# Notes on Stochastic Control

Haosheng Zhou

Dec, 2022

**The following contents about HJE & HJB refers to the Evans book.**

# Hamilton-Jacobi Equation (HJE)

The Hamilton-Jacobi equation is a non-linear first-order PDE with the form

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

where $u = u(x,t) : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is the function to solve out and $Du = (u_{x_1}, ..., u_{x_n})$ is the gradient of $u$ w.r.t. space variable $x$. Here the **Hamiltonian** $H : \mathbb{R}^n \to \mathbb{R}$ is given and the initial condition $g : \mathbb{R}^n \to \mathbb{R}$ is given.

## Connection with Hamilton's Equations

Let's first apply the method of characteristics to get some intuition by noticing that this equation is a first-order equation. We know that the method of characteristics does not necessarily hold in general (since it requires the existence of $C^2$ solution), but this may tell us how to proceed. In this section, we assume that HJE looks like

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

where the Hamiltonian also depends on $x$.

Notice that here we **merge the time variable $t$ with the space variable $x$ and denote it as $x \in \mathbb{R}^{n+1}$, where $x^1, ..., x^n$ are components of $x$ and $x^{n+1}$ denotes the time**. Define

$$z(s) = u(x(s)) \tag{3}$$

as the version of $u$ along the characteristic curve and

$$p(s) = Du(x(s)) \in \mathbb{R}^{n+1} \tag{4}$$

as the version of $Du$ along the characteristic curve, note that here $p^1, ..., p^n$ are partial derivatives w.r.t. $x$-components and $p_{n+1}$ is the partial derivative w.r.t. $t$. One would always set the characteristic direction to be

$$x'(s) = D_p F \tag{5}$$

where the original PDE can be written as $F(Du, u, x) = 0$ and in this case

$$F(p, z, y) = p^{n+1} + H(p^1, ..., p^n, x^1, ..., x^n) \tag{6}$$

As a result, one get the ODE system from the method of characteristics

$$\begin{cases} [x^i(s)]' = H_{p_i}(p^1, ..., p^n, x^1, ..., x^n) \ (i = 1, 2, ..., n) \\ [x^{n+1}(s)]' = 1 \end{cases} \tag{7}$$

so one can identify $x^{n+1}(s)$ as $s$, meaning that the parameter $s$ is the same as the time variable $t = x^{n+1}$. The equation for $z(s)$ is $z'(s) = D_pF \cdot p(s)$, so

$$z'(s) = \sum_{i=1}^{n} H_{p_i}(p^1, ..., p^n, x^1, ..., x^n) \cdot p^i(s) + p^{n+1}(s) \tag{8}$$

$$= \sum_{i=1}^{n} H_{p_i}(p^1, ..., p^n, x^1, ..., x^n) \cdot p^i(s) - H(p^1, ..., p^n, x^1, ..., x^n) \tag{9}$$

The equation for $p(s)$ is $p'(s) = -D_xF - D_zF \cdot p(s)$, so

$$\begin{cases} [p^i(s)]' = -H_{x_i}(p^1, ..., p^n, x^1, ..., x^n) \ (i = 1, 2, ..., n) \\ [p^{n+1}(s)]' = 0 \end{cases} \tag{10}$$

with the last equation $[p^{n+1}(s)]' = 0$ as the redundant one since $x^{n+1}$ has already been parameterized as $s$.

By cancelling all redundant equations and reorganizing the variables, we get the **characteristic ODE system** for HJE

$$\begin{cases} x'(s) = D_pH(p(s), x(s)) \\ z'(s) = D_pH(p(s), x(s)) \cdot p(s) - H(p(s), x(s)) \\ p'(s) = -D_xH(p(s), x(s)) \end{cases} \tag{11}$$

where $p(s) = (p^1(s), ..., p^n(s))$ and $x(s) = (x^1(s), ..., x^n(s))$ (the last component in $x(s), p(s)$ is ignored). **The Hamilton's equation** is defined as the system consisting of the first and third equation, i.e.

$$\begin{cases} x'(s) = D_pH(p(s), x(s)) \\ p'(s) = -D_xH(p(s), x(s)) \end{cases} \tag{12}$$

**Remark.** *The reason that we only take the equations w.r.t $x(s)$ and $p(s)$ in the Hamilton's equations is that those two equations have nothing to do with $z$, they already have $2n$ unknowns and $2n$ equations. In other words, the equation w.r.t. $z(s)$ does not provide any effective information for the derivation of $x(s), p(s)$, and after solving out $x(s), p(s)$, one can immediately know $z(s)$.*

## A Problem in the Calculus of Variation

The connection between HJE and Hamilton's equations can also be shown in another perspective by considering a problem in the calculus of variation. The problem is formed as finding a best curve in an admissible class. The **admissible class** is defined as

$$\mathscr{A} = \left\{w \in C^2, w : [0,t] \to \mathbb{R}^n : w(0) = y, w(t) = x\right\} \tag{13}$$

so any admissible curve is a $C^2$ path in $\mathbb{R}^n$ such that it starts from point $y$ and ends at point $x$ with $x, y \in \mathbb{R}^n, t > 0$ given. Imagine $w(s) \in \mathscr{A}$ as the moving trajectory of a particle, then $w'(s)$ is actually the speed of the particle at each time. The **action functional** is then defined as

$$I[w] = \int_0^t L(w'(s), w(s))\, ds \tag{14}$$

where $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a given smooth function called **Lagrangian** and we hope to find a curve $x(s) \in \mathscr{A}$ such that the action functional is minimized

$$I[x] = \inf_{w(s) \in \mathscr{A}} I[w] \tag{15}$$

**Remark.** *The Lagrangian has the meaning as the kinetic energy minus the potential energy in physics which has the meaning of "increments of distance". Here among all possible and smooth enough curves between two fixed points, we want to find $x(s)$ such that it minimizes the integral of the Lagrangian along the path, equivalent to saying that the optimal path is the one that* ***takes the "shortest" path****. If one still finds it hard to understand, think about how light travels, it always travels in the path such that the distance it goes through is the shortest, a natural minimization of a "trivial" action functional.*

Let's **assume** that the Lagrangian is given by $L = L(v, x)$ $(v, x \in \mathbb{R}^n)$ for the convenience of notations and that **the inf of $I[w]$ can be achieved by some $x(s) \in \mathscr{A}$ as the optimal path**. To build up a PDE for $x(s)$, choose smooth $y : [0,t] \to \mathbb{R}^n$ with $y(s) = (y^1(s), ..., y^n(s))$ such that $y(0) = y(t) = 0$ and consider perturbing the optimal path $x(s)$ by a small multiple of $y(s)$ to get

$$w(s) = x(s) + \tau y(s) \ (\tau \in \mathbb{R}) \tag{16}$$

since $w(s) \in \mathscr{A}$, one immediately sees that

$$I[w] \geq I[x] \tag{17}$$

Consider the action functional of the perturbed path

$$i : \mathbb{R} \to \mathbb{R}, i(\tau) = I[x + \tau y] \tag{18}$$

it's easy to see that it has minimum at $\tau = 0$ (assume it's differentiable) with

$$i'(\tau) = \frac{d}{d\tau}\int_0^t L(x'(s) + \tau y'(s), x(s) + \tau y(s))\,ds \tag{19}$$

$$= \int_0^t y'(s)\cdot L_v(x'(s) + \tau y'(s), x(s) + \tau y(s)) + y(s)\cdot L_x(x'(s) + \tau y'(s), x(s) + \tau y(s))\,ds \tag{20}$$

so

$$i'(0) = \int_0^t y'(s)\cdot L_v(x'(s), x(s)) + y(s)\cdot L_x(x'(s), x(s))\,ds \tag{21}$$

$$= \int_0^t \sum_{i=1}^n \left([y^i(s)]'\cdot L_{v_i}(x'(s), x(s)) + y^i(s)\cdot L_{x_i}(x'(s), x(s))\right)\,ds \tag{22}$$

$$= 0 \tag{23}$$

Do transformations to this integral to find

$$\sum_{i=1}^n \int_0^t \left([y^i(s)]'\cdot L_{v_i}(x'(s), x(s)) + y^i(s)\cdot L_{x_i}(x'(s), x(s))\right)\,ds \tag{24}$$

$$= \sum_{i=1}^n \int_0^t L_{v_i}(x'(s), x(s))\,dy^i(s) + \int_0^t L_{x_i}(x'(s), x(s))\cdot y^i(s)\,ds \tag{25}$$

$$= \sum_{i=1}^n -\int_0^t y^i(s)\,dL_{v_i}(x'(s), x(s)) + \int_0^t L_{x_i}(x'(s), x(s))\cdot y^i(s)\,ds \tag{26}$$

$$= \sum_{i=1}^n \int_0^t \left[-\frac{d}{ds}L_{v_i}(x'(s), x(s)) + L_{x_i}(x'(s), x(s))\right] y^i(s)\,ds \tag{27}$$

$$= 0 \tag{28}$$

which is valid for any smooth $y$ such that $y(0) = y(t) = 0$. By density argument,

$$\forall i = 1, 2, ..., n, \forall s \in [0, t], -\frac{d}{ds}L_{v_i}(x'(s), x(s)) + L_{x_i}(x'(s), x(s)) = 0 \tag{29}$$

**Theorem 1.** ***(Euler-Lagrange Equation)*** *If path $x(s)$ is the optimal path and solves the variational problem mentioned above, then it must satisfy Euler-Lagrange equation that*

$$\forall s \in [0, t], -\frac{d}{ds}D_v L(x'(s), x(s)) + D_x L(x'(s), x(s)) = 0 \tag{30}$$

**Remark.** *The Euler-Lagrange equation consists of $n$* ***second-order ODEs****. Note that when $x(s)$ is the solution to the Euler-Lagrange equation, it does not necessarily achieve the inf of the action functional in the variational problem so* ***the converse of this theorem is not true****.*

In order to link Euler-Lagrange equation back to Hamilton's equations, let's first define

$$p(s) = D_v L(x'(s), x(s)) \tag{31}$$

as the **generalized momentum for position $x(s)$ and velocity $x'(s)$** (we will see why this has something to do with momentum later). We have to **assume that given $x, p \in \mathbb{R}^n$ the equation $p = D_v L(v, x)$ can be uniquely solved for $v$ as a smooth function of $p$ and $x$ as $v(p, x)$.** The **Hamiltonian $H$ associated with Lagrangian $L$** is defined as

$$H(p, x) = p \cdot v(p, x) - L(v(p, x), x) \ (p, x \in \mathbb{R}^n) \tag{32}$$

for $v(p, x)$ satisfying $p = D_v L(v, x)$ for given $x, p$ defined implicitly.

**Remark.** *To understand the motivation of those definitions, let's consider the classical setting in physics that*

$$L(v, x) = \frac{1}{2} m||v||_2^2 - \phi(x) \tag{33}$$

*where $\frac{1}{2}m||v||_2^2$ is the kinetic energy and $\phi$ is the potential energy with the mass $m > 0$. The Lagrangian immediately tells us that the Euler-Lagrange equation is*

$$-\frac{d}{ds} m x'(s) - D\phi(x(s)) = 0 \tag{34}$$

$$m \cdot x''(s) = -D\phi(x(s)) \tag{35}$$

*where $D\phi$ is the force field generated by $\phi$ and this is **Newton's second law** for the acceleration of a particle with mass $m$ in such force field.*

*Let's then try to calculate the generated momentum*

$$p(s) = m \cdot x'(s) \tag{36}$$

*which is consistent with the true momentum in this case. The implicit definition of $v$ is then*

$$p(s) = D_v L(v(s), x(s)) \tag{37}$$

$$m \cdot x'(s) = m \cdot v(s) \tag{38}$$

*since it can be uniquely solved for $v$ as a smooth function, it must be true that $v(p, x) = x'(s)$, just the velocity of the particle. As a result, the Hamiltonian for such Lagrangian is*

$$H(p, x) = m \cdot v \cdot v - L(v, x) \tag{39}$$

$$= \frac{1}{2} m||v||_2^2 + \phi(x) \tag{40}$$

*so the Hamiltonian is the **total energy** as the sum of kinetic and potential energy.*

**Theorem 2.** ***(Connection with Hamilton's Equation)*** *Let $x(s)$ be the optimal solution to the variational problem and $p(s)$ be its generalized momentum defined as $p(s) = D_v L(x'(s), x(s))$ above, then those two quantities satisfy Hamilton's equations*

$$\begin{cases} x'(s) = D_p H(p(s), x(s)) \\ p'(s) = -D_x H(p(s), x(s)) \end{cases} \tag{41}$$

*for $s \in [0, t]$ and the mapping*

$$s \to H(p(s), x(s)) \tag{42}$$

*is constant.*

*Proof.* Here is where the assumption that $p = D_v L(v, x)$ has unique smooth solution $v = v(p, x)$ comes in. By such assumption, we conclude that $v(p(s), x(s)) = x'(s)$.

After noticing this fact, we are left with pure calculations for $D_p H, D_x H$. By definition, $H(p, x) = p \cdot v(p, x) - L(v(p, x), x)$, so

$$\begin{aligned} H_{p_i}(p, x) &= \sum_{j=1, j\neq i}^{n} p_j \cdot v^j_{p_i}(p, x) + v^i(p, x) + p_i \cdot v^i_{p_i}(p, x) - D_v L(v(p, x), x) \cdot D_{p_i} v(p, x) && (43) \\ &= \sum_{j=1}^{n} [p_j \cdot v^j_{p_i}(p, x) - L_{v_j}(v(p, x), x) \cdot v^j_{p_i}(p, x)] + v^i(p, x) && (44) \\ &= \sum_{j=1}^{n} [p_j - L_{v_j}(v(p, x), x)] \cdot v^j_{p_i}(p, x) + v^i(p, x) && (45) \\ &= v^i(p, x) && (46) \end{aligned}$$

since $p = D_v L(v, x)$ by the definition of $v$. As a result,

$$H_{p_i}(p(s), x(s)) = v^i(p(s), x(s)) = [x_i(s)]' \tag{47}$$

proved how the first $n$ equations come.

For the next $n$ equations, the calculation is similar

$$H_{x_i}(p,x) = \sum_{j=1}^n p_j v^j_{x_i}(p,x) - D_v L(v(p,x),x)\cdot D_{x_i}v(p,x) - D_x L(v(p,x),x) \tag{48}$$
$$= \sum_{j=1}^n [p_j v^j_{x_i}(p,x) - p_j \cdot v^j_{x_i}(p,x)] - D_x L(v(p,x),x) \tag{49}$$
$$= -D_x L(v(p,x),x) \tag{50}$$

by applying the definition of $v$ once more. As a result,

$$H_{x_i}(p(s),x(s)) = -D_x L(v(p(s),x(s)),x(s)) = -D_x L(x'(s),x(s)) \tag{51}$$

proves the Hamilton's equations.

Moreover,

$$\frac{d}{ds}H(p(s),x(s)) = D_p H(p(s),x(s))\cdot p'(s) + D_x H(p(s),x(s))\cdot x'(s) \tag{52}$$
$$= x'(s)\cdot p'(s) - p'(s)\cdot x'(s) \tag{53}$$
$$= 0 \tag{54}$$

and this is telling us that the Hamiltonian is invariant w.r.t. time. □

**Remark.** *To briefly conclude what we have talked about in this section, we start from introducing Lagrangian as the "running loss function" of the variational problem and hope to find the optimal path $x(s)$ minimizing the loss. Such optimal path shall then satisfy the Euler-Lagrange equation consisting of $n$ second-order ODEs.*

*From the Euler-Lagrange equations, one can further introduce the generalized momentum $p(s)$ and the velocity $v(p,x)$ as the unique smooth solution to $p = D_v L(v,x)$ (such $v(s) = x'(s)$ is the unique velocity such that the generalized momentum is the given $p$). The Hamiltonian is defined and the optimal path $x(s)$ and the generalized momentum $p(s)$ must satisfy the Hamilton's equation. Moreover,* ***the Hamiltonian won't change as time goes by****.*

*As a result, we have interpreted the meaning of the Hamilton equations derived from the method of characteristics.*

## Legendre Transform & Frenchel Conjugate

Now let's turn back to HJE

$$\begin{cases} u_t + H(Du) = 0 & in\ \mathbb{R}^n \times (0,\infty) \\ u = g & on\ \mathbb{R}^n \times \{t=0\} \end{cases} \tag{55}$$

with **the dependence on $x$ of Hamiltonian $H$ cancelled**. Now the Lagrangian $L(v)$ only depends on $v$. Let's **assume that the Lagrangian is a convex function with** $\lim_{||v||\to\infty} \frac{L(v)}{||v||} = +\infty$ so of course it's continuous.

The **Legendre transform** provides the **Frenchel conjugate** of the Lagrangian as

$$L^*(p) = \sup_{v\in\mathbb{R}^n} \{p\cdot v - L(v)\} \tag{56}$$

The motivation of considering Frenchel conjugate comes from the fact that in previous discussions the Hamiltonian is defined as $H(p,x) = p\cdot v(p,x) - L(v(p,x),x)$, a form very similar to the conjugate of Lagrangian. To figure out the relationship between Hamiltonian and Lagrangian, notice that under the assumptions for Lagrangian, $p\cdot v - L(v)$ is concave and continuous in $v$. For each fixed $p\in\mathbb{R}^n$,

$$\frac{p\cdot v - L(v)}{||v||} = p\cdot\frac{v}{||v||} - \frac{L(v)}{||v||} \to -\infty\ (||v||\to\infty) \tag{57}$$

so there must **exist** $v^*\in\mathbb{R}^n$ **such that the sup can be attained**, i.e. $L^*(p) = p\cdot v^* - L(v^*)$. Note that **if $L$ is differentiable at** $v^*$, then

$$p - DL(v^*) = 0 \tag{58}$$

since $v^*$ achieves the sup. This gives us the equation $p = DL(v^*)$ which is just the definition equation for velocity $v(p)$ in the context above. As a result, $v(p) = v^*$ is the solution (although no uniqueness ensured). Replacing $v^*$ with the velocity $v(p)$ one can see

$$p\cdot v(p) - L(v(p)) = L^*(p) \tag{59}$$

and the LHS is an analogue to the definition of the Hamiltonian at $p$! Heuristically, this gives rise to the convex duality construction of Lagrangian and Hamiltonian.

**Theorem 3.** ***(Convex Duality of Lagrangian and Hamiltonian)*** *Assume that Lagrangian $L = L(v)$ is convex and $\lim_{||v||\to\infty}\frac{L(v)}{||v||} = +\infty$ and* ***define*** *Hamiltonian $H = L^*$, then $H$ is still convex, $\lim_{||p||\to\infty}\frac{H(p)}{||p||} = +\infty$ and $L = H^*$.*

*In particular, when $H$ is differentiable at $p$ and $L$ is differentiable at $v$, then the followings are equivalent:*

$$\begin{cases} p \cdot v = L(v) + H(p) \\ p = DL(v) \\ v = DH(p) \end{cases} \tag{60}$$

*Proof.* Note that $H = L^*$ so $H^* = L^{**}$. Note that since $L$ is a convex and closed function, its Frenchel conjugate must be itself (since double Frenchel conjugate gives the convex envelope), so $H^* = L$ is still convex and closed.

Notice that $H(p) = \sup_{v\in\mathbb{R}^n} \{p \cdot v - L(v)\}$, so

$$\forall \lambda > 0, H(p) \geq p \cdot \lambda \frac{p}{||p||} - L\left(\lambda \frac{p}{||p||}\right) \tag{61}$$

$$\geq \lambda ||p|| - \sup_{B(0,\lambda)} L \tag{62}$$

it's then obvious that $\lim_{||p||\to\infty} \frac{H(p)}{||p||} \geq \lambda$, so $\lim_{||p||\to\infty} \frac{H(p)}{||p||} = +\infty$.

When $H$ is differentiable at $p$ and $L$ is differentiable at $v$, note that if $p \cdot v = L(v) + H(p)$ then $v$ is achieving the sup in $H(p) = \sup_{v\in\mathbb{R}^n} \{p \cdot v - L(v)\}$ so

$$p - DL(v) = 0 \tag{63}$$

and $p$ is achieving the sup in $L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\}$ so

$$v - DH(p) = 0 \tag{64}$$

Conversely, if $p = DL(v)$, then it's true that $H(p) = p \cdot v - L(v)$ so it's proved. □

**Remark.** *Consider the previous example that*

$$L(v) = \frac{1}{2} m ||v||^2 \tag{65}$$

*then $H(p) = \sup_{v\in\mathbb{R}^n} \left\{p \cdot v - \frac{1}{2} m ||v||^2\right\}$ and the sup is achieved at $v^* = \frac{1}{m} p$*

$$H(p) = \frac{1}{2m} ||p||^2 \tag{66}$$

*if $p = DL(v) = mv$, then the Hamiltonian is actually*

$$H(p) = \frac{1}{2} m ||v||^2 \tag{67}$$

*which is equal to the Lagrangian when there's no potential and $H(p) + L(v) = p \cdot v$.*

**Remark.** *Let's compute some more examples to illustrate the connection between Lagrangian and Hamiltonian.*

*Consider* $H(p) = \frac{1}{r}||p||^r$ $(1 < r < \infty)$, *so*

$$L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\} \tag{68}$$

*and the sup is achieved when* $v = p \cdot ||p||^{r-2}$, *so* $v$ *is parallel to* $p$ *with* $p = kv$ $(k > 0)$. *Plug in to find*

$$L(v) = \sup_{k>0} \left\{k||v||^2 - \frac{k^r}{r}||v||^r\right\} \tag{69}$$

*and take another derivative w.r.t.* $k$ *to find that the sup is achieved when* $k = ||v||^{\frac{2-r}{r-1}}$, *so*

$$L(v) = \frac{r-1}{r}||v||^{\frac{r}{r-1}} \tag{70}$$

$$= \frac{1}{s}||v||^s \tag{71}$$

*where* $\frac{1}{s} + \frac{1}{r} = 1$, *so* $s$ *is the Holder conjugate of* $r$.

*Consider* $H(p) = \frac{1}{2}p^T Ap + b \cdot p$, *where* $A$ *is symmetric, positive definite and* $b \in \mathbb{R}^n$, *then*

$$L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\} \tag{72}$$

*and the sup is achieved when* $p = A^{-1}(v - b)$, *so*

$$L(v) = \frac{1}{2}(v - b)^T A^{-1}(v - b) \tag{73}$$

**Remark.** *For convex function, one can define the subdifferential of* $H$ *at* $p$ *so that the Frenchel inequality holds*

$$H(p) + L(v) \geq p \cdot v \tag{74}$$

*and the equality holds if and only if* $v \in \partial H(p)$ *if and only if* $p \in \partial L(v)$, *a generalization of the conclusion in the theorem above.*

## Hopf-Lax Formula

We still consider the HJE with Hamiltonian $H$ not depend on $x$ but only depends on $Du$. The characteristic ODEs then become

$$\begin{cases} p'(s) = 0 \\ z'(s) = DH(p(s)) \cdot p(s) - H(p(s)) \\ x'(s) = DH(p(s)) \end{cases} \tag{75}$$

with the equation for $p'(s), x'(s)$ being Hamilton's equations. Note that since $x'(s) = DH(p(s))$, by the theorem we have proved above, $L(x'(s)) + H(p(s)) = p \cdot x'(s)$. So **the equation of $z'(s)$ is describing the fact that** $z'(s) = L(x'(s))$. From the method of characteristics,

$$z(t) = u(x(t), t) = \int_0^t L(x'(s))\, ds + g(x(0)) \tag{76}$$

since $z(0) = u(x(0), 0) = g(x(0))$ by the initial value condition, providing us an ansatz of the solution. However, this construction of the solution $u(x, t)$ assumes the smoothness of the solution, which is often not the case for HJE. To think about modifying the construction of the solution such that it also works for non-smooth solution $u(x, t)$, we notice that

$$\int_0^t L(x'(s))\, ds \tag{77}$$

is the "running loss function" of the variational problem we have mentioned above and $x(s)$ is the optimal path found in that problem. As a result, we can think about **defining**

$$u(x, t) \overset{def}{=} \inf_w \left\{ \int_0^t L(w'(s))\, ds + g(w(0)) : w : [0, t] \to \mathbb{R}^n, w \in C^1, w(t) = x \right\} \tag{78}$$

**as the optimal "loss" determined by the Lagrangian among all paths that hits $x$ at time** $t$. To see how this works as the solution to the HJE, refer to the following theorem. We **assume that $H$ is smooth and convex with $\lim_{||p|| \to \infty} \frac{H(p)}{||p||} = +\infty$ and $g : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz in the following context**.

**Theorem 4.** ***(Hopf-Lax Formula)*** *For fixed $x \in \mathbb{R}^n, t > 0$,*

$$u(x, t) = \inf_w \left\{ \int_0^t L(w'(s))\, ds + g(w(0)) : w : [0, t] \to \mathbb{R}^n, w \in C^1, w(t) = x \right\} \tag{79}$$

$$= \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x - y}{t}\right) + g(y) \right\} \tag{80}$$

*Proof.* Consider $\forall y \in \mathbb{R}^n$ and the path $w(s) = y + \frac{s}{t}(x - y)$ so $w(t) = x$ (constructed based on $\frac{x-y}{t}$ inside the Lagrangian), it's obvious that

$$u(x, t) \leq \int_0^t L\left(\frac{x - y}{t}\right) ds + g(y) \tag{81}$$

so by taking inf w.r.t. $y$ on both sides

$$u(x, t) \leq \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x - y}{t}\right) + g(y) \right\} \tag{82}$$

Conversely, for any $C^1$ path $w$ such that $w(t) = x$, take $y = w(0)$

$$
\begin{aligned}
tL\left(\frac{x-y}{t}\right) + g(y) &= tL\left(\frac{x-w(0)}{t}\right) + g(w(0)) && (83)\\
&= tL\left(\frac{1}{t}\int_0^t w(s)\, ds\right) + g(w(0)) && (84)\\
&\leq \int_0^t L(w'(s))\, ds + g(w(0)) && (85)
\end{aligned}
$$

because of Jensen's inequality applied for $\frac{1}{t}\int_0^t f(s)\, ds$, the integral average of $f$ on $[0,t]$

$$
\frac{1}{t}\int_0^t L\left(w'(s)\right)\, ds \geq L\left(\frac{1}{t}\int_0^t w'(s)\, ds\right) \tag{86}
$$

by taking inf w.r.t. all paths $w$ on both sides, one can conclude that

$$
u(x,t) \geq \inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + g(y)\right\} \tag{87}
$$

so the theorem is proved. □

**Remark.** *The shifting from the ansatz $u(x(t),t) = \int_0^t L(x'(s))\, ds + g(x(0))$ to Hopf-Lax formula is critical! The main thought comes from the variational problem viewing $x(s)$ as the optimal path and the integral of Lagrangian as running loss.*

*Actually, from another perspective, one may be able to see the spirit of stochastic control out of the Hopf-Lax formula. Notice that $w$ can be view as a stochastic process instead of a deterministic function, and the $\int_0^t L(w'(s))\, ds$ can be viewed as a running loss with $g(w(0))$ as terminal loss (conditional on the filtration $\mathscr{F}_t$, i.e. all information available until time $t$, that's why the domain of $w$ is $[0,t]$. Then $u(x,t)$ is essentially a value function conditioning on $w(t) = x$, i.e. the process passes through $x$ at time $t$. In such sense,* ***HJE is actually characterizing the value function of a stochastic control problem in a deterministic way!***

**Remark.** ***In Hopf-Lax formula, the inf can always be attained****. Note that $f(y) = tL\left(\frac{x-y}{t}\right) + g(y)$ is continuous in $y$ and*

$$
\frac{f(y)}{||y||} = \frac{L\left(\frac{x-y}{t}\right)}{\frac{||y||}{t}} + \frac{g(y)}{||y||} \tag{88}
$$

*with $L = H^*$ so $\lim_{||v||\to\infty}\frac{L(v)}{||v||} = +\infty$ and since $g$ is Lipschitz, $\frac{g(y)}{||y||} \leq \frac{g(0)+Lips(g)||y||}{||y||} \leq Lips(g) + \varepsilon$ for large enough $||y||$. As a result,*

$$
\frac{f(y)}{||y||} \to +\infty \ (||y|| \to \infty) \tag{89}
$$

*combining with continuity, we see that the minimum of $f(y)$ must be attained by some $y \in \mathbb{R}^n$.*

## Hopf-Lax Formula as Solution to HJE

Now let's argue that the heuristic definition of such $u(x,t)$ by the Hopf-Lax formula is actually a solution to HJE. In order to prove this, let's first consider some useful propositions.

**Theorem 5.** ***(Flow Property)*** *For each $x \in \mathbb{R}^n$ and $s \in [0,t]$,*

$$u(x,t) = \inf_{y \in \mathbb{R}^n} \left\{ (t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \right\} \tag{90}$$

*Proof.* Let's start by noticing that for $\forall y \in \mathbb{R}^n, s \in [0,t]$, there always exists $z \in \mathbb{R}^n$ such that the inf in Hopf-Lax formula is attained, i.e.

$$u(y,s) = sL\left(\frac{y-z}{s}\right) + g(z) \tag{91}$$

in order to connect it with $\frac{x-y}{t-s}$, consider the convex representation and apply the convexity of $L$ that

$$\frac{t-s}{t}\frac{x-y}{t-s} + \frac{s}{t}\frac{y-z}{s} = \frac{x-z}{t} \tag{92}$$

$$\frac{t-s}{t}L\left(\frac{x-y}{t-s}\right) + \frac{s}{t}L\left(\frac{y-z}{s}\right) \geq L\left(\frac{x-z}{t}\right) \tag{93}$$

so that

$$(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) = (t-s)L\left(\frac{x-y}{t-s}\right) + sL\left(\frac{y-z}{s}\right) + g(z) \geq tL\left(\frac{x-z}{t}\right) + g(z) \tag{94}$$

take inf w.r.t. $y$ on both sides, one would see that

$$\inf_{y \in \mathbb{R}^n} \left\{ (t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \right\} \geq tL\left(\frac{x-z}{t}\right) + g(z) \geq u(x,t) \tag{95}$$

On the other hand, let's try to find $y \in \mathbb{R}^n$ such that $(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \leq u(x,t)$. Apply the Hopf-Lax formula again to find $w \in \mathbb{R}^n$ such that $u(x,t) = tL\left(\frac{x-w}{t}\right) + g(w)$. Consider applying the convexity of $L$ again, to set

$$y = \frac{s}{t}x + \frac{t-s}{t}w \tag{96}$$

$$\frac{x-y}{t-s} = \frac{x-w}{t} \tag{97}$$

and apply Hopf-Lax formula for $u(y,s)$ once more to find

$$(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \leq (t-s)L\left(\frac{x-w}{t}\right) + u(y,s) \leq (t-s)L\left(\frac{x-w}{t}\right) + sL\left(\frac{y-w}{s}\right) + g(w) \tag{98}$$

note that $\frac{y-w}{s} = \frac{x-w}{t}$, so

$$(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \leq tL\left(\frac{x-w}{t}\right) + g(w) = u(x,t) \tag{99}$$

by taking inf w.r.t. $y$ on both sides, we proved the conclusion. □

**Remark.** *Note that* ***the inf in this theorem can always be attained****. This requires proving the fact that $y \to u(y,s)$ is continuous, which will be proved in a later context.*

**Remark.** *The reason why we are calling this property the flow property is that this is telling us that we can act as if we are starting at time $s < t$ with initial value $u(y,s)$. Then the Hopf-Lax formula still holds for such problem and will generate the same $u$ as what we would derive with an initial value condition at time 0. This is actually very similar to the flow property of diffusion process.*

Under the assumption that $g$ is Lipschitz, one would see that such $u$ is also Lipschitz in $\mathbb{R}^n \times [0,\infty)$ and it agrees with the initial value condition $g$, i.e. $\forall x \in \mathbb{R}^n, u(x,0) = g(x)$.

**Theorem 6.** ***(Lipschitz Continuity)*** *Such $u$ is Lipschitz in $\mathbb{R}^n \times [0,\infty)$, and $\forall x \in \mathbb{R}^n, u(x,0) = g(x)$.*

*Proof.* First prove that $u(x,t)$ is Lipschitz in $x$. By Hopf-Lax formula, there exists $y \in \mathbb{R}^n$ such that $u(x,t) = tL\left(\frac{x-y}{t}\right) + g(y)$. As a result, for $\forall x, x' \in \mathbb{R}^n$,

$$u(x',t) - u(x,t) = \inf_z \left\{tL\left(\frac{x'-z}{t}\right) + g(z)\right\} - tL\left(\frac{x-y}{t}\right) - g(y) \tag{100}$$

$$\leq tL\left(\frac{x'-(x'-x+y)}{t}\right) + g(x'-x+y) - tL\left(\frac{x-y}{t}\right) - g(y) \tag{101}$$

$$= g(x'-x+y) - g(y) \leq Lips(g) \cdot ||x'-x|| \tag{102}$$

so

$$|u(x',t) - u(x,t)| \leq Lips(g) \cdot ||x'-x|| \tag{103}$$

by interchanging $x$ and $x'$.

Now let's prove that $u$ and $g$ agree when $t = 0$. Note that by Hopf-Lax formula, $u(x,t) \leq tL(0) + g(x)$. Set $t = 0$ to find $u(x,0) \leq g(x)$. For the other direction, we would need to use the conjugacy of Lagrangian and Hamiltonian.

$$u(x,t) = \inf_{y \in \mathbb{R}^n} \left\{tL\left(\frac{x-y}{t}\right) + g(y)\right\} \tag{104}$$

$$= g(x) + \inf_{y \in \mathbb{R}^n} \left\{tL\left(\frac{x-y}{t}\right) + g(y) - g(x)\right\} \tag{105}$$

$$\geq g(x) - t \sup_{y \in \mathbb{R}^n} \left\{-L\left(\frac{x-y}{t}\right) + Lips(g) \cdot \frac{||y-x||}{t}\right\} \tag{106}$$

by setting $z = \frac{x-y}{t}$ as a new variable, one can see the structure of this sup

$$u(x,t) \geq g(x) - t \sup_{z\in\mathbb{R}^n} \left\{-L(z) + Lips(g)\cdot ||z||\right\} \tag{107}$$

in order to connect this sup with the Frenchel conjugate of Lagrangian which is the Hamiltonian, we would like to see the forms like $w\cdot z - L(z)$. That's why we view $Lips(g)\cdot ||z||$ as $Lips(g)\frac{z}{||z||}\cdot z$ with $w = Lips(g)\frac{z}{||z||}$

$$u(x,t) \geq g(x) - t \sup_{w\in B(0,Lips(g))} \sup_{z\in\mathbb{R}^n} \left\{-L(z) + w\cdot z\right\} \tag{108}$$

$$= g(x) - t \sup_{w\in B(0,Lips(g))} H(w) \tag{109}$$

and since $H$ is continuous and convex, $\sup_{w\in B(0,Lips(g))} H(w) < \infty$, setting $t = 0$ to see

$$u(x,0) \geq g(x) \tag{110}$$

and we conclude that such $u$ is equal to $g$ when $t = 0$.

At last, prove that $u(x,t)$ is Lipschitz in $t$. For $\forall 0 < t < t'$, by the flow property,

$$u(x,t') - u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{(t'-t)L\left(\frac{x-y}{t'-t}\right) + u(y,t)\right\} - u(x,t) \tag{111}$$

$$\leq (t'-t)L(0) + u(x,t) - u(x,t) \tag{112}$$

$$= (t'-t)\cdot L(0) \tag{113}$$

on the other hand, let's apply the trick above one more time

$$u(x,t') = u(x,t) + \inf_{y\in\mathbb{R}^n} \left\{(t'-t)L\left(\frac{x-y}{t'-t}\right) + u(y,t) - u(x,t)\right\} \tag{114}$$

$$\geq u(x,t) + (t'-t) \inf_{y\in\mathbb{R}^n} \left\{L\left(\frac{x-y}{t'-t}\right) - Lips(u)\cdot \frac{||y-x||}{t'-t}\right\} \tag{115}$$

consider $z = \frac{y-x}{t'-t}$ and transform $Lips(u)\cdot\frac{||y-x||}{t'-t}$ into the inner product form to see

$$u(x,t') - u(x,t) \geq -(t'-t) \sup_{z\in\mathbb{R}^n} \left\{-L(z) + Lips(u)\cdot ||z||\right\} \tag{116}$$

$$= -(t'-t) \sup_{w\in B(0,Lips(u))} \sup_{z\in\mathbb{R}^n} \left\{-L(z) + w\cdot z\right\} \tag{117}$$

$$= -(t'-t) \sup_{w\in B(0,Lips(u))} H(w) \tag{118}$$

in all, we see that

$$|u(x,t') - u(x,t)| \leq C \cdot |t'-t|, C = \max\left\{|L(0)|, \sup_{w\in B(0,Lips(u))} |H(w)|\right\} \tag{119}$$

and such constant $C$ has no dependence on $x$ and $t$, that's why $u$ is also Lipschitz w.r.t. time $t$.

□

**Theorem 7.** ***(Hopf-Lax Formula as Solution to HJE)*** *For $u$ defined by the Hopf-Lax formula, if it's differentiable at a point $(x,t)$, then $u_t(x,t) + H(Du(x,t)) = 0$. In particular, such $u$ is differentiable almost everywhere and it's the solution to HJE in the almost everywhere sense.*

*Proof.* By Rademacher's theorem, Lipschitz function on an open subset of $\mathbb{R}^n$ is almost everywhere differentiable. So we only have to prove that HJE holds whenever $u$ is differentiable at $(x,t)$.

let's first calculate the directional derivative of $u$ along any vector $v$. By flow property,

$$u(x+hv, t+h) = \inf_{y\in\mathbb{R}^n}\left\{hL\left(\frac{x+hv-y}{h}\right) + u(y,t)\right\} \tag{120}$$

$$\leq hL(v) + u(x,t) \tag{121}$$

as a result,

$$\forall v \in \mathbb{R}^n, v\cdot Du(x,t) + u_t(x,t) = \lim_{h\to 0^+} \frac{u(x+hv,t+h) - u(x,t)}{h} \leq L(v) \tag{122}$$

note that the Hamiltonian is the Frenchel conjugate of Lagrangian, so

$$u_t(x,t) + H(Du(x,t)) = u_t(x,t) + \sup_{v\in\mathbb{R}^n}\{v\cdot Du(x,t) - L(v)\} \leq 0 \tag{123}$$

To prove the other side, we have to choose $v$ in the sup carefully. By Hopf-Lax formula, there exists $z \in \mathbb{R}^n$ such that $u(x,t) = tL\left(\frac{x-z}{t}\right) + g(z)$. Take $v = \frac{x-z}{t}$ in the sup to find

$$u_t(x,t) + H(Du(x,t)) \geq u_t(x,t) + \frac{x-z}{t}\cdot Du(x,t) - L\left(\frac{x-z}{t}\right) \tag{124}$$

again we have to use finite difference to approximate the partial derivatives

$$u(x,t) - u\left(\frac{t-h}{t}x + \frac{h}{t}z, t-h\right) = tL\left(\frac{x-z}{t}\right) + g(z) - u\left(\frac{t-h}{t}x + \frac{h}{t}z, t-h\right) \tag{125}$$

$$\geq tL\left(\frac{x-z}{t}\right) + g(z) - (t-h)L\left(\frac{x-z}{t}\right) - g(z) \tag{126}$$

$$= hL\left(\frac{x-z}{t}\right) \tag{127}$$

setting $h \to 0^+$ to know

$$u_t(x,t) + \frac{x-z}{t} \cdot Du(x,t) \geq L\left(\frac{x-z}{t}\right) \tag{128}$$

Finally, we have proved that

$$u_t(x,t) + H(Du(x,t)) = 0 \tag{129}$$

□

The theorem above ends our discussion on the solution to **a particular kind of HJE (Hamiltonian only depends on $Du$ and is convex with Lipschitz initial value condition)**. To see a direct example of the application of Hopf-Lax formula, consider the following PDE

$$\begin{cases} u_t + ||Du||^2 = 0 & in\ \mathbb{R}^n \times (0,\infty) \\ u = +\infty \cdot \mathbb{I}_{E^c} & on\ \mathbb{R}^n \times \{0\} \end{cases} \tag{130}$$

with $E$ as a closed subset in $\mathbb{R}^n$. Now the Hamiltonian is $H(p) = ||p||^2$ so Lagrangian is its Frenchel conjugate

$$L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\} = \frac{1}{4}||v||^2 \tag{131}$$

Apply the Hopf-Lax formula to find

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \tag{132}$$

$$= \inf_{y\in E} \left\{ \frac{1}{4t}||x-y||^2 \right\} \tag{133}$$

$$= \frac{1}{4t} dist^2(x,E) \tag{134}$$

the solution has something to do with the distance between $x$ and $E$.

# Optimal Control Problem

In this section we state the deterministic optimal control problem and find the connection between optimal control problem, HJE, Hamilton-Jacobi-Bellman equation (HJBE) and -Lax formula.

## Problem Formulation

All control problems have a certain dynamics telling us how the system evolves. In optimal control problem, the dynamics is given by an ODE

$$\begin{cases} x'(s) = f(x(s), \alpha(s)) \ (s \in [t, T]) \\ x(t) = x \end{cases} \tag{135}$$

where the dynamics works in time interval $[t, T]$ with $T$ fixed and an initial value condition given at time $t$. We will be varying the time $t$ and the initial value $x$ shortly afterwards to get a PDE describing such an optimal control problem. **Note that $x(s)$ denotes the state of the problem at time $s$ while $x$ denotes the initial value condition.** Viewing $x'(s)$ as $\frac{x(s+h)-x(s)}{h}$ for $h \to 0^+$, the ODE is describing how the change of state from time $s$ to time $s+h$ happens given the current state $x(s)$ and the current **control** $\alpha(s)$. (so it's actually a **Markovian** setting since $x'(s)$ has nothing to do with $\{x(t)\}|_{t<s}$ given $x(s)$.) The control can be understood as the "action" in discrete-time Markov decision process that changes the state evolution and has something to do with the rewards.

The control is nothing complicated but a set of parameters given at each time that will change the dynamics of the system, eventually changing the state evolution of the system. Let's denote $A \subset \mathbb{R}^m$ as some given compact set consisting of all possible values the control **at a given time** $\alpha(s)$ can take. The **admissible set**

$$\mathscr{A} = \{\alpha : [0, T] \to A : \alpha(\cdot) \ measurable\} \tag{136}$$

then denotes all possible controls across the whole time interval $[0, T]$ (since control may change over time, it maps each time point to the value of control at that time point). It's then clear that the function

$$f : \mathbb{R}^n \times A \to \mathbb{R}^n \tag{137}$$

is mapping a $m+n$-dimensional vector to a $n$-dimensional vector. Let's **assume that $f$ is a given bounded Lipschitz function**. This assumption is made to ensure that the ODE always has unique solution for every given control $\alpha(\cdot) \in \mathscr{A}$ denoted $x(\cdot) = x^{\alpha(\cdot)}(\cdot)$. **Our goal in optimal control problem is to find the optimal control $\alpha^*(\cdot)$ under some criteria**.

In order to define the optimality, we introduce the **cost functional** that represents the cost one has to pay selecting control $\alpha$ with initial value condition $x(t) = x$

$$C_{x,t}[\alpha] = \int_t^T r(x(s), \alpha(s))\, ds + g(x(T)) \tag{138}$$

here $\int_t^T r(x(s),\alpha(s))\,ds$ is the running cost and $g(x(T))$ is the terminal cost where $r:\mathbb{R}^n\times A\to\mathbb{R}, g:\mathbb{R}^n\to\mathbb{R}$ are **assumed to be bounded and Lipschitz in variable** $x$.

To sum up, given time $t\in[0,T]$ and the initial value condition $x(t)=x$, we want to find **the optimal control** $\alpha^*$ such that

$$C_{x,t}[\alpha^*]=\inf_{\alpha\in\mathscr{A}}C_{x,t}[\alpha] \tag{139}$$

## Value Function

Let's consider the **value function** $u(x,t)$ as the least possible cost among all admissible control with initial value condition $x(t)=x$ (with dynamic programming approach), i.e.

$$u(x,t)=\inf_{\alpha\in\mathscr{A}}C_{x,t}[\alpha] \tag{140}$$

then we hope to find a PDE that characterizes such value function $u$.

**Theorem 8.** ***(Optimality Condition)*** *For fixed $x\in\mathbb{R}^n, 0\le t<T$ and $h>0$ such that $t+h\le T$,*

$$u(x,t)=\inf_{\alpha\in\mathscr{A}}\left\{\int_t^{t+h}r(x(s),\alpha(s))\,ds+u(x(t+h),t+h)\right\} \tag{141}$$

*where $x(\cdot)=x^{\alpha(\cdot)}(\cdot)$ is the solution to the ODE for fixed control $\alpha(\cdot)$.*

*Proof.* For any control $\alpha_1\in\mathscr{A}$, the ODE has solution $x_1(\cdot)$. Now we want to prove that LHS is less than RHS, so we have to argue that $\forall\varepsilon>0$,

$$u(x,t)\le\int_t^{t+h}r(x_1(s),\alpha_1(s))\,ds+u(x_1(t+h),t+h)+\varepsilon \tag{142}$$

In order to achieve this goal, expand the inf in the definition of $u(x,t)$ for time $t+h$ and initial value $x_1(t+h)$ to find that $\forall\varepsilon>0$, there exists $\alpha_2\in\mathscr{A}$ and the solution to the ODE for fixed control $\alpha_2$ which is $x_2(\cdot)$ such that

$$u(x_1(t+h),t+h)+\varepsilon\ge C_{x_1(t+h),t+h}[\alpha_2]=\int_{t+h}^T r(x_2(s),\alpha_2(s))\,ds+g(x_2(T)) \tag{143}$$

so far, we have successfully figured out a lower bound for $u(x_1(t+h),t+h)$. To connect it with $u(x,t)$ and any control $\alpha_1$, we can construct a new control $\alpha_3$ that sticks to $\alpha_1$ before time $t+h$ but shifts to $\alpha_2$ after time $t+h$.

$$\alpha_3(s)=\alpha_1(s)\cdot\mathbb{I}_{t\le s\le t+h}+\alpha_2(s)\cdot\mathbb{I}_{t+h\le s\le T} \tag{144}$$

under our assumption, the original ODE has unique solution, and it's easy to see that

$$x_3(s)=x_1(s)\cdot\mathbb{I}_{t\le s\le t+h}+x_2(s)\cdot\mathbb{I}_{t+h\le s\le T} \tag{145}$$

is the solution to the ODE for fixed control $\alpha_3$ since

$$\forall t \le s \le t+h, x_3'(s) = x_1'(s) = f(x_1(s), \alpha_1(s)) = f(x_3(s), \alpha_3(s)) \tag{146}$$

$$\forall t+h \le s \le T, x_3'(s) = x_2'(s) = f(x_2(s), \alpha_2(s)) = f(x_3(s), \alpha_3(s)) \tag{147}$$

$$x_3(t) = x_1(t) = x \tag{148}$$

now we can see that

$$u(x,t) \le C_{x,t}[\alpha_3] \tag{149}$$

$$= \int_t^T r(x_3(s), \alpha_3(s))\, ds + g(x_3(T)) \tag{150}$$

$$= \int_t^{t+h} r(x_1(s), \alpha_1(s))\, ds + \int_{t+h}^T r(x_2(s), \alpha_2(s))\, ds + g(x_2(T)) \tag{151}$$

$$\le \int_t^{t+h} r(x_1(s), \alpha_1(s))\, ds + u(x_1(t+h), t+h) + \varepsilon \tag{152}$$

so we have proved that

$$u(x,t) \le \inf_{\alpha \in \mathscr{A}} \left\{ \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) \right\} \tag{153}$$

On the other hand, $\forall \varepsilon > 0$, there exists control $\alpha_4 \in \mathscr{A}$ such that

$$u(x,t) + \varepsilon \ge C_{x,t}[\alpha_4] \tag{154}$$

$$= \int_t^T r(x_4(s), \alpha_4(s))\, ds + g(x_4(T)) \tag{155}$$

$$= \int_t^{t+h} r(x_4(s), \alpha_4(s))\, ds + \int_{t+h}^T r(x_4(s), \alpha_4(s))\, ds + g(x_4(T)) \tag{156}$$

by the inf in the definition of value function. However, by applying again the inf for $u(x_4(t+h), t+h)$

$$u(x_4(t+h), t+h) \le C_{x_4(t+h), t+h}[\alpha_4] \tag{157}$$

$$= \int_{t+h}^T r(x_4(s), \alpha_4(s))\, ds + g(x_4(T)) \tag{158}$$

so we have proved that

$$u(x,t) + \varepsilon \ge \int_t^{t+h} r(x_4(s), \alpha_4(s))\, ds + u(x_4(t+h), t+h) \tag{159}$$

and we will find

$$u(x,t) \geq \inf_{\alpha \in \mathscr{A}} \left\{ \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) \right\} \tag{160}$$

so the theorem is proved. $\square$

**Remark.** *The optimality condition is telling us a very intuitive fact: the optimal control for the process starting from $x$ at time $t$, has already taken the optimal control for the process starting from $x(t+h)$ at time $t+h$ into consideration. As a result, we can view $u(x(t+h), t+h)$ as the terminal cost (only depends on the endpoint $x(t+h)$) and $\int_t^{t+h} r(x(s), \alpha(s))\, ds$ as the running cost (depends on how $x(t)$ behaves in time $[t, t+h]$). One might be able to see the Markovian structure again from this expression.*

To set up a PDE for value function $u(x,t)$, it's natural for us to prove that $u$ is Lipschitz (so it's almost everywhere differentiable and the PDE can hold in the almost everywhere sense).

**Theorem 9.** ***(Boundedness and Lipschitz Continuity of Value Function)*** *The value function $u(x,t)$ under the assumptions above is bounded and Lipschitz on $\mathbb{R}^n \times [0,T]$.*

*Proof.* Since $u(x,t) = \inf_{\alpha \in \mathscr{A}} C_{x,t}[\alpha]$ and $r, g$ are assumed to be bounded, it's obvious that $u$ is also bounded

$$u(x,t) \leq \sup |r| \cdot T + \sup |g| \tag{161}$$

Now fix $t \in [0,T]$ and consider $x, \hat{x} \in \mathbb{R}$, apply the inf in the definition of value function, so $\forall \varepsilon > 0$, there exists control $\hat{\alpha}$ and $\hat{x}(s)$ as the solution to the ODE with fixed control $\hat{\alpha}$ and initial value condition $\hat{x}(t) = \hat{x}$ such that

$$u(\hat{x}, t) + \varepsilon \geq \int_t^T r(\hat{x}(s), \hat{\alpha}(s))\, ds + g(\hat{x}(T)) \tag{162}$$

so let's estimate the difference

$$u(x,t) - u(\hat{x}, t) \leq u(x,t) - \int_t^T r(\hat{x}(s), \hat{\alpha}(s))\, ds - g(\hat{x}(T)) + \varepsilon \tag{163}$$

$$\leq \int_t^T r(x(s), \hat{\alpha}(s))\, ds + g(x(T)) - \int_t^T r(\hat{x}(s), \hat{\alpha}(s))\, ds - g(\hat{x}(T)) + \varepsilon \tag{164}$$

note that here we are taking $x(s)$ as the solution to the ODE with initial value condition $x(t) = x$ that

$$x'(s) = f(x(s), \hat{\alpha}(s)) \tag{165}$$

since $r, g$ are Lipschitz with Lipschitz constant $C_r, C_g$,

$$u(x,t) - u(\hat{x}, t) \leq C_r \int_t^T ||x(s) - \hat{x}(s)||\, ds + C_g ||x(T) - \hat{x}(T)|| + \varepsilon \tag{166}$$

in order to estimate $\int_t^T ||x(s)-\hat{x}(s)||\,ds$, note that since $f$ is also Lipschitz with constant $C_f$,

$$\begin{aligned} ||x'(s)-\hat{x}'(s)|| &= ||f(x(s),\hat{\alpha}(s))-f(\hat{x}(s),\hat{\alpha}(s))|| && (167)\\ &\le C_f||x(s)-\hat{x}(s)|| && (168)\end{aligned}$$

by Grownwall's inequality,

$$||x(s)-\hat{x}(s)|| \le C||x(t)-\hat{x}(t)|| = C||x-\hat{x}|| \tag{169}$$

that's why

$$u(x,t)-u(\hat{x},t)\le CT||x-\hat{x}||+\varepsilon \tag{170}$$

for some constant $C$ and thus $u$ is Lipschitz in variable $x$ (the other side is similar).

To prove that it's also Lipschitz in variable $t$, let's fix $x\in\mathbb{R}^n$ and consider $t,\hat{t}\in[0,T]$. For $\forall\varepsilon>0$, there exists control $\alpha$ and the solution $x(\cdot)$ to the ODE with fixed control $\alpha$ such that

$$u(x,t)+\varepsilon \ge C_{x,t}[\alpha] = \int_t^T r(x(s),\alpha(s))\,ds+g(x(T)) \tag{171}$$

consider the time-shifted control $\hat{\alpha}(s)=\alpha(s+t-\hat{t})$ and $\hat{x}$ as the solution to the ODE with fixed control $\hat{\alpha}$, one may find $\hat{x}'(s)=f(\hat{x}(s),\hat{\alpha}(s))$ and $\frac{d}{ds}x(s+t-\hat{t})=x'(s+t-\hat{t})=f(x(s+t-\hat{t}),\alpha(s+t-\hat{t}))=f(x(s+t-\hat{t}),\hat{\alpha}(s))$. By the uniqueness of the solution, we know that $\hat{x}(s)=x(s+t-\hat{t}),\hat{x}(\hat{t})=x(t)=x$, so

$$\begin{aligned} u(x,\hat{t})-u(x,t) &\le u(x,\hat{t})-\int_t^T r(x(s),\alpha(s))\,ds-g(x(T))+\varepsilon && (172)\\ &\le \int_{\hat{t}}^T r(\hat{x}(s),\hat{\alpha}(s))\,ds+g(\hat{x}(T))-\int_t^T r(x(s),\alpha(s))\,ds-g(x(T))+\varepsilon && (173)\\ &\le \int_T^{T-\hat{t}+t} r(x(s),\alpha(s))\,ds+g(\hat{x}(T))-g(x(T))+\varepsilon && (174)\\ &\le \sup|r|\cdot|t-\hat{t}|+C_g\cdot||\hat{x}(T)-x(T)||+\varepsilon && (175)\\ &\le C\cdot|t-\hat{t}|+\varepsilon && (176)\end{aligned}$$

since $||\hat{x}(T)-x(T)||\le\sup|f|\cdot|T+t-\hat{t}-T|=\sup|f|\cdot|t-\hat{t}|$ so we have proved that $u(x,t)$ is also Lipschitz in $t$ (the other side is similar). $\square$

## Hamilton-Jacobi-Bellman Equation (HJBE)

Now from the optimality condition and the Lipschitz continuity of the value function derived above, we can set up a PDE describing the evolution of value function $u(x,t)$.

**Theorem 10.** ***(HJBE for Value Function)*** *The value function under assumptions above satisfies the HJBE*

$$\begin{cases} u_t + \inf_{\alpha\in\mathscr{A}} \{f(x,\alpha)\cdot Du + r(x,\alpha)\} = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ u = g & \text{on } \mathbb{R}^n \times \{t\} \end{cases} \tag{177}$$

*Proof.* When $t = T$, $u = \inf_{\alpha\in\mathscr{A}} C_{x,T}[\alpha] = \int_T^T r(x(s),\alpha(s))\,ds + g(x(T)) = g(x)$ gives the terminal condition.

When $0 < t < T$, recall the optimality condition that for $h > 0$ such that $t + h \leq T$,

$$u(x,t) = \inf_{\alpha\in\mathscr{A}} \left\{ \int_t^{t+h} r(x(s),\alpha(s))\,ds + u(x(t+h),t+h) \right\} \tag{178}$$

where $x(\cdot)$ is the solution to the ODE for fixed control $\alpha$. Let's modify both sides of this property to get HJBE, be careful with the difference between $x$ and $x(\cdot)$ since the previous one denotes the initial value while the latter one denotes the solution to the PDE

$$\frac{u(x,t) - u(x,t+h)}{h} = \inf_{\alpha\in\mathscr{A}} \left\{ \frac{1}{h}\int_t^{t+h} r(x(s),\alpha(s))\,ds + \frac{u(x(t+h),t+h) - u(x,t+h)}{h} \right\} \tag{179}$$

$$= \inf_{\alpha\in\mathscr{A}} \left\{ \frac{1}{h}\int_t^{t+h} r(x(s),\alpha(s))\,ds + \frac{u(x(t+h),t+h) - u(x(t),t+h)}{h} \right\} \tag{180}$$

setting $h \to 0^+$ on both sides to find

$$-u_t(x,t) = \inf_{\alpha\in\mathscr{A}} \{r(x(t),\alpha(t)) + Du(x(t),t)\cdot x'(t)\} \tag{181}$$

$$= \inf_{\alpha\in\mathscr{A}} \{r(x(t),\alpha(t)) + Du(x,t)\cdot f(x(t),\alpha(t))\} \tag{182}$$

now let's neglect the initial time $t$ and initial value $x$ to denote the PDE as

$$u_t + \inf_{\alpha\in\mathscr{A}} \{r(x,\alpha) + Du\cdot f(x,\alpha)\} = 0 \tag{183}$$

note that $u$ being Lipschitz guarantees that the partial derivatives w.r.t. each variable exists almost everywhere. □

**Remark.** *We can find the connection between HJE and HJBE that if we set the Hamiltonian as*

$$H(p,x) = \inf_{\alpha\in\mathscr{A}} \{f(x,\alpha)\cdot p + r(x,\alpha)\} \tag{184}$$

*then HJBE is just HJE $u_t + H(Du,x) = 0$ but with a* ***terminal value condition*** *instead of an initial value condition.*

**Remark.** *One may still recall the Hopf-Lax formula mentioned above to solve HJE $u_t + H(Du) = 0$ with initial value condition $u(x,0) = g(x)$ that*

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \tag{185}$$

*with the Lagrangian $L$ as the Frenchel conjugate of the Hamiltonian $H$. We can verify that such $u(x,t)$ also provides us with the solution to a special kind of HJBE.*

*Now that HJE has initial value condition but HJBE has terminal value condition, the most natural way is to do the time reflection $v(x,t) = u(x, T-t)$ such that the terminal value condition of $u$ actually gives the initial value condition of $v$. It's easy to see that*

$$v(x,0) = u(x,T) = g(x) \tag{186}$$

*then notice that $v_t = -u_t, Dv = Du$, so the HJBE for $u$ can be reformulated as the following HJE for $v$ that*

$$\begin{cases} v_t + H(Dv, x) = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ v = g & \text{on } \mathbb{R}^n \times \{0\} \end{cases} \tag{187}$$

*with Hamiltonian*

$$H(p,x) = -\inf_{\alpha\in\mathscr{A}} \{f(x,\alpha)\cdot p + r(x,\alpha)\} \tag{188}$$

*However, in order to let the Hopf-Lax formula work, we have to* ***assume that*** $r(x,\alpha) = r(\alpha), f(x,\alpha) = f(\alpha)$***, i.e. both running reward and the dynamics does not depend on the state*** $x$*. So the HJE and the Hamiltonian becomes*

$$\begin{cases} v_t + H(Dv) = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ v = g & \text{on } \mathbb{R}^n \times \{0\} \end{cases} \tag{189}$$

*and*

$$H(p) = -\inf_{\alpha\in\mathscr{A}} \{f(\alpha)\cdot p + r(\alpha)\} \tag{190}$$

*So the Frenchel conjugate is*

$$L(v) = \sup_{p\in\mathbb{R}^n} \left\{ p\cdot v + \inf_{\alpha\in\mathscr{A}} \{f(\alpha)\cdot p + r(\alpha)\} \right\} \tag{191}$$

*and the solution to HJE is given by*

$$v(x,t) = \inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right)+g(y)\right\} \tag{192}$$

$$= \inf_{y\in\mathbb{R}^n}\left\{\sup_{p\in\mathbb{R}^n}\left\{p\cdot(x-y)+t\inf_{\alpha\in\mathscr{A}}\{f(\alpha)\cdot p+r(\alpha)\}\right\}+g(y)\right\} \tag{193}$$

*as a result,* ***the solution to HJBE*** *is*

$$u(x,t) = v(x,T-t) = \inf_{y\in\mathbb{R}^n}\left\{\sup_{p\in\mathbb{R}^n}\left\{p\cdot(x-y)+(T-t)\inf_{\alpha\in\mathscr{A}}\{f(\alpha)\cdot p+r(\alpha)\}\right\}+g(y)\right\} \tag{194}$$

***under the assumption that*** $g$ ***is Lipschitz,*** $H$ ***is convex and*** $\lim_{||p||\to\infty}\frac{H(p)}{||p||}=+\infty$.

*However, one might realize that although we have got an analytic solution for HJBE, the assumption that the running reward and the dymanics both do not depend on state is too strong that most of the interesting examples would not satisfy such assumption. This assumption only works well for a problem setting with a single state and many actions to be chosen, i.e. the continuous-time bandit problem but fails for most reinforcement learning problems.*

*Although one would not be able to solve the HJBE analytically in all cases, our previous discussion about general HJE* $u_t+H(Du,x)=0$ *still provides some insights. One can consider the Hamilton's equation and the Euler-Lagrange equations associated with the HJBE.*

## Infinite-Horizon Problem

Among our discussion, we are assuming that there exists some upper time limit $T<\infty$ and the dynamics works in time interval $[0,T]$. However, one can also consider the infinite-horizon problem by taking $T=\infty$. Let's adopt all same assumptions for $A,f,r,g$ above, and consider the admissible set

$$\mathscr{A} = \{\alpha:[0,\infty)\to A : \alpha(\cdot)\ \textit{measurable}\} \tag{195}$$

with $x(\cdot)$ as the unique solution to ODE

$$\begin{cases} x'(s) = f(x(s),\alpha(s)) \\ x(0) = x \end{cases} \tag{196}$$

for fixed control $\alpha$. In order to ensure that the cost is well-defined on infinite time horizon, let's introduce $\lambda>0$ as continuous-time discount rate and define the cost as

$$C_x[\alpha] = \int_0^\infty e^{-\lambda s} r(x(s),\alpha(s))\,ds \tag{197}$$

and the value function as

$$u(x) = \inf_{\alpha \in \mathscr{A}} C_x[\alpha] \tag{198}$$

note that **the biggest difference is that infinite time horizon problem under the Markovian setting has time-homogeneous value function**.

**Remark.** *To see this, let's assume that former definition still applies*

$$C_{x,t}[\alpha] = \int_t^\infty e^{-\lambda s} r(x(s), \alpha(s))\, ds \tag{199}$$

*and the value function is*

$$u(x,t) = \inf_{\alpha \in \mathscr{A}} C_{x,t}[\alpha] \tag{200}$$

*with the ODE having initial value condition $x(t) = x$. Now consider $\forall t > 0$,*

$$C_{x,t}[\alpha] = \int_t^\infty e^{-\lambda s} r(x(s), \alpha(s))\, ds \tag{201}$$

$$= \int_0^\infty e^{-\lambda(s+t)} r(x(s+t), \alpha(s+t))\, ds \tag{202}$$

*where $x'(s) = f(x(s), \alpha(s)), x(t) = x$. However, let's consider another solution $\hat{x}(s)$ to the ODE with fixed control $\hat{\alpha}(s) = \alpha(s+t)$ such that $\hat{x}'(s) = f(\hat{x}(s), \hat{\alpha}(s)), \hat{x}(0) = x$, according to the uniqueness of the solution to the ODE, we immediately know that $\hat{x}(s) = x(s+t)$. So now*

$$C_{x,t}[\alpha] = \int_0^\infty e^{-\lambda(s+t)} r(x(s+t), \alpha(s+t))\, ds \tag{203}$$

$$= e^{-\lambda t} \cdot \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds \tag{204}$$

$$= e^{-\lambda t} \cdot C_{x,0}[\hat{\alpha}] \tag{205}$$

*and by taking inf on both sides, one would see that*

$$u(x,t) = e^{-\lambda t} \cdot u(x,0) \tag{206}$$

*so the time $t$ only appears in the discount factor $e^{-\lambda t}$. That's why we only need to consider $u(x,0)$ and denote it as $u(x)$ by taking the time $t$ as 0 by default.*

Under all assumptions made above, one can see that $u$ **is bounded and if** $\lambda > Lips(f)$ **then** $u$ **is Lipschitz**. To argue this, one do the similar thing as done in the previous proofs. $\forall x, \hat{x} \in \mathbb{R}^n, \forall \varepsilon > 0$, there exists control

$\hat{\alpha} \in \mathscr{A}$ and the solution $\hat{x}(s)$ to the ODE with fixed control $\hat{\alpha}$ and initial value condition $\hat{x}(0) = \hat{x}$ such that

$$u(\hat{x}) + \varepsilon \geq \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds \tag{207}$$

now by definition,

$$u(x) - u(\hat{x}) \leq u(x) - \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds + \varepsilon \tag{208}$$

$$\leq \int_0^\infty e^{-\lambda s} r(x(s), \hat{\alpha}(s))\, ds - \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds + \varepsilon \tag{209}$$

where $x(s)$ is the solution to the ODE with fixed control $\hat{\alpha}$ and initial value condition $x(0) = x$. So we know that

$$u(x) - u(\hat{x}) \leq \int_0^\infty e^{-\lambda s} [r(x(s), \hat{\alpha}(s)) - r(\hat{x}(s), \hat{\alpha}(s))]\, ds + \varepsilon \tag{210}$$

$$\leq C_r \cdot \int_0^\infty e^{-\lambda s} \cdot ||x(s) - \hat{x}(s)||\, ds + \varepsilon \tag{211}$$

and $||x'(s) - \hat{x}'(s)|| = ||f(x(s), \hat{\alpha}(s)) - f(\hat{x}(s), \hat{\alpha}(s))|| \leq C_f \cdot ||x(s) - \hat{x}(s)||$ so by Grownwall's inequality, we conclude that

$$||x(s) - \hat{x}(s)|| \leq e^{C_f s} \cdot ||x(0) - \hat{x}(0)|| = e^{C_f s} \cdot ||x - \hat{x}|| \tag{212}$$

so the estimates look like

$$u(x) - u(\hat{x}) \leq C_r \cdot ||x - \hat{x}|| \cdot \int_0^\infty e^{(C_f - \lambda)s}\, ds + \varepsilon \tag{213}$$

so when $C_f = Lips(f) < \lambda$, the integral converges and is a constant, that's why $u$ is Lipschitz and is differentiable almost everywhere.

To get the HJBE for such value function $u(x)$, let's plug in

$$u(x, t) = e^{-\lambda t} \cdot u(x, 0) \tag{214}$$

into the HJBE we derived for general optimal control problem to see that

$$u_t(x, t)|_{t=0} = -\lambda \cdot u(x, 0) \tag{215}$$

so

$$-\lambda \cdot u(x, 0) + \inf_{\alpha \in \mathscr{A}} \{f(x, \alpha) \cdot Du + r(x, \alpha)\} = 0 \tag{216}$$

and we get **the HJBE for infinite-horizon optimal control problem**

$$\lambda u - \inf_{\alpha \in \mathscr{A}} \{f(x,\alpha) \cdot Du + r(x,\alpha)\} = 0 \tag{217}$$

for value function $u = u(x)$.