for (P), then the following integral

$$\int_a^b \left[\langle f_{uu}(t, \bar{u}, \dot{\bar{u}})\, \psi, \psi \rangle + 2\left\langle f_{u\xi}(t, \bar{u}, \dot{\bar{u}})\, \psi, \dot{\psi} \right\rangle + \left\langle f_{\xi\xi}(t, \bar{u}, \dot{\bar{u}})\, \dot{\psi}, \dot{\psi} \right\rangle \right] \qquad (5)$$

*is **nonnegative** for any $\psi \in C_c^1\left([a,b]; \mathbb{R}^N\right)$.*

Proof. As we did in deriving the EL equations, we take $\psi \in C_c^1\left([a,b]; \mathbb{R}^N\right)$. Thus for any $h \in \mathbb{R}$, we have $\bar{u} + h\psi \in X$. Now we define the function $g : \mathbb{R} \to \mathbb{R}$ by $g(h) := I(\bar{u} + h\psi)$.

Then $g \in C^2(\mathbb{R})$ (Check!) and since $\bar{u}$ is a minimizer, g must have a local minima at 0. Thus we must have $g''(0) \geq 0$. But

$$g''(0) = \left. \frac{d^2}{dh^2}\left[I(\bar{u} + h\psi) \right] \right|_{h=0}.$$

The rest is a straight forward calculation. Note that since $f \in C^3$, thus in particular C^2 and thus $f_{u\xi}$, f_{uu} and $f_{\xi\xi}$ are all symmetric matrices. $\square$

Quadratic functional related to second variation

To understand the expression for the second variation better, we integrate by parts in the mixed term and obtain

$$2\int_a^b \left\langle f_{u\xi}(t, \bar{u}, \dot{\bar{u}})\, \psi, \dot{\psi} \right\rangle = -\int_a^b \left\langle \frac{d}{dt}\left[f_{u\xi}(t, \bar{u}, \dot{\bar{u}}) \right] \psi, \psi \right\rangle.$$

Note that here again we used the fact that $f_{u\xi}$ takes values in the space of symmetric matrices. In view of this, we can rewrite the expression (5) as

$$\int_a^b \left[\left\langle \left[f_{uu}(t, \bar{u}, \dot{\bar{u}}) - \frac{d}{dt} f_{u\xi}(t, \bar{u}, \dot{\bar{u}}) \right] \psi, \psi \right\rangle + \left\langle f_{\xi\xi}(t, \bar{u}, \dot{\bar{u}})\, \dot{\psi}, \dot{\psi} \right\rangle \right].$$

The important point here is that the matrix $f_{\xi\xi}$ plays the dominant role here in determining whether this quadratic form will be nonnegative or not.

The heuristic argument is, since ψ vanishes at the boundary, we have a Poincaré inequality. Roughly, since the function vanishes at the boundary, the value of the function itself can not be large while keeping its derivative small, since it has climb up from zero to the high values. But the converse is quite possible! The function can be small with large derivative! Why? It can oscillate a lot!

We now formalize the heuristic argument.

Lemma 4. *If the following inequality*

$$\int_a^b \left[\left\langle \left[f_{uu}\left(t, \bar{u}, \dot{\bar{u}}\right) - \frac{d}{dt} f_{u\xi}\left(t, \bar{u}, \dot{\bar{u}}\right) \right] \psi, \psi \right\rangle + \left\langle f_{\xi\xi}\left(t, \bar{u}, \dot{\bar{u}}\right) \dot{\psi}, \dot{\psi} \right\rangle \right] \; \mathrm{d}t \; \geq \; 0,$$

holds for every $\psi \in C_c^1\left([a,b]; \mathbb{R}^N\right)$, *then the matrix* $f_{\xi\xi}\left(t, \bar{u}, \dot{\bar{u}}\right)$ *is nonnegative definite for every* $t \in (a,b)$.

Bibliography

[1] Gelfand, I. M., and Fomin, S. V. *Calculus of variations.* Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[2] Zeidler, E. *Nonlinear functional analysis and its applications. III.* Springer-Verlag, New York, 1985. Variational methods and optimization, Translated from the German by Leo F. Boron.