Hamilton – Jacobi Equations

The main problem to be discussed in this paper is to solve the following:

$$\begin{cases} u_t + H(D_x u, x) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \quad (1) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \quad (2) \end{cases}$$

This is known as the Hamilton-Jacobi equation; physically it represents an elegant and mathematically sophisticated reformulation of Newtonian mechanics (1687), and is equivalent to Lagrangian mechanics (1788) and Hamiltonian mechanics (1833), which are other reformulations of Newtonian mechanics that trade the use of vectors to solve problems for more differential equations.

The Hamiltonian, represented by $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, as well as the initial function $g : \mathbb{R}^n \to \mathbb{R}$, are specified. The problem is to find the function $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$, (known as Hamilton's principle function) where $u = u(x,t)$ and $D_x u = (u_{x_1}, \ldots.., u_{x_n})$. We define $p \equiv D_x u$, $p_{n+1} \equiv u_t$ so that now (1) becomes $p_{n+1} + H(p,x) = 0$.

We will start by considering the following nonlinear first order partial differential equation:

$$F(D_x u, x) = 0 \ \text{ in } \Omega,$$

where $x = (x_1, \ldots x_n)$, F is a smooth function, and u=u(x) $\in C^2(\Omega)$. Now making use of the method of characteristics, we let $x = x(s)$, $p(s) \equiv D_x u(x(s))$, $\dot{} \equiv \frac{d}{ds}$, and we compute derivatives:

$$\dot{p}^k = u_{x_k x_j}(x(s))\dot{x}^j \quad (3)$$

where we employ Einstein summation notation over the repeated index $j = 1, \ldots., n$ (we will use this notation throughout this paper). We next differentiate $F(D_x u, x)$ with respect to x_k and evaluate it at x=x(s), and get

$$\frac{\partial F}{\partial p_j}(p(s), x(s)) u_{x_j} x_k + \frac{\partial F}{\partial x_k}(p(s), x(s)) = 0$$

$$\implies \frac{\partial F}{\partial p_j}(p(s), x(s)) u_{x_j} x_k = -\frac{\partial F}{\partial x_k}(p(s), x(s)) \quad (4)$$

Thus setting

$$\dot{x^j}(s) = \frac{\partial F}{\partial p_j}(p(s), x(s)) \quad (5)$$

and substituting this into (4), we can use (3) to see that

$$\dot{p}^k = -\frac{\partial F}{\partial x_k}(p(s), x(s))\,. \quad (6)\,.$$

Putting together (5) and (6) and using more compact notation, we have

$$\begin{cases} \dot{p}(s) = -D_x F(p(s), x(s)) & (7) \\ \dot{x}(s) = D_p F(p(s), x(s)) & (8) \end{cases}$$

We will now make use of (7), (8) for the main problem (with $p \to (p,\, p_{n+1}),\ x \to (x, t)$): Let

$$F((p,\, p_{n+1})\,, (x\,, t)) = p_{n+1} + H(p, x) = 0,$$

$$\implies D_{(p,\, p_{n+1})} F = (D_p H, 1),\ D_{(x,\, t)} F = (D_x H, 0),$$

then by comparing with (7) and (8), we get

$$\begin{cases} \dot{p}(s) = -D_x H(p(s),\, x(s)) & (9) \\ \dot{x}(s) = D_p H(p(s),\, x(s))\,. & (10) \end{cases}$$

This is a system of 2n first order ordinary differential equations, and it is comprised of the characteristic equations for the Hamilton-Jacobi equations; they are known as Hamilton's equations.

Lagrangian:

Consider a specified smooth function, $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, which we will call the Lagrangian. We introduce the functional, called the action, defined by

$$I[x] = \int_0^t \mathcal{L}(\dot{x}(s), x(s))\, ds$$

where $x \in A \equiv \{f \in C^2([0, t]; \mathbb{R}^n) : f(0) = y,\ f(t) = z\}$. We will look for the $y \in A$ such that $I[y] = \min_{x \in A} I[x]$. (11).

Physically, letting $\mathcal{L} = T - V$ (the kinetic energy associated with a particle minus its potential energy), then Hamilton's principle of least action says nature minimizes the action.

Theorem : If the function y solves (11), then y satisfies the Euler-Lagrange equations:

$$\frac{d}{ds} D_{\dot{y}} \mathcal{L}(\dot{y}, y) - D_y \mathcal{L}(\dot{y}, y) = 0\,.$$

Proof : Let $\epsilon \in \mathbb{R}$, $v = (v_1, ..., v_n) : [0,t] \to \mathbb{R}^n$ be a smooth function satisfying $v(0) = v(t) = 0$. Then

$$I[y] \leq I[y + \epsilon v]$$

so the right hand side has a minimum at $\epsilon = 0$, so differentiating the right hand side with respect to epsilon, we have that

$$I'[y] = 0.$$

$$I[y + \epsilon v] = \int_0^t \mathcal{L}(\dot{y} + \epsilon \dot{v}, y + \epsilon v)\, ds$$

$$\implies I'[y + \epsilon v] = \int_0^t \mathcal{L}_{\dot{y}_i}(\dot{y} + \epsilon \dot{v}, y + \epsilon v)\dot{v}_i + \mathcal{L}_{y_i}(\dot{y} + \epsilon \dot{v}, y + \epsilon v) v_i\, ds$$

$$\implies 0 = I'[y] = \int_0^t \mathcal{L}_{\dot{y}_i}(\dot{y}, y)\dot{v}_i + \mathcal{L}_{y_i}(\dot{y}, y) v_i\, ds\,,$$

and after integrating by parts on the left term of the integrand (and using the fact that v vanishes at the end points), we get

$$0 = I'[y] = \int_0^t \Big(- \frac{d}{ds} \mathcal{L}_{\dot{y}_i}(\dot{y}, y) + \mathcal{L}_{y_i}(\dot{y}, y) \Big) v_i\, ds\,.$$

Since this is true for all smooth v which satisfy the boundary conditions, this implies that

$$-\frac{d}{ds} \mathcal{L}_{\dot{y}_i}(\dot{y}, y) + \mathcal{L}_{y_i}(\dot{y}, y) = 0$$

for all $0 \leq s \leq t$, $i = 1, ..., n$. $\square$

Note that the converse is not true; a function that satisfies the Euler-Lagrange equations is called a critical point of $I[\cdot]$, and not all critical points are minimizers.

Connecting the Hamiltonian with the Lagrangian :

Suppose that for all x, $p \in \mathbb{R}^n$ we can uniquely solve for a smooth function $\dot{x} = \dot{x}(p, x)$ the equation $p = D_{\dot{x}}\mathcal{L}(\dot{x}, x)$, then for a particilular Lagrangian $\mathcal{L}$, with $q \equiv \dot{x}(x, p)$,the associated Hamiltonian is

$$\mathcal{H}(p, x) \equiv p \cdot q(p, x) - \mathcal{L}(q(p, x), x).$$

Theorem : The functions x and p satisfy Hamilton's equations ((9) and (10)) and $H(p(s), x(s))$ is constant in s.

Proof :

$$\partial_{x_i}\mathcal{H} = p_k \cdot \partial_{x_i} q_k - \partial_{q_k}\mathcal{L}\, \partial_{x_i} q_k - \partial_{x_i}\mathcal{L}$$

and using that $p = D_q\mathcal{L}$, we have that

$$\partial_{x_i}\mathcal{H} = -\partial_{x_i}\mathcal{L} = -d_s \partial_{\dot{x}_i}\mathcal{L} = -\dot{p}_i \,.$$

$$\partial_{p_i}\mathcal{H} = q_i + p_k \partial_{p_i} q_k - \partial_{q_k}\mathcal{L}\, \partial_{p_i} q_k$$

and for the same reasoning, we thus have

$$\partial_{p_i} H = q_i = \dot{x}_i.$$

Finally, we have

$$d_s\mathcal{H}(p(s), x(s)) = \dot{p}_i \partial_{p_i}\mathcal{H} + \dot{x}_i \partial_{x_i}\mathcal{H}$$

$$= -\partial_{x_i}\mathcal{H}\, \partial_{p_i}\mathcal{H} + \partial_{p_i}\mathcal{H} \partial_{x_i}\mathcal{H} = 0\,. \;\square$$

This is not the only connection between the Hamiltonian and the Lagrangian; they are acutally dual convex functions. Basically, if we postulate that the mapping $q \mapsto \mathcal{L}(q)$ is convex (for all $t \in [0,1]$ and $q_1,\, q_2 \in \mathbb{R}^n$, $\mathcal{L}(q_2 - t[q_2 - q_1]) \leq \mathcal{L}(q_2) - t[\mathcal{L}(q_2) - \mathcal{L}(q_1)]$), and that $\lim_{q \to \infty} \frac{\mathcal{L}(q)}{|q|} = \infty$, we can use the Legendre transform and write $\mathcal{H}(p) = \max_{q \in \mathbb{R}^n} [pq - \mathcal{L}(q)]$ for $p \in \mathbb{R}^n$; note that the maximum necessarily exists due to the postulates. Now it turns out that $\mathcal{H}$ satisfies the postulates we laid out on $\mathcal{L}$, and that $\mathcal{L}$ is the Legendre transform of $\mathcal{H}$; in othe words, like Kelvin transform, the Legendre transform is an involution.

Suppose for a minute that the equation is differentiable, then if we want to find q^* that maximizes it, we can differentiate and set to zero:

$$\frac{d\mathcal{H}(q_i^*)}{dq_i} = p - \partial_{q_i}\mathcal{L}(q_i^*) = 0 \implies p = \partial_{q_i}\mathcal{L}(q_i^*)$$

Now we just need to invert this last for $q^* = q^*(p)$, and then we would have our Hamiltonian.

Proof of Duality : Let $t \in [0,1]$. Then

$$\mathcal{H}(p_2 - t(p_2 - p_1)) = \sup_q [(p_2 - t(p_2 - p_1))q - \mathcal{L}(q)] = \sup_q ((tp_1 + (1-t)p_2)q - \mathcal{L}(q))$$

$$\leq t \sup_q (p_1 q - \mathcal{L}(q)) + (1-t) \sup_q (p_2 q - \mathcal{L}(q))$$

$$= t\mathcal{H}(p_1) + (1-t)\mathcal{H}(p_2) = \mathcal{H}(p_2) - t[\mathcal{H}(p_2) - \mathcal{H}(p_1)]$$

and so $\mathcal{H}(p)$ is convex.

Now pick a $\lambda > 0$, and let $q = \lambda \frac{p}{|p|}$, then

$$\mathcal{H}(p) = \sup_{q \in \mathbb{R}^n} [pq - \mathcal{L}(q)] \geq \lambda |p| - \mathcal{L}\Big(\lambda \frac{p}{|p|}\Big)$$

$$\geq \lambda |p| - \sup_{B(0,\lambda)} \mathcal{L}$$

and since λ was arbitrary,

$$\implies \lim_{|p| \to \infty} \frac{\mathcal{H}(p)}{|p|} = \infty.$$

Now getting back to the "involutionism", it is clear that

$$\mathcal{H}(p) \geq pq - \mathcal{L}(q) \implies \mathcal{L}(q) \geq pq - \mathcal{H}(p)$$

$$\implies \mathcal{L}(q) \geq \sup_p (pq - \mathcal{H}(p)) = H^*(q)$$

where the * indicates the Legendre transform of the function. Now we also have that

$$H^*(q) = \sup_p [pq - \sup_r (pr - \mathcal{L}(r))]$$

and using that $\sup(-x) = -\inf(x)$, we get

$$H^*(q) = \sup_p \inf_r [p(q-r) + \mathcal{L}(r)] \, .$$

Now $\mathcal{L}(q)$ is convex, so by the supporting hyperplanes theorem, there exists an $s \in \mathbb{R}^n$ such that $\mathcal{L}(r) \geq \mathcal{L}(q) + s(r-q)$ for all $r \in \mathbb{R}^n$ (if $\mathcal{L}$ is differentiable at q, then letting $s = \nabla \mathcal{L}(q)$ gives the more common condition for convexity: that the function lies above all its tangents) . So letting $p = s$, we have that

$$H^*(q) = H^*(q) = \sup_p \inf_r [p(q-r) + \mathcal{L}(r)] \geq \inf_r [s(q-r) + \mathcal{L}(r)] \, ,$$

and since $s(q-r) + \mathcal{L}(r) \geq \mathcal{L}(q)$, letting $r = q$, we get

$$H^*(q) \geq \mathcal{L}(q)$$

and so finally we see that $\mathcal{H}^*(q) = \mathcal{L}(q)$. $\square$

Now going back to the original problem, the Hamilton-Jacobi equation, we have already seen the deep connection between the Euler-Lagrange equations, Hamilton's ODEs, and the action, we can make an ansantz guess that there is also a connection between the action and the Hamilton-Jacobi equation.

Claim:

$$u(x,t) = \inf \left\{ \int_0^t \mathcal{L}(\dot{w}(s)) \, ds + g(y) : w \in C^1, \ w(0) = y, \ w(t) = x \right\}$$

solves the Hamilton-Jacobi equation. For simplification purposes it will be assumed that $g : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous.

Theorem: (Hopf-Lax forumula). If $x \in \mathbb{R}^n$ and $t > 0$, then

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left\{ t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y) \right\}.$$

Proof : Let $y \in \mathbb{R}^n$, and $w(s) \equiv y + \frac{s}{t}(x-y)$ for $0 \le s \le t$. Then we have

$$u(x,t) \le \int_0^t \mathcal{L}(\dot{w}(s))\,ds + g(y) = t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y)$$

$$\implies u(x,t) \le \inf_{y \in \mathbb{R}^n} \left\{ t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y) \right\}.$$

Now, let $w \in C^1$ be such that $w(0) = y,\ w(t) = x$, then since $\mathcal{L}$ is convex, we can use Jensen's inequality and write

$$\mathcal{L}\left(\frac{1}{t}\int_0^t \dot{w}(s)\,ds\right) \le \frac{1}{t}\int_0^t \mathcal{L}(\dot{w}(s))\,ds$$

and upon carrying out the integration in the argument of the left hand side, we see that

$$t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y) \le \int_0^t \mathcal{L}(\dot{w}(s))\,ds + g(y).$$

Since y was arbitrary, this gives

$$\inf_{y \in \mathbb{R}^n} \left\{ t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y) \right\} \le \int_0^t \mathcal{L}(\dot{w}(s))\,ds + g(y)$$

taking the infimum over w, and combining the result the last inequality gives

$$\inf_{y \in \mathbb{R}^n} \left\{ t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y) \right\} = u(x,t)\,.$$

It is left to show that the infimum is a minimum: Set $y = x$, then we have $u(x,t) \le t\mathcal{L}(0) + g(x)$. Then since $\lim_{|q| \to \infty} \frac{\mathcal{L}(q)}{|q|} = \infty$, there exists a constant $C > 0$ such that $t\mathcal{L}(q) \ge 2(K+1)q$ for all $|q| > C$ where K is the Lipschitz constant associated with g. Then if $|x-y| \ge tC$, we have

$$t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y) \ge 2(K+1)|x-y| + g(y)$$

$$= (K+2)|x-y| + K|x-y| + g(y) \geq (K+2)|x-y| + g(x)$$

(I used that $K|x-y| = \sup_{x \neq y \in \mathbb{R}^n} \left\{ \frac{|g(x)-g(y)|}{|x-y|} \right\} |x-y| \geq g(x) - g(y)$)

$$\geq (K+2)|x-y| - t\mathcal{L}(0) + u(x,t).$$

Thus letting $D = \max\left\{C, \frac{\mathcal{L}(0)}{K+2}\right\}$, we have that

$$t\mathcal{L}\left(\frac{x-y}{t}\right) + g(y) \geq u(x,t)$$

provided $|x-y| \geq tD$, and so evidently the infimum is a minimum. $\square$.

Theorem : Suppose $x \in \mathbb{R}^n$, $t > 0$, and that u is defined by the Hopf-Lax formula and is differentiable at (x,t). Then

$$u_t(x,t) + \mathcal{H}(D(u)(x,t)) = 0.$$

Remark : For all $x \in \mathbb{R}^n$ and $0 \leq s < t$, we have that

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left\{ (t-s)\mathcal{L}\left(\frac{x-y}{t-s}\right) + u(y,s) \right\}$$

Proof of Theorem : Let $q \in \mathbb{R}^n$, $h > 0$. Then using the remark, we see that

$$u(x+hq, t+h) = \min_{y \in \mathbb{R}^n} \left\{ h\mathcal{L}\left(\frac{x+hq-y}{h}\right) + u(y,t) \right\}$$
$$\leq h\mathcal{L}(q) + u(x,t)$$

as we see by setting y=x. Therefore we have that

$$\frac{u(x+hq,t+h) - u(x,t)}{h} \leq \mathcal{L}(q)$$

Now taking $h \to 0^+$, we get that

$$q \cdot Du(x,t) + u_t(x,t) \leq \mathcal{L}(q)$$

and since this is true for all $q \in \mathbb{R}^n$, we get that

$$u_t(x,t) + \mathcal{H}(Du(x,t)) = u_t(x,t) + \max_{q \in \mathbb{R}^n}\{q \cdot Du(x,t) - \mathcal{L}(q)\} \leq 0$$

$$\implies u_t(x,t) + \mathcal{H}(Du(x,t)) \leq 0.$$

Now we just need to get the same inequality reversed. Let $h > 0$, and take z such that $u(x,t) = t\mathcal{L}\big(\frac{x-z}{t}\big) + g(z)$. Let $s = t - h$, $y = \frac{s}{t}x + (1 - \frac{s}{t})z$; then $\frac{x-z}{t} = \frac{y-z}{s}$, and so we get

$$u(x,t) - u(y,s) \geq t\mathcal{L}\left(\frac{x-z}{t}\right) + g(z) - \left(s\mathcal{L}\left(\frac{y-z}{s}\right) + g(z)\right)$$
$$= (t-s)\mathcal{L}\left(\frac{x-z}{t}\right).$$

Making the substitutions for h, we find that

$$\frac{u(x,t) - u(\big(1 - \frac{h}{t}\big) + \frac{h}{t}z, t-h)}{h} \geq \mathcal{L}\left(\frac{x-z}{t}\right).$$

Now take $h \to 0^+$, and arrive at

$$\frac{x-z}{t} \cdot Du(x,t) + u_t(x,t) \geq \mathcal{L}\left(\frac{x-z}{t}\right).$$

Finally we see that

$$u_t(x,t) + \mathcal{H}(Du(x,t)) = u_t(x,t) + \max_{q \in \mathbb{R}^n}\{q \cdot Du(x,t) - \mathcal{L}(q)\}$$

$$\geq u_t(x,t) + \frac{x-z}{t} \cdot Du(x,t) - \mathcal{L}\left(\frac{x-z}{t}\right) \geq 0$$
$$\implies u_t(x,t) + \mathcal{H}(Du(x,t)) \geq 0$$

and so we have equality. $\square$

So we have arrived at the climax of the paper, that u defined by the Hopf-Lax formula solves

$$\begin{cases} u_t + H(D_x u, x) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

References

(1) Lawrence Craig Evans, Partial differential equations. AMS 1998.

(2) http://idv.sinica.edu.tw/ftliang/pde/*PDE_1/first_order/hamilton_jacobi/hopf-lax.pdf

(3) https://www.math.ualberta.ca/~xinweiyu/527.1.11f/lec06.pdf