Hamilton Jacobi equations

Intoduction to PDE

The rigorous stuff from Evans, mostly.

We discuss first

$$\partial_t u + H(\nabla u) = 0, \tag{1}$$

where $H(p)$ is convex, and superlinear at infinity,

$$\lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty$$

This by comes by integration from special hyperbolic systems of the form $(n = m)$

$$\partial_t v + F_j(v)\partial_j v = 0$$

when there exists a pontental for F_j, i.e. $F_j = \partial_j H(v)$, and when we seek solutions as gradients $v = \nabla u$. Typical example $(n = 1)$ is Burgers equation

$$v_t + v v_x = 0$$

which by integration $v = u_x$ gives, omitting constants,

$$u_t + \frac{1}{2}u_x^2 = 0.$$

We know solutions of Burgers or conservation laws conserve L^∞ norms and produce shocks. So we expect solutions of HJ equations to be Lipschitz, and loose second derivatives. That they do. We recall that the Legendre transform of a convex superlinear-at-infinity function L is

$$L^*(p) = \sup_{q\in\mathbb{R}^n}[p\cdot q - L(q)]$$

Note that L^* is convex and superlinear at infinity and

$$(L^*)^* = L.$$

Let $H = L^*$. Note that, if L is smooth (and it is, if H is strictly convex), then at a maximum (they are attained becuse of superlinearity)

$$\nabla_q L(q) = p,$$

so, solving for q gives a function of p. (This is unique if L is strictly convex). So,

$$H(p) = p \cdot q(p) - L(q(p))$$

Differentiating this, we have

$$\nabla_p H(p) = q(p) + p \cdot \nabla_p q - \nabla_q L(q(p)) \nabla_p q = q(p)$$

This can be understood: the Legendre transform of L is given by

$$H(p) = p \cdot q(p) - L(q(p))$$

where q solves $\nabla_q L = p$, or by the equation

$$H(p) = p \cdot \nabla_p H - L(\nabla_p H(p))$$

which looks complicated, and itself it is a steady Hamilton-Jacobi equation. Similarly, $H^*(q) = L(q)$ is computed by solving $\nabla_p H(p) = q$ and setting $L(q) = p(q) \cdot q - H(p(q))$. Of course, if L is not strictly convex we do not have necessarily a unique solution, and the minimization is necessary.

Now we have a magical solution of (1),

$$u(x,t) = \min_y \left\{ tL\left(\frac{x-y}{t}\right) + u_0(y) \right\} \tag{2}$$

where u_0 is the value of u at $t = 0$. This is termed the Hopf-Lax formula in Evans. The magic is then enveloped in mystery:

$$u(x,t) = \min_A \left\{ \int_0^t L(\dot{w}(s))ds + u_0(w(0)) \right\}$$

where

$$A = \{w \in C^1([0,t], \mathbb{R}^n) | w(t) = x\}$$

and $\dot{w}(s) = \frac{dw}{ds}$. This is easier to prove than it looks (see Evans), and makes contact with minimum action principles, or control, but does not illuminate the Hopf Lax formula by even a single lumen. By what divination process could somebody arrive at such an amazing formula?

Let us differentiate our equation (1), motivated by where it came from. We obtain

$$\partial_t v + (\nabla_p H(v))\nabla v = 0$$

where $v = \nabla u$. So, let us introduce characteristics,

$$\frac{dx}{dt} = \nabla_p H(v(x,t)), \quad x(0) = x_0.$$

Denote for a second $x(t, x_0)$ the characteristic issued from x_0. On it, v is constant, $v(x(t, x_0), t) = v(x_0, 0)$. But $v = \nabla u$, which is the sole argument of $\nabla_p H$. So the characteristics are straight lines

$$x(t, x_0) = x_0 + tq, \quad q = \nabla_p H(p)$$

and

$$\nabla u(x,t) = p, \qquad \text{for } x = x_0 + t\nabla_p H(p).$$

Now let us look at u on the characteristic. We differentiate

$$\frac{d}{ds}u(x_0 + sq, s) = \partial_s u(x_0 + sq, s) + q \cdot \nabla u(x_0 + sq, s)$$

But we are on the characteristic, so $\nabla u = p$, and $q = \nabla_p H(p)$. Using the equation (1) we also have $\partial_s u(x_0 + sq, s) = -H(p)$. The right-hand side is a constant, $p \cdot q - H(p)$. That we know how to integrate in time:

$$u(x,t) - u(x_0, 0) = t\,(p \cdot q - H(p))$$

We recall that $q = \nabla_p H(p)$. Wait a minute, that means that $p \cdot q - H(p) = L(q)$. So

$$u(x,t) = tL(q) + u_0(x_0)$$

Now we express q in terms of x and x_0: From $x = x_0 + tq$, it follows that $q = \frac{x - x_0}{t}$. So, without magic, we arrived at

$$u(x,t) = tL\left(\frac{x - x_0}{t}\right) + u_0(x_0).$$

Now, we assumed strict convexity, and smooth initial data, and this would work only for short time. Remarkably, the variational formula is true for all time and gives the good weak solution even after ∇u develops shocks. When they do, x_0 is not uniquely determined on characteristics by $\nabla u(x,t)$, x an t. We start with a semigroup property:

Lemma 1. *If H is convex and superlinear at infinity and if u is given by (2), then*

$$u(x,t) = \min_y \left\{ (t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \right\} \tag{3}$$

for $0 \leq s < t$.

Proof. Take z so that $u(y,s) = sL\left(\frac{y-z}{s}\right) + u_0(z)$ and use convexity of L to show

$$L\left(\frac{x-z}{t}\right) \leq \left(1 - \frac{s}{t}\right) L\left(\frac{x-y}{t-s}\right) + \frac{s}{t} L\left(\frac{y-z}{s}\right)$$

Then

$$\begin{aligned}&u(x,t) \leq tL\left(\tfrac{x-z}{t}\right) + u_0(z) \leq (t-s)\, L\left(\tfrac{x-y}{t-s}\right) + sL\left(\tfrac{y-z}{s}\right) + u_0(z)\\ &= (t-s)L\left(\tfrac{x-y}{t-s}\right) + u(y,s).\end{aligned}$$

On the other hand, pick w so that $u(x,t) = tL\left(\frac{x-w}{t}\right) + u_0(w)$ and take $y = \frac{s}{t}x + (1-\frac{s}{t})w$. Then

$$\frac{x-y}{t-s} = \frac{x-w}{t} = \frac{y-w}{s}$$

and consequently

$$\begin{aligned}&(t-s)L\left(\tfrac{x-y}{t-s}\right) + u(y,s) \leq (t-s)L\left(\tfrac{x-w}{t}\right) + sL\left(\tfrac{y-w}{s}\right) + u_0(w)\\ &= tL\left(\tfrac{x-w}{t}\right) + u_0(w) = u(x,t).\end{aligned}$$

Proposition 1. *If u_0 is Lipschitz, then the function $u(x,t)$ given in (2) is Lipschitz continuous and continuous up to $t = 0$.*

Proof. Fix $t > 0$. Pick x, x_1. Find y so that

$$u(x,t) = tL\left(\frac{x-y}{t}\right) + u_0(y)$$

Then

$$\begin{aligned}&u(x_1,t) - u(x,t) = \min_z \left\{tL\left(\tfrac{x_1-z}{t}\right) + u_0(z)\right\} - tL\left(\tfrac{x-y}{t}\right) - u_0(y)\\ &\leq u_0(x_1 - x + y) - u_0(y) \leq C|x - x_1|.\end{aligned}$$

Then we switch the roles of x_1 and x. For the continuity at $t = 0$, on one hand we have

$$u(x,t) \leq tL(0) + u_0(x)$$

directly from (2) by taking $y = x$. On the other hand, using $u_0(y) \geq u_0(x) - C|x-y|$ in the definition (2) we have

$$\begin{aligned}&u(x,t) \geq u_0(x) + \min_y \left\{-C|x-y| + tL\left(\tfrac{x-y}{t}\right)\right\}\\&= u_0(x) - t\max_w\{C|w| - L(w)\}\\&= u_0(x) - t\max_{|z|\leq C}\max_w\{z\cdot w - L(w)\}\\&= u_0(x) - t\max_{|z|\leq C} H(z)\end{aligned}$$

We also have $|u(x,t) - u(x,s)| \leq C|t-s|$ using the semigroup property and the bounds above.

Proposition 2. *Let u_0 be Lipschitz and take a point (x,t) where the function $u(x,t)$ defined by (2) is differentiable. (This happens a.e. by Rademacher's theorem and the result above). Then*

$$\partial_t u(x,t) + H(\nabla u(x,t)) = 0.$$

Proof. let $q \in \mathbb{R}^n$, $h > 0$. By the semigroup property

$$\begin{aligned}&u(x+hq, t+h) = \min_y \left\{hL\left(\tfrac{x+hq-y}{h}\right) + u(y,t)\right\}\\&\leq hL(q) + u(x,t),\end{aligned}$$

which implies

$$q\cdot\nabla u(x,t) + \partial_t u(x,t) \leq L(q)$$

This is valid for all q, so

$$\partial_t u(x,t) + H(\nabla u(x,t)) = \partial_t u(x,t) + \min_q\{q\cdot\nabla u(x,t) - L(q)\} \leq 0.$$

On the other hand, choose z so that $u(x,t) = tL\left(\frac{x-z}{t}\right) + u_0(z)$. Take $h > 0$, put $s = t-h$, $y = \frac{s}{t}x + (1-\frac{s}{t})z$. Then $\frac{x-z}{t} = \frac{y-z}{s}$ and so,

$$\begin{aligned}&u(x,t) - u(y,s) \geq tL\left(\tfrac{x-z}{t}\right) + u_0(z) - \left\{sL\left(\tfrac{y-z}{s}\right) + u_0(z)\right\}\\&= (t-s)L\left(\tfrac{x-z}{t}\right).\end{aligned}$$

This means

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

Letting $h \to 0$

$$\partial_t u(x,t) + \frac{x-z}{t}\cdot\nabla u(x,t) \geq L\left(\frac{x-z}{t}\right)$$

Then

$$\partial_t u(x,t) + H(\nabla u(x,t)) = \partial_t u(x,t) + \max_w \{q \cdot \nabla u(x,t) - L(q)\}$$
$$\geq \partial_t u(x,t) + \tfrac{x-z}{t} \nabla u(x,t) - L\left(\tfrac{x-z}{t}\right) \geq 0.$$

This concludes this verification. Now it turns out that an additional property holds, semi-concavity.

Lemma 2. *If there exists a constant such that*

$$u_0(x+z) + u_0(x-z) - 2u_0(x) \leq C|z|^2$$

holds for all x, z, then the solution (2) satisfies, with the same C

$$u(x+z,t) + u(x-z,t) - 2u(x,t) \leq C|z|^2$$

Proof. Let y so that $u(x,t) = tL\left(\frac{x-y}{t}\right) + u_0(y)$ and use this $y+z$ and $y-z$ in the definitions of $u(x+z,t)$, $u(x-z,t)$. This will cancel the L terms and give the result.

Semi-concavity of u implies that $u_\epsilon = \phi_\epsilon * u$ where ϕ_ϵ is a standard mollifier satisfies

$$\nabla\nabla u_\epsilon \leq C\mathbb{I}$$

It turns out that if H is strictly convex, then for $t > 0$ solutions become semiconcave, even if u_0 was not, with bounds that explode at $t = 0$.

Lemma 3. *If H is strictly convex*

$$H\left(\frac{p_1+p_2}{2}\right) \leq \frac{1}{2}H(p_1) + \frac{1}{2}H(p_2) - c|p_1 - p_2|^2$$

then the solution (2) satisfies

$$u(x+z,t) + u(x-z,t) - 2u(x,t) \leq \frac{C}{t}|z|^2.$$

Proof. Using definitions, it turns out that

$$\frac{1}{2}L(q_1) + \frac{1}{2}L(q_2) \leq L\left(\frac{q_1+q_2}{2}\right) + C|q-1-q_2|^2$$

Then, choosing y to have $u(x,t) = tL\left(\frac{x-y}{t}\right) + u_0(y)$ and using the same y for $u(x+z,t)$ and $u(x-z,t)$ the u_0 terms cancel out, and the result emerges.

Theorem 1. *Let $H \in C^2$, strictly convex, superlinear at infinity. Assume u_0 is Lipschitz. Then there is only one Lipschitz continuous function u satisfying*

$$u(x,0) = u_0(x)$$

$$\partial_t u + H(\nabla u(x,t)) = 0, \quad \text{a.e.}$$

and

$$u(x+z,t) + u(x-z,t) - 2u(x) \le C(1+\frac{1}{t})|z|^2$$

Proof. Taking the difference u of two solutions u_1, u_2 we arrive at

$$\partial_t u + v \cdot \nabla u = 0$$

where v is bounded, and given by

$$v(x,t) = \int_0^1 \nabla_p H(\lambda \nabla u_1(x,t) + (1-\lambda)u_2(x,t))d\lambda$$

We mollify the solutions u_1 u_2 and define

$$v_\epsilon(x,t) = \int_0^1 \nabla_p H(\lambda \nabla u_{1\epsilon}(x,t) + (1-\lambda)u_{2\epsilon}(x,t))d\lambda$$

where $u_{j_\epsilon} = u_j * \phi_\epsilon$. The divergence satisfies a one-sided bound

$$\text{div} v_\epsilon \le C(1+\frac{1}{t}).$$

This used the semiconcavity of both solutions. Now because both solutions are Lipschitz

$$|u(x,t)| \le ct, \quad |\nabla u| \le C.$$

We write

$$\partial_t u + v_\epsilon \cdot \nabla u = (v_\epsilon - v) \cdot \nabla u$$

We take a function $w = f(u)$ with f smooth and nonnegative. This still solves

$$\partial_t w + v_\epsilon \cdot \nabla w = (v_\epsilon - v) \cdot \nabla w$$

We integrate on a fastly shrinking ball

$$e(t) = \int_{|x-x_0| \le V(t_0-t)} w dx$$

with $V > \|b\|_{L^\infty}$. Then

$$\frac{de}{dt} \leq C(1+\frac{1}{t})e + \int_{|x-x_0|\leq V(t_0-t)} (v-v_\epsilon)\nabla w$$

We let $\epsilon \to 0$:

$$\frac{de}{dt} \leq C(1+\frac{1}{t})e$$

Now we select f to vanish for $|u| \leq c\epsilon$, positive otherwise. Then $e = 0$ for $t \leq \epsilon$ and from then on, by Gronwall $e(t) = 0$. It follows that $|u| \leq c\epsilon$, but ϵ was arbitrary.