Week 4. Derivation of the Hamilton–Jacobi–Bellman Equation

Primary references: Hung Vinh Tran, *Hamilton–Jacobi Equations*, §2.1, Theorem 2.4, and §2.3, Theorem 2.11

1. Overview

In this course, we do not follow the textbook in a strictly linear order.
Instead, we first formulate deterministic optimal control problems and value functions, then study the Dynamic Programming Principle (DPP), and only afterward derive the associated Hamilton–Jacobi equation.

The goal of this lecture is to explain how the stationary Hamilton–Jacobi–Bellman (HJB) equation arises from the DPP. The derivation is initially formal, since it assumes differentiability of the value function. Nevertheless, it reveals the correct PDE structure and explains why the Hamiltonian is defined in the way used throughout this course.

A second goal of this lecture is to prepare for the transition to viscosity solutions. We will state that the value function is a viscosity solution of the HJB equation, but the formal definition of viscosity solution will be postponed until Week 6.

Thus the conceptual flow of Weeks 2–4 is:

$$\text{controlled dynamics} \longrightarrow \text{value function} \longrightarrow \text{DPP} \longrightarrow \text{HJB equation}.$$

2. Learning Objectives

By the end of this lecture, students should be able to:

1. recall the discounted infinite-horizon optimal control problem and its value function;
2. explain formally how the DPP leads to the stationary HJB equation;
3. define the Hamiltonian under the sign convention used in this course;
4. state the basic structural properties of the Hamiltonian;
5. understand why the value function should be interpreted as a viscosity solution;
6. compute the Hamiltonian in several standard examples.

3. Recall: Controlled Dynamics, Cost, and Value Function

We consider the deterministic controlled system

$$\begin{cases} y'(s) = b(y(s), v(s)), & s > 0, \\ y(0) = x, \end{cases}$$

where $x \in \mathbb{R}^n$ is the initial state, V is a nonempty compact metric control space, and an admissible control is a measurable map $v(\cdot) : [0, \infty) \to V$.

Throughout this lecture, we assume:

1. $b \in C(\mathbb{R}^n \times V; \mathbb{R}^n)$;
2. b is bounded and Lipschitz continuous in x, uniformly in v;
3. the running cost $f : \mathbb{R}^n \times V \to \mathbb{R}$ is continuous, bounded, and Lipschitz continuous in x, uniformly in v;
4. the discount factor satisfies $\lambda > 0$.

For each admissible control $v(\cdot)$, let $y_{x,v(\cdot)}(\cdot)$ denote the corresponding trajectory starting from x. The infinite-horizon discounted cost is

$$J(x, v(\cdot)) := \int_0^\infty e^{-\lambda s} f(y_{x,v(\cdot)}(s), v(s))\, ds,$$

and the value function is defined by

$$u(x) := \inf_{v(\cdot)} J(x, v(\cdot)).$$

Thus $u(x)$ is the infimal total discounted cost when the system starts from x.

4. Recall: The Dynamic Programming Principle

The key result from Week 3 is the Dynamic Programming Principle:

$$u(x) = \inf_{v(\cdot)} \left\{ \int_0^t e^{-\lambda s} f(y_x(s), v(s))\, ds + e^{-\lambda t} u(y_x(t)) \right\} \qquad \text{for every } t > 0.$$

This identity expresses the principle of optimality: the infimal cost from time 0 to ∞ can be decomposed into

1. the cost accumulated on the short interval $[0, t]$, and
2. the discounted continuation value starting from the state reached at time t.

The HJB equation is obtained by analyzing this identity in the small-time limit.

5. Why a PDE Should Emerge

The DPP is valid for every $t > 0$.
If t is very small, then the trajectory remains close to its initial point x, and the short-time cost can be approximated by a first-order expansion. Therefore, one expects the DPP to admit an infinitesimal form.

This is the fundamental idea: the value function is defined globally by optimization over an infinite time horizon, but the DPP makes it possible to express that global optimization in a local-in-time form. Once a local-in-time formula is available, a differential equation naturally appears.

6. Formal Derivation of the HJB Equation

We now derive the equation formally.
For this purpose only, assume temporarily that

$$u \in C^1(\mathbb{R}^n).$$

Starting from the DPP,

$$u(x) = \inf_{v(\cdot)} \left\{ \int_0^t e^{-\lambda s} f(y_x(s), v(s))\, ds + e^{-\lambda t} u(y_x(t)) \right\},$$

subtract $u(x)$ from both sides and divide by t:

$$0 = \inf_{v(\cdot)} \left\{ \frac{1}{t} \int_0^t e^{-\lambda s} f(y_x(s), v(s))\, ds + \frac{e^{-\lambda t} u(y_x(t)) - u(x)}{t} \right\}.$$

To identify the first-order terms, freeze the control at a constant value $v(\cdot) \equiv v$. Then, for small t,

$$y_x(t) = x + t\, b(x, v) + o(t),$$

and

$$e^{-\lambda t} = 1 - \lambda t + o(t).$$

Moreover,

$$\int_0^t e^{-\lambda s} f(y_x(s), v)\, ds = t\, f(x, v) + o(t),$$

and, by Taylor expansion,

$$u(y_x(t)) = u(x) + t\, b(x,v)\cdot Du(x) + o(t).$$

Since every constant control is admissible, substituting these expansions into the DPP first gives

$$f(x,v) - \lambda u(x) + b(x,v)\cdot Du(x) \geq 0 \qquad \text{for every } v \in V.$$

For the reverse inequality, one chooses, for each small t, a control whose DPP expression is within $o(t)$ of the infimum and uses the uniform short-time expansions. Thus the formal argument gives

$$0 = \inf_{v\in V}\{f(x,v) - \lambda u(x) + b(x,v)\cdot Du(x)\}.$$

This is the Bellman form of the equation.

7. Definition of the Hamiltonian

To write the equation more compactly, we define the Hamiltonian by

$$H(x,p) := \sup_{v\in V}\{-b(x,v)\cdot p - f(x,v)\}.$$

With this definition, the Bellman form becomes

$$\boxed{\lambda u(x) + H(x, Du(x)) = 0.}$$

This is the **stationary Hamilton–Jacobi–Bellman equation** associated with the discounted infinite-horizon control problem.

For the one-dimensional eikonal control model

$$b(x,v) = v, \qquad v \in [-1,1], \qquad f(x,v) = 1,$$

each fixed control produces an affine function of the gradient variable,

$$p \longmapsto -vp - 1.$$

Taking the supremum over the control set forms the upper envelope $H(p) = |p| - 1$. The pointwise maximizing branch also identifies the feedback selected by the Hamiltonian.

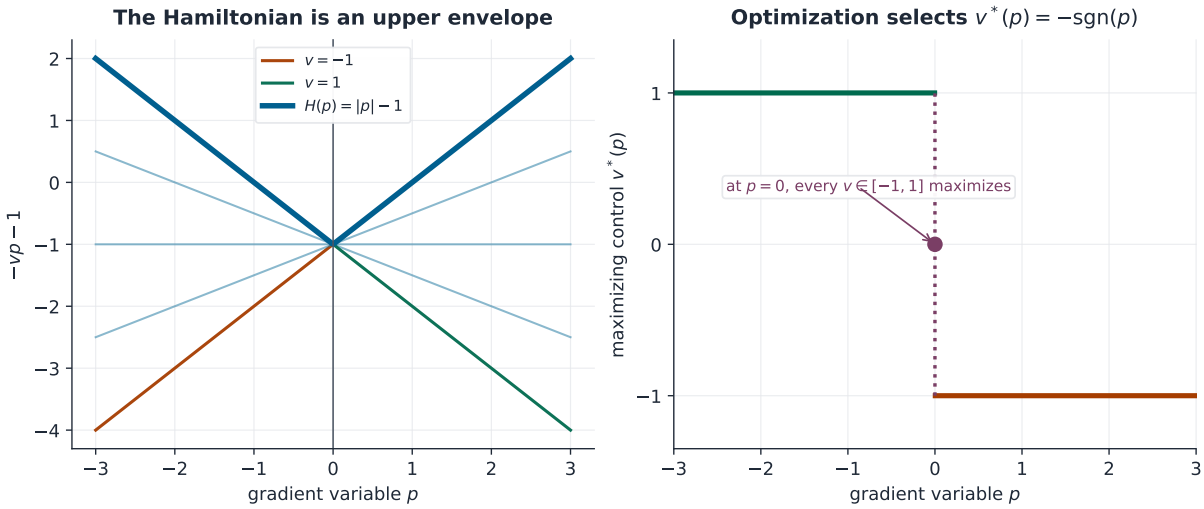


Figure 4.1. Left: $H(p) = \sup_{v\in[-1,1]}(-vp-1) = |p| - 1$ is the upper envelope of the affine branches associated with fixed controls. Right: for $p \neq 0$, the maximizing control is $v^*(p) = -\operatorname{sgn}(p)$; at $p = 0$, every control is a maximizer.

8. Interpretation of the Terms

Each term in the HJB equation has a clear origin.

8.1 The running cost term

The term

$$f(x, v)$$

comes from the short-time cost accumulated on the interval $[0, t]$.

8.2 The discount term

The term

$$-\lambda u(x)$$

comes from the expansion of the discount factor

$$e^{-\lambda t} = 1 - \lambda t + o(t).$$

8.3 The transport term

The term

$$b(x, v) \cdot Du(x)$$

describes the first-order change of the value function along the controlled trajectory.

Thus, the HJB equation is the infinitesimal form of dynamic optimality.

9. Theorem 2.11: The Value Function Is a Viscosity Solution

The previous derivation assumed that u is C^1, but in general the value function need not be differentiable. The correct result is the following.

Theorem 9.1 (Tran, Theorem 2.11)

The value function u is a viscosity solution of

$$\lambda u(x) + H(x, Du(x)) = 0 \qquad \text{in } \mathbb{R}^n.$$

At this stage of the course, we do **not** yet give the formal definition of viscosity solution. That definition will be introduced in Week 6. For now, the theorem should be understood as saying that the HJB equation remains valid even when the value function is not classically differentiable.

Proof idea (preview only)

We only indicate the main strategy.

Subsolution side

Let ϕ be a smooth test function touching u from above at x_0.
Using the DPP with a constant control $v(\cdot) \equiv v$, dividing by small time, and passing to the limit, one obtains

$$\lambda \phi(x_0) + H(x_0, D\phi(x_0)) \leq 0.$$

Supersolution side

Let ψ be a smooth test function touching u from below at x_0.
This direction is more delicate. One combines the lower bound from the DPP with uniform control of trajectories, based on the Lipschitz dependence of the flow, and then passes to the limit to obtain

$$\lambda\psi(x_0) + H(x_0, D\psi(x_0)) \geq 0.$$

We will return to this argument after the formal definition of viscosity solution is introduced.

10. Basic Properties of the Hamiltonian

The Hamiltonian

$$H(x,p) = \sup_{v\in V}\{-b(x,v)\cdot p - f(x,v)\}$$

inherits important structure from the control problem.

Theorem 10.1 (Tran, Theorem 2.4)

The Hamiltonian satisfies the following properties:

1. $H \in C(\mathbb{R}^n \times \mathbb{R}^n)$.
2. For each fixed x, the map $p \mapsto H(x,p)$ is convex.
3. There exists a constant $C > 0$ such that

$$|H(x,p) - H(x,q)| \leq C|p-q|$$

for all $x, p, q \in \mathbb{R}^n$.

4. There exists a constant $C > 0$ such that

$$|H(x,p) - H(y,p)| \leq C(1+|p|)|x-y|$$

for all $x, y, p \in \mathbb{R}^n$.

Why convexity holds

For each fixed v, the function

$$p \mapsto -b(x,v)\cdot p - f(x,v)$$

is affine in p. Since the supremum of affine functions is convex, it follows that $p \mapsto H(x,p)$ is convex.
This convexity is one of the most important structural features of Hamiltonians arising from optimal control.

11. Sign Convention

Different books use different sign conventions for the Hamiltonian.
To avoid confusion, we state our convention explicitly.

Convention	Formula	Comment
Tran / this course	$H(x,p) = \sup_{v\in V}\{-b(x,v)\cdot p - f(x,v)\}$, $\lambda u + H(x,Du) = 0$	**Used throughout this course**
Some control texts	$\inf_{v\in V}\{b(x,v)\cdot p + f(x,v)\} = 0$	Equivalent up to a minus sign
Some PDE texts	$H(x,p) = \sup_{v\in V}\{b(x,v)\cdot p - f(x,v)\}$	Opposite sign in the drift term

We will consistently use the **Tran convention**:

$$H(x,p) = \sup_{v\in V}\{-b(x,v)\cdot p - f(x,v)\}, \qquad \lambda u(x) + H(x,Du(x)) = 0.$$

12. Standard Examples

We now compute the Hamiltonian in several model cases.

Example 4.1 (Eikonal type)

Let

$$V = \overline{B(0,1)}, \qquad b(x,v) = v, \qquad f(x,v) = f(x).$$

Then

$$H(x,p) = \sup_{|v|\le 1}\{-v\cdot p - f(x)\} = |p| - f(x).$$

Hence the HJB equation is

$$\lambda u(x) + |Du(x)| - f(x) = 0.$$

This is the basic eikonal-type structure.

Example 4.2 (Minimum time to a target)

In this example, we replace the discounted infinite-horizon cost by an **undiscounted cost stopped upon reaching a target**.

Let $K \subsetneq \mathbb{R}^n$ be a nonempty closed target set. Assume that a is Lipschitz continuous and

$$0 < a_* \le a(x) \le a^* < \infty.$$

Consider

$$V = \overline{B(0,1)}, \qquad b(x,v) = a(x)v, \qquad f \equiv 1,$$

so that

$$y'(s) = a(y(s))v(s), \qquad y(0) = x, \qquad |v(s)| \le 1.$$

For an admissible control, define the first hitting time

$$\tau_K(x, v(\cdot)) := \inf\{s \ge 0 : y_{x,v(\cdot)}(s) \in K\},$$

with $\inf\emptyset = \infty$. The total cost and the minimum-time value function are

$$J(x, v(\cdot)) := \int_0^{\tau_K(x,v(\cdot))} 1\, ds = \tau_K(x, v(\cdot)), \qquad T(x) := \inf_{v(\cdot)} \tau_K(x, v(\cdot)).$$

Thus $T(x)$ is the least time needed to reach K. The lower bound $a \ge a_* > 0$ ensures that the target can be reached in finite time from every initial state.

Under the sign convention of this course,

$$H(x,p) = \sup_{|v|\le 1}\{-a(x)v\cdot p - 1\} = a(x)|p| - 1.$$

The stopped DPP gives $H(x, DT(x)) = 0$ outside the target. Hence the minimum-time equation is

$$\boxed{\begin{cases} a(x)|DT(x)| = 1 & \text{in } \mathbb{R}^n \setminus K, \\ T(x) = 0 & \text{on } K. \end{cases}}$$

The PDE is understood in the viscosity sense when T is not differentiable. There is no λT term because the cost is not discounted; the discounted whole-space theorem stated above is not being applied to this stopping problem.

For example, if $a \equiv 1$ and $K = \{0\}$, then $T(x) = |x|$: the fastest path moves straight toward the origin at unit speed. The equation $|DT| = 1$ is imposed only outside the target, with $T(0) = 0$.

Example 4.3 (Anisotropic minimum-time type structure)

Let V be compact, let $b(x,v)$ be general, and let

$$f \equiv 1.$$

Then

$$H(x,p) = \sup_{v\in V}\{-b(x,v)\cdot p - 1\} = \left(\sup_{v\in V}\{-b(x,v)\cdot p\}\right) - 1.$$

If $B(x) := \{b(x,v) : v \in V\}$ and $\sigma_A(p) := \sup_{q\in A} q\cdot p$, then

$$H(x,p) = \sigma_{B(x)}(-p) - 1 = \sigma_{-B(x)}(p) - 1.$$

Thus the Hamiltonian is the support function of the reflected admissible velocity set, shifted by the constant running cost 1.

This is the natural anisotropic generalization of the eikonal structure and has the flavor of a minimum-time problem.

Example 4.4 (One-dimensional LQ model)

Let

$$V = \mathbb{R}, \qquad b(x,v) = v, \qquad f(x,v) = \frac{1}{2}(x^2+v^2).$$

Then

$$H(x,p) = \sup_{v\in\mathbb{R}}\left\{-vp - \frac{1}{2}(x^2+v^2)\right\}.$$

Complete the square:

$$-vp - \frac{1}{2}v^2 = -\frac{1}{2}(v+p)^2 + \frac{1}{2}p^2.$$

Hence

$$H(x,p) = \frac{1}{2}p^2 - \frac{1}{2}x^2.$$

So the HJB equation becomes

$$\lambda u(x) + \frac{1}{2}|u_x(x)|^2 - \frac{1}{2}x^2 = 0.$$

Remark.
This example lies outside the standing framework: $V = \mathbb{R}$ is not compact, $b(x,v) = v$ is not uniformly bounded, and $f(x,v) = \frac{1}{2}(x^2+v^2)$ is neither bounded nor globally Lipschitz in x. It should therefore be understood as a model computation and a preview of the linear-quadratic theory.

Example 4.5 (G-equation type Hamiltonian)

The PDE form of the G-equation is

$$u_t + |Du| + W(x)\cdot Du = 0,$$

so the associated Hamiltonian is

$$H(x,p) = |p| + W(x)\cdot p.$$

Under the sign convention of this course, this Hamiltonian is obtained by taking

$$V = \overline{B(0,1)}, \qquad b(x,v) = v - W(x), \qquad f \equiv 0.$$

Indeed,

$$H(x,p) = \sup_{|v|\le 1}\{-(v - W(x)) \cdot p\} = |p| + W(x) \cdot p.$$

Important sign remark.
If one instead writes $b(x,v) = v + W(x)$, then

$$H(x,p) = \sup_{|v|\le 1}\{-(v + W(x)) \cdot p\} = |p| - W(x) \cdot p,$$

which gives the opposite sign in front of $W(x) \cdot p$.
Therefore, one must check the drift carefully when matching a control system with a PDE sign convention.

13. Summary

The main lesson of this lecture is that the HJB equation is the infinitesimal form of the Dynamic Programming Principle.

Starting from the value function and the DPP, and assuming temporarily that u is differentiable, we obtain the Bellman form

$$0 = \inf_{v\in V}\{f(x,v) - \lambda u(x) + b(x,v) \cdot Du(x)\},$$

which can be rewritten as

$$\lambda u(x) + H(x, Du(x)) = 0.$$

Thus the essential logic of Weeks 2–4 is:

1. define the value function by optimal control;
2. establish the Dynamic Programming Principle;
3. pass formally to the small-time limit;
4. obtain the stationary HJB equation;
5. interpret the value function as a viscosity solution when classical differentiability fails.

The four key formulas are:

$$u(x) = \inf_{v(\cdot)} \int_0^\infty e^{-\lambda s} f(y_x(s), v(s))\, ds,$$

$$u(x) = \inf_{v(\cdot)} \left\{ \int_0^t e^{-\lambda s} f(y_x(s), v(s))\, ds + e^{-\lambda t} u(y_x(t)) \right\},$$

$$H(x,p) = \sup_{v\in V}\{-b(x,v) \cdot p - f(x,v)\},$$

and

$$\boxed{\lambda u(x) + H(x, Du(x)) = 0.}$$

14. Exercises

Exercise 4.1

Starting from the DPP, derive formally the Bellman form

$$0 = \inf_{v \in V} \{ f(x,v) - \lambda u(x) + b(x,v) \cdot Du(x) \}.$$

Clearly indicate where the assumptions $u \in C^1$ and $v(\cdot) \equiv v$ are used.

Exercise 4.2

Show that if

$$H(x,p) = \sup_{v \in V} \{ -b(x,v) \cdot p - f(x,v) \},$$

then $p \mapsto H(x,p)$ is convex for each fixed x.

Exercise 4.3

Compute the Hamiltonian for the front propagation model

$$V = \overline{B(0,1)}, \qquad b(x,v) = a(x)v, \qquad f \equiv 0.$$

Exercise 4.4

For the one-dimensional LQ example, verify directly that

$$H(x,p) = \frac{1}{2}p^2 - \frac{1}{2}x^2.$$

Exercise 4.5

Check carefully the sign in the G-equation example and explain why the choice

$$b(x,v) = v - W(x)$$

produces the Hamiltonian

$$H(x,p) = |p| + W(x) \cdot p.$$

References

- Hung Vinh Tran, *Hamilton–Jacobi Equations*, §2.1, Theorem 2.4, and §2.3, Theorem 2.11.