The Hamilton Jacobi Equation: An example demonstrating that Hamilton's principle function need not be everywhere differentiable

Tibor Döme

May 29, 2019

We have proven in Theorem 2.5.3 that Hamilton's principle function u is differentiable a.e. in $\mathcal{D} \equiv \mathbb{R}^n \times (0, +\infty)$. Yet, we will see that it is easy to construct initial conditions such that u will not be differentiable on some subset $U \subseteq \mathcal{D}$ with Lebesgue measure $\mu(U) = 0$.

1 The fully nonlinear Hamilton Jacobi Equation (HJE)

The problem we are discussing is the following:

$$\begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R} \times (0, +\infty) \\ u = g & \text{in } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{1}$$

Here, $u : \mathbb{R} \times [0, +\infty) \to \mathbb{R}$ is the unknown, also called Hamilton's principle function, with the gradient $Du = D_x u = u_x$. The Lagrangian $L : \mathbb{R} \to \mathbb{R}$ is given by $L(q) = \frac{q^2}{2}$, representing a free particle of mass $m = 1$. Its Legendre transform $H = L^* : \mathbb{R} \to \mathbb{R}$, the Hamiltonian, reveals to have the same functional form, $H(p) = \frac{p^2}{2}$. The initial function $g : \mathbb{R} \to \mathbb{R}$ is chosen to ensure that u will satisfy the desired property of non-differentiability:

$$g(x) = \begin{cases} -x^2 & |x| < 1 \\ 1 - 2|x| & |x| \geq 1. \end{cases}$$

We will present two "distinct" ways of solving problem (1). In our first approach, we will apply the Hopf-Lax formula (Theorem 2.5.2). In our second approach, we will apply the method of characteristics (MoC). A priori, the problem (1) could have more than one solution in the class of Lipschitz continuous functions as Remark (3.1) shows. Yet, the Hopf-Lax Formula is essentially equivalent to Hamilton's equations, which are nothing but two characteristic equations. Hence, we certainly expect method 1 to yield the same result as method 2.

2 Via the Hopf-Lax Formula

In the special case when H is independent of x, as in our case, we know that the Hopf-Lax formula

$$u(x,t) = \inf_{y \in \mathbb{R}} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} = \inf_{y \in \mathbb{R}} \left\{ \frac{(x-y)^2}{2t} + g(y) \right\} \equiv \inf_{y \in \mathbb{R}} \left\{ f(x,y,t) \right\} \tag{2}$$

indeed solves problem (1) (Theorem 2.5.3), albeit only a.e. Above, we have chosen L to be coercive,

$$\lim_{|q| \to +\infty} \frac{L(q)}{|q|} = +\infty, \tag{3}$$

i.e. it grows superlinearly at infinity. As a consequence, if we were able to show that g is locally Lipschitz continuous, then the infimum were actually a minimum. What we do know, by looking at the graph 1 of g, is that g is concave. Now, we can apply the fact that concavity (also convexity of course) implies local Lipschitz continuity (at least in finite-dimensional spaces such as $\mathbb{R}$). In total, we found that in our case, Eq. (2) contains a minimum rather than an infimum. This is crucial since for each given (x, t), it allows us to calculate the differential of $f(x, y, t)$ with respect to y, set it to zero, and choose the solution y^* with the smallest corresponding $f(x, y^*, t)$, without bothering about the boundary regions of $\mathbb{R} \ni y$.

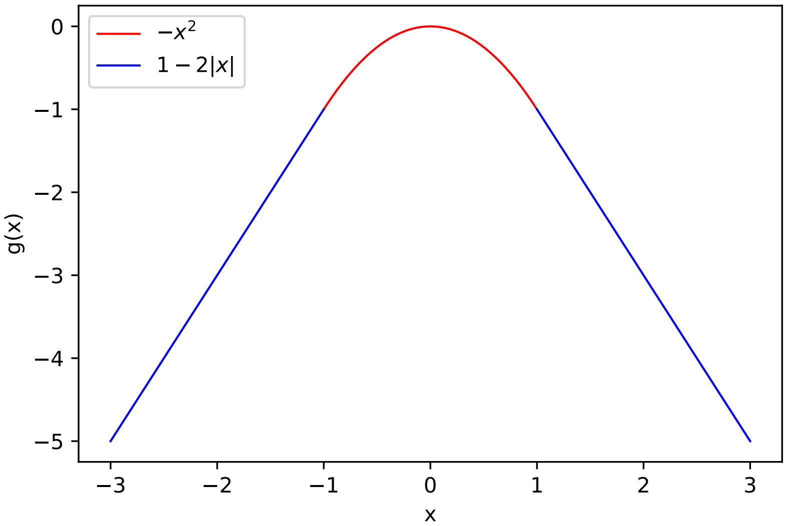

Figure 1: The initial function $g(x)$ is concave.

Our strategy is thus to case differentiate into

(a) $|y| < 1$

(b) $|y| \geq 1$,

and to perform the minimization in each section.

2.1 Case (a)

A trivial calculation yields

$$\frac{\partial}{\partial y} f(x, y, t) \overset{!}{=} 0 \quad \Leftrightarrow \quad y^{(a)} = \frac{x}{1-2t}. \tag{4}$$

Inserting back into $f(x, y, t)$ gives

$$u^{(a)}(x, t) = f(x, y, t)\Big|_{y=y^{(a)}} = -\frac{x^2}{1-2t}. \tag{5}$$

Now that we have found one part of the solution, $u^{(a)}(x, t)$, it might be good to know for which (x, t) it actually holds. For that, we recall that in case (a), $|y| < 1$, so in particular $|y^{(a)}| < 1$. This translates to the condition $t < \frac{1}{2} \ \wedge \ |x| < 1 - 2t$; this region is highlighted green ● in Fig. 2.

2.2 Case (b)

We repeat the minimization procedure and find

$$\frac{\partial}{\partial y} f(x, y, t) \overset{!}{=} 0 \quad \Leftrightarrow \quad |y^{(b)}| = \frac{2y^{(b)}t}{y^{(b)} - x}, \tag{6}$$

yielding

$$y^{(b)} = \begin{cases} 2t + x & \text{if } \; y^{(b)} \geq 1 \quad \color{red}{\bullet} \\ -2t + x & \text{if } \; y^{(b)} \leq -1 \quad \color{blue}{\bullet}, \end{cases}$$

where we have identified the corresponding regions in the (x,t)-plane of Fig. 2 by the colors red and blue. The intersection of the two regions is purple. For the region marked red, we find

$$u(x,t) = f(x,y,t)\Big|_{y=y^{(b)}} = 2t + 1 - 2|2t + x| = 2t + 1 - 2(|x| + 2t) = 1 - 2(|x| + t), \tag{7}$$

while for the region marked blue, we find

$$u(x,t) = f(x,y,t)\Big|_{y=y^{(b)}} = 2t + 1 - 2| - 2t + x| = 2t + 1 - 2(|x| + 2t) = 1 - 2(|x| + t). \tag{8}$$

For the purple region $\bullet$, we need to take the minimum of the two possibilities

$$u(x,t) = f(x,y,t)\Big|_{y=y^{(b)}} = \min_{\{+,-\}} \left\{ 2t + 1 - 2| \pm 2t + x| \right\} = 2t + 1 - 2(|x| + 2t) = 1 - 2(|x| + t), \tag{9}$$

which luckily turns out to have the same functional form. Although it is evident from Fig. 2, we still

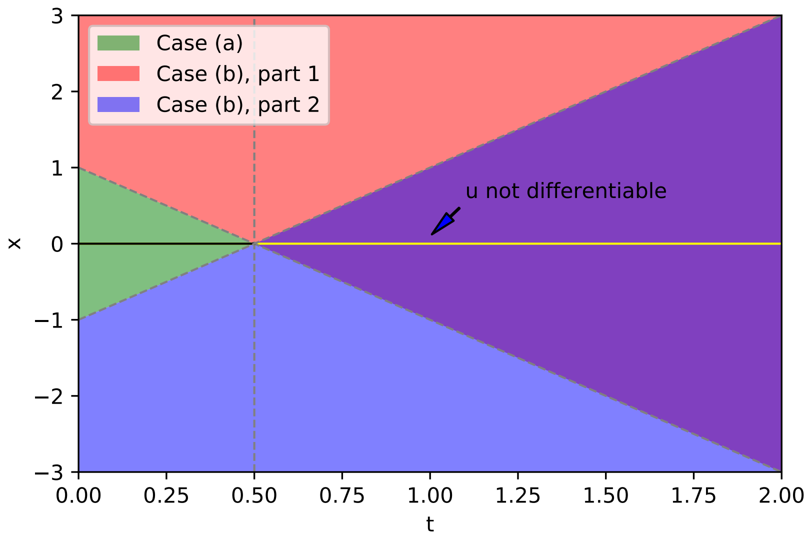

Figure 2: The (x,t)-plane is splitted into different regions, depending on the value of y. On the set $0 \times \mathbb{R}_{>\frac{1}{2}}$ of Lebesgue measure zero, $u(x,t)$ is not differentiable.

need to rewrite the condition $|y| \geq 1$ into a condition in the (x,t)-plane: $\left(t < \frac{1}{2} \wedge |x| \geq 1 - 2t\right) \vee t > \frac{1}{2}$.

2.3 The result for Hamilton's principle function

Hence our full result reads

$$u(x,t) = \begin{cases} -\frac{x^2}{1-2t} & \text{if } \; t < \frac{1}{2} \; \wedge \; |x| < 1 - 2t \\ 1 - 2(|x| + t) & \text{if } \; \left(t < \frac{1}{2} \; \wedge \; |x| \geq 1 - 2t\right) \; \vee \; t > \frac{1}{2}. \end{cases} \tag{10}$$

First, we observe that indeed, $u(x, t = 0) = g(x)$. Secondly, we see that $u(x,t)$ is not differentiably on the set $0 \times \mathbb{R}_{>\frac{1}{2}}$, owing to the absolute value function in case (b).

3 Via the MoC

For any given (fully nonlinear) PDE, one can try to apply the MoC, but its success is limited to some very special cases, including our problem (1). First, we rewrite (1) in an appropriate language:

$$\begin{cases} F(x,t,u,u_x,u_t) = \frac{u_x^2}{2} + u_t = 0 \\ u|_\Gamma = g \end{cases} \tag{11}$$

Defining $z \equiv u, p \equiv u_x, q \equiv u_t$, we identify our HJE in the traditional nomenclature as

$$F(x,t,z,p,q) = \frac{p^2}{2} + q = 0 \tag{12}$$

The system of characteristic equations turns out to be

$$\begin{cases} \dot{x}(s) = F_p = p \quad \text{(HE)} \\ \dot{t}(s) = F_q = 1 \\ \dot{z}(s) = pF_p + qF_q = p^2 + q = \frac{p^2}{2} \\ \dot{p}(s) = -F_x - pF_z = 0 \quad \text{(HE)} \\ \dot{q}(s) = -F_t - qF_z = 0, \end{cases} \tag{13}$$

where we used s as a characteristic curve parameter and labeled two equations by HE to indicate that those are the physically relevant Hamilton's equations. Identifying p as the momentum of the particle, the second HE represents a freely moving particle. We should emphasize that the last two equations in the system (13) are necessary only in the fully nonlinear case: We need the system to be determined, i.e. complete. As seen in (11), we have prescribed data for u on $\Gamma = \{(r,0)\}$. Therefore, as usual, our initial conditions for $x(r,s), y(r,s), z(r,s)$ are given by

$$x(r,0) \equiv \gamma_1(r) = r \tag{14}$$

$$t(r,0) \equiv \gamma_2(r) = 0 \tag{15}$$

$$z(r,0) \equiv \phi(r) = g(r). \tag{16}$$

In order to solve the system (13) of five ODEs, however, we need to prescribe initial conditions $p(r,0), q(r,0)$ as well. In so doing, we replace the initial curve $(\Gamma,\phi) = (\gamma_1(r),\gamma_2(r),\phi(r))$ with an *initial strip* $(\Gamma,\phi,\Psi) = (\gamma_1,\gamma_2,\phi,\psi_1,\psi_2)$. Initial values for p and q may not be prescribed arbitrarily, however. In particular, they must satisfy the following conditions. First, they must satisfy the PDE. Letting $\psi_1(r) = p(r,0), \psi_2(r) = q(r,0)$, we need

$$F(\gamma_1(r),\gamma_2(r),\phi(r),\psi_1(r),\psi_2(r)) = 0. \tag{17}$$

Second, using the chain rule,

$$\frac{du}{dr}(r) = \frac{du}{dx}x'(r) + \frac{du}{dt}t'(r) \tag{18}$$

suggests that we need ψ_1, ψ_2 to satisfy

$$\phi'(r) = \psi_1(r)\gamma_1'(r) + \psi_2(r)\gamma_2'(r). \tag{19}$$

Initial data (Γ,ϕ,Ψ) which satisfies Eq. (17) and Eq. (19) is said to be *admissible.*

Remark 3.1. *Uniqueness of Solutions: There may be no functions $\psi_1(r),\psi_2(r)$ which satisfy these equations, or there may not be unique functions. We will be able to invert the function $G(r,s) \equiv (x(r,s),y(r,s))$ as long as the noncharacteristic condition*

$$\begin{pmatrix} F_p(\gamma_1(r),\gamma_2(r),\phi(r),\psi_1(r),\psi_2(r)) \\ F_q(\gamma_1(r),\gamma_2(r),\phi(r),\psi_1(r),\psi_2(r)) \end{pmatrix} \cdot \begin{pmatrix} -\gamma_2'(r) \\ \gamma_1'(r) \end{pmatrix} \neq 0 \tag{20}$$

holds. Hence, if we are able to find functions $\psi_1(r), \psi_2(r)$ such that the initial data (Γ, ϕ, Ψ) is admissible and noncharacteristic, then, close to Γ at least, i.e. for small times, there exists a unique solution of the problem (11) with initial data $x(r,0) = \gamma_1(r), y(r,0) = \gamma_2(r), z(r,0) = \phi(r), p(r,0) = \psi_1(r)$ and $q(r,0) = \psi_2(r)$. This is essentially Lemma 2.4.2. In general of course, one cannot expect the existence and uniqueness of classical solutions $u(x,t)$ on the entire domain $\mathcal{A} = \mathbb{R} \times [0, +\infty)$, due to possible intersections of characteristic curves.

For our problem, Eq. (17) and Eq. (19) translate to

$$\begin{cases} \frac{\psi(r)^2}{2} + \psi_2(r) = 0 \\ \phi'(r) = \psi_1(r), \end{cases} \tag{21}$$

whose solution is

$$\psi_1(r) = \begin{cases} -2r & \text{if } |r| < 1 \\ -2\frac{r}{|r|} & \text{if } |r| \geq 1, \end{cases} \tag{22}$$

together with

$$\psi_2(r) = -\frac{\psi_1(r)^2}{2} = \begin{cases} -4r^2 & \text{if } |r| < 1 \\ -4 & \text{if } |r| \geq 1. \end{cases} \tag{23}$$

Further, the noncharacteristic condition is also satisfied:

$$\begin{pmatrix} \psi_1(r) \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \neq 0. \tag{24}$$

Even though the functions ψ_1 and ψ_2 are insignificant for the subsequent inversion of $G(r,s)$, we can rest assured that the MoC is guaranteed to yield a unique solution, for small times at least. In summary, we have the following system of IVPs

$$\begin{cases} \dot{x} = p & x(r,0) = r \\ \dot{t} = 1 & t(r,0) = r \\ \dot{z} = \frac{p^2}{2} & z(r,0) = g(r) \\ \dot{p} = 0 & p(r,0) = g'(r) \\ \dot{q} = 0 & q(r,0) = -\frac{g'(r)^2}{2}, \end{cases} \tag{25}$$

whose solution is

$$\begin{cases} x(r,s) = ps + r \\ t(r,s) = s \\ z(r,s) = s\frac{p^2}{2} + g(r) \\ p(r,s) = g'(r) \\ q(r,s) = -\frac{g'(r)^2}{2}. \end{cases} \tag{26}$$

In particular,

$$x(r,s) = \begin{cases} r(1-2t) & \text{if } |r| < 1 \\ r(1 - \frac{2t}{|r|} & \text{if } |r| \geq 1, \end{cases} \tag{27}$$

which can be inverted to give

$$r(x,t) = \begin{cases} \frac{x}{1-2t} & \text{if } |r| < 1 \\ x \pm 2t & \text{if } |r| \geq 1, \end{cases} \tag{28}$$

where the $\pm$ stands for the case $x > 0$ and $x \leq 0$, respectively. The last step consists of expressing z in terms of (x, t):

$$u(x,t) = z(r(x,t), s(x,t)) = t\frac{\psi_1(r(x,t))^2}{2} + g(r(x,t)) \tag{29}$$

$$= \begin{cases} -\frac{x^2}{1-2t} & \text{if } \ t < \frac{1}{2} \ \wedge \ |x| < 1-2t \\ 1-2(|x|+t) & \text{if } \ \left(t < \frac{1}{2} \ \wedge \ |x| \geq 1-2t\right) \ \vee \ t > \frac{1}{2}, \end{cases} \tag{30}$$

which is the same result as via the Hopf-Lax formula, as expected. One subtlety was swept under the carpet though: Right after Eq. (28), we claimed that the plus-solution corresponds to the case $x > 0$ and the minus-solution to the case $x \leq 0$. This was the only possible choice, among the mathematically admissible solutions of the inversion $x(r, s), t(r, s) \to r(x, t), s(x, t)$, that ensures continuity of $u(x, t)$ everywhere in $\mathcal{A}$. Other choices of $\pm$ in the (x, t)-plane yield non-continuity at the border line between the red (respectively blue) region and the purple region in Fig. 2. Since we are searching for $u(x, t)$ in the space of Lipschitz continuous functions, we do not care about non-continous solutions. Regarding the Hopf-Lax formula, however, this ambiguity did never even arise since we took the minimum of $f(x, y, t)$ over both possibilities, yielding the everywhere continuous solution, as expected. Still, how is it possible that the solution of the MoC was not unique, in the face of Remark (3.1)? The answer is that the characteristic curves intersect in the purple region of Fig. 2. In that case, statements on existence and uniqueness of solutions become much more involved. In general, however, such an occurrence signals the breakdown of uniqueness, just as we have seen in this example.

It might be interesting to note that Hamilton's principle function $u(x, t)$, that we have just found, has no physical meaning, in contrast to $x(t)$ and $p(t)$.

4 Conclusion

We have constructed initial conditions $g(x)$ that guarantee that the resulting principle function $u(x, t)$ is not differentiable on some subset $U \subset \mathcal{D}$ of Lebesgue measure zero. That principle function $u(x, t)$ could be found by the Hopf-Lax formula, since H did not depend on x. We were also able to apply the MoC to find a solution on the entire domain $\mathcal{A}$, though the solution is not unique due to the crossing of characteristic curves. Still, the Hopf-Lax solution was identical to the everywhere continuous MoC-solution, which was expected since the Hopf-Lax formula is equivalent to HEs, which are just two characteristic equations.