Choose $0 < \epsilon,\ \lambda < 1$ and set

$$\Phi(x,y,t,s) := u(x,t) - \tilde{u}(y,s) - \lambda(t+s) - \frac{1}{\epsilon^2}(|x-y|^2 + (t-s)^2) - \epsilon(|x|^2 + |y|^2), \tag{5}$$

for $x, y \in \mathbb{R}^n,\ t, s \geq 0$. Then there exists a point $(x_0, y_0, t_0, s_0) \in \mathbb{R}^{2n} \times [0,T]^2$ such that

$$\Phi(x_0, y_0, t_0, s_0) = \max_{\mathbb{R}^{2n}\times[0,T]^2} \Phi(x,y,t,s). \tag{6}$$

2. We may fix $0 < \epsilon,\ \lambda < 1$ so small that (4) implies

$$\Phi(x_0, y_0, t_0, s_0) \geq \sup_{\mathbb{R}^n\times[0,T]} \Phi(x,x,t,t) \geq \frac{\sigma}{2}. \tag{7}$$

In addition, $\Phi(x_0, y_0, t_0, s_0) \geq \Phi(0,0,0,0)$; and therefore

$$\begin{aligned} \lambda(t_0 + s_0) + \frac{1}{\epsilon^2}(|x_0 - y_0|^2 + (t_0 - s_0)^2) + \epsilon(|x_0|^2 + |y_0|^2) \\ \leq u(x_0, t_0) - \tilde{u}(y_0, s_0) - u(0,0) + \tilde{u}(0,0). \end{aligned} \tag{8}$$

Since u and $\tilde{u}$ are bounded, we deduce

$$|x_0 - y_0|, |t_0 - s_0| = O(\epsilon) \quad \text{as } \epsilon \to 0. \tag{9}$$

Furthermore (8) implies $\epsilon(|x_0|^2 + |y_0|^2) = O(1)$, and consequently

$$\begin{aligned} \epsilon(|x_0| + |y_0|) &= \epsilon^{1/4}\epsilon^{3/4}(|x_0| + |y_0|) \\ &\leq \epsilon^{1/2} + C\epsilon^{3/2}(|x_0|^2 + |y_0|^2) \\ &\leq C\epsilon^{1/2}. \end{aligned}$$

Thus

$$\epsilon(|x_0| + |y_0|) = O(\epsilon^{1/2}). \tag{10}$$

3. Since $\Phi(x_0, y_0, t_0, s_0) \geq \Phi(x_0, x_0, t_0, t_0)$, we also have

$$\begin{aligned} u(x_0, t_0) - \tilde{u}(y_0, s_0) - \lambda(t_0 + s_0) - \frac{1}{\epsilon^2}(|x_0 - y_0|^2 + (t_0 - s_0)^2) \\ - \epsilon(|x_0|^2 + |y_0|^2) \geq u(x_0, t_0) - \tilde{u}(x_0, t_0) - 2\lambda t_0 - 2\epsilon|x_0|^2. \end{aligned}$$

Hence

$$\begin{aligned} \frac{1}{\epsilon^2}(|x_0 - y_0|^2 + (t_0 - s_0)^2) \leq \tilde{u}(x_0, t_0) - \tilde{u}(y_0, s_0) + \lambda(t_0 - s_0) \\ + \epsilon(x_0 + y_0)\cdot(x_0 - y_0). \end{aligned}$$

In view of (9), (10) and the uniform continuity of $\tilde{u}$, we deduce

$$|x_0 - y_0|, |t_0 - s_0| = o(\epsilon). \tag{11}$$

4. Now write $\omega(\cdot)$ to denote the modulus of continuity of u; that is,

$$|u(x,t) - u(y,s)| \leq \omega(|x-y| + |t-s|)$$

for all $x, y \in \mathbb{R}^n$, $0 \leq t$, $s \leq T$, and $\omega(r) \to 0$ as $r \to 0$. Similarly, $\tilde{\omega}(\cdot)$ will denote the modulus of continuity of $\tilde{u}$.

Then (7) implies

$$\begin{aligned}\frac{\sigma}{2} \leq u(x_0,t_0) - \tilde{u}(y_0,s_0) &= u(x_0,t_0) - u(x_0,0) + u(x_0,0) - \tilde{u}(x_0,0)\\ &\quad + \tilde{u}(x_0,0) - \tilde{u}(x_0,t_0) + \tilde{u}(x_0,t_0) - \tilde{u}(y_0,s_0)\\ &\leq \omega(t_0) + \tilde{\omega}(t_0) + \tilde{\omega}(o(\epsilon)),\end{aligned}$$

by (9), (11) and the initial condition. We can now take $\epsilon > 0$ to be so small that the foregoing implies $\frac{\sigma}{4} \leq \omega(t_0) + \tilde{\omega}(t_0)$; and this in turn implies $t_0 \geq \mu > 0$ for some constant $\mu > 0$. Similarly we have $s_0 \geq \mu > 0$.

5. Now observe in light of (6) that the mapping $(x,t) \mapsto \Phi(x, y_0, t, s_0)$ has a maximum at the point (x_0, t_0). In view of (5) then,

$$u - v \text{ has a maximum at } (x_0, t_0)$$

for

$$v(x,t) := \tilde{u}(y_0,s_0) + \lambda(t+s_0) + \frac{1}{\epsilon^2}(|x-y_0|^2 + (t-s_0)^2) + \epsilon(|x|^2 + |y_0|^2).$$

Since u is a viscosity solution of (1), we conclude, using the lemma if necessary, that

$$v_t(x_0,t_0) + H(D_x v(x_0,t_0), x_0) \leq 0.$$

Therefore

$$\lambda + \frac{2(t_0 - s_0)}{\epsilon^2} + H\left(\frac{2}{\epsilon^2}(x_0 - y_0) + 2\epsilon x_0, x_0\right) \leq 0. \tag{12}$$

We further observe that since the mapping $(y,s) \mapsto -\Phi(x_0, y, t_0, s)$ has a minimum at the point (y_0, s_0),

$$\tilde{u} - \tilde{v} \text{ has a minimum at } (y_0, s_0)$$

for

$$\tilde{v}(y,s) := u(x_0,t_0) - \lambda(t_0+s) - \frac{1}{\epsilon^2}(|x_0-y|^2 + (t_0-s)^2) - \epsilon(|x_0|^2 + |y|^2).$$

As $\tilde{u}$ is a viscosity solution of (1), we know then that

$$\tilde{v}_s(y_0,s_0) + H(D_y\tilde{v}(y_0,s_0),y_0) \geq 0.$$

Consequently

$$(13) \qquad -\lambda + \frac{2(t_0-s_0)}{\epsilon^2} + H\left(\frac{2}{\epsilon^2}(x_0-y_0) - 2\epsilon y_0, y_0\right) \geq 0.$$

6. Next, subtract (13) from (12):

$$(14) \qquad 2\lambda \leq H\left(\frac{2}{\epsilon^2}(x_0-y_0) - 2\epsilon y_0, y_0\right) - H\left(\frac{2}{\epsilon^2}(x_0-y_0) + 2\epsilon x_0, x_0\right).$$

In view of hypothesis (3) therefore,

$$(15) \qquad \lambda \leq C\epsilon(|x_0| + |y_0|) + C|x_0-y_0|\left(1 + \frac{|x_0-y_0|}{\epsilon^2} + \epsilon(|x_0|+|y_0|)\right).$$

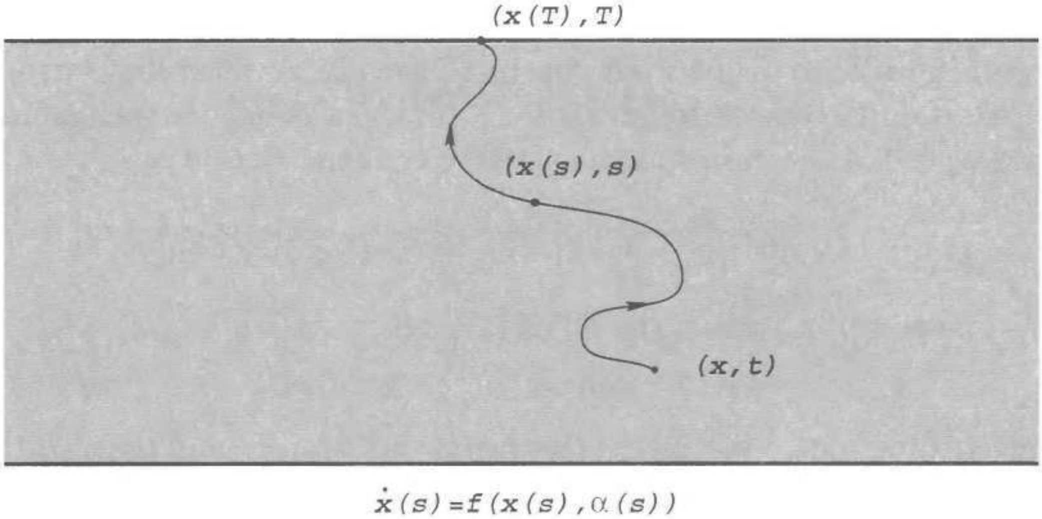
We employ estimates (10), (11) in (15) and then let $\epsilon \to 0$, to discover $0 < \lambda \leq 0$. This contradiction completes the proof. □

10.3. CONTROL THEORY, DYNAMIC PROGRAMMING

It remains for us to establish the existence of a viscosity solution to our initial-value problem for the Hamilton–Jacobi partial differential equation. One method would be now to prove the existence of a smooth solution u^ϵ of the regularized equation (2) in §10.1 and then to make good enough uniform estimates. This technique in fact works but requires knowledge of certain bounds for the heat equation beyond the scope of this book.

In this section we provide an alternative approach of independent interest, which is suitable for Hamiltonians which are convex in p.

We will first of all introduce some of the basic issues concerning *control theory* for ordinary differential equations and the connection with Hamilton–Jacobi PDE afforded by the method of *dynamic programming.* This discussion will make clearer the connections of the theory developed above in §§10.1–10.2 with that set forth earlier in §3.3.1. The remarkable fact is that *the defining viscosity solution inequalities* (16), (17) *in* §10.1.1 *are a consequence of the optimality conditions of control theory.*



Response of system to the control $\alpha(\cdot)$

10.3.1. Introduction to optimal control theory.

We will now study the possibility of optimally controlling the solution $\mathbf{x}(\cdot)$ of the ordinary differential equation

$$(1) \qquad \begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s)) & (t < s < T) \\ \mathbf{x}(t) = x. \end{cases}$$

Here $\dot{} = \frac{d}{ds}$, $T > 0$ is a fixed terminal time, and $x \in \mathbb{R}^n$ is a given initial point, taken on by our solution $\mathbf{x}(\cdot)$ at the starting time $t \geq 0$. At later times $t < s < T$, $\mathbf{x}(\cdot)$ evolves according to the ODE, where

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

is a given bounded, Lipschitz continuous function and A is some given compact subset of, say, $\mathbb{R}^m$. The function $\boldsymbol{\alpha}(\cdot)$ appearing in (1) is a *control*, that is, some appropriate scheme for adjusting parameters from the set A as time evolves, thereby affecting the dynamics of the system modeled by (1).

Let us write

$$(2) \qquad \mathcal{A} := \{\boldsymbol{\alpha} : [0, T] \to A \mid \boldsymbol{\alpha}(\cdot) \text{ is measurable}\}$$

to denote the set of *admissible controls*. Then since

$$(3) \quad |\mathbf{f}(x, a)| \leq C, \ |\mathbf{f}(x, a) - \mathbf{f}(y, a)| \leq C|x - y| \quad (x, y \in \mathbb{R}^n, \ a \in A)$$

for some constant C, we see that for each control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$, the ODE (1) has a unique, Lipschitz continuous solution $\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$, existing on the time interval $[t, T]$ and solving the ODE for a.e. time $t < s < T$. We call $\mathbf{x}(\cdot)$ the *response* of the system to the control $\boldsymbol{\alpha}(\cdot)$, and $\mathbf{x}(s)$ the *state* of the system at time s.

Our goal is to find a control $\boldsymbol{\alpha}^*(\cdot)$ which optimally steers the system. However in order to define what "optimal" means, we must first introduce a *cost criterion.* Given $x \in \mathbb{R}^n$ and $0 \leq t \leq T$, let us define for each admissible control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ the corresponding *cost functional*

$$C_{x,t}[\boldsymbol{\alpha}(\cdot)] := \int_t^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)), \tag{4}$$

where $\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$ solves the ODE (1) and

$$r : \mathbb{R}^n \times A \to \mathbb{R}, \quad g : \mathbb{R}^n \to \mathbb{R}$$

are given functions. We call r the *running cost per unit time* and g the *terminal cost*, and will henceforth assume

$$\begin{cases} |r(x,a)|,\ |g(x)| \leq C \\ |r(x,a) - r(y,a)|,\ |g(x) - g(y)| \leq C|x-y| \end{cases} \qquad (x, y \in \mathbb{R}^n, a \in A) \tag{5}$$

for some constant C.

Given now $x \in \mathbb{R}^n$ and $0 \leq t \leq T$, we would like to find if possible a control $\boldsymbol{\alpha}^*(\cdot)$ which minimizes the cost functional (4) among all other admissible controls. This is a *finite horizon* optimal control problem. (See Problems 10 and 11 for infinite horizon problems.)

10.3.2. Dynamic programming.

The method of *dynamic programming* investigates the above problem by turning attention to the *value function*

$$u(x,t) := \inf_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} C_{x,t}[\boldsymbol{\alpha}(\cdot)] \quad (x \in \mathbb{R}^n,\ 0 \leq t \leq T). \tag{6}$$

The plan is this: having defined $u(x,t)$ as the least cost given that we start at the position x at time t, we want to study u as a function of x and t. We are therefore embedding our given control problem (1), (4) into the larger class of all such problems, as x and t vary. The idea then is to show that u solves a certain Hamilton–Jacobi type PDE and to show conversely that a solution of this PDE helps us to synthesize an optimal feedback control.

Hereafter, we fix $x \in \mathbb{R}^n$, $0 \leq t < T$.

THEOREM 1 (Optimality conditions). *For each $h > 0$ so small that $t + h \leq T$, we have*

$$u(x,t) = \inf_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + u(\mathbf{x}(t+h), t+h) \right\}, \tag{7}$$

where $\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$ solves the ODE (1) *for the control $\boldsymbol{\alpha}(\cdot)$.*

Proof. 1. Choose any control $\boldsymbol{\alpha}_1(\cdot) \in \mathcal{A}$ and solve the ODE

$$\text{(8)} \qquad \begin{cases} \dot{\mathbf{x}}_1(s) = \mathbf{f}(\mathbf{x}_1(s), \boldsymbol{\alpha}_1(s)) & (t < s < t+h) \\ \mathbf{x}_1(t) = x. \end{cases}$$

Fix $\epsilon > 0$ and choose then $\boldsymbol{\alpha}_2(\cdot) \in \mathcal{A}$ so that

$$\text{(9)} \qquad u(\mathbf{x}_1(t+h), t+h) + \epsilon \geq \int_{t+h}^{T} r(\mathbf{x}_2(s), \boldsymbol{\alpha}_2(s))\, ds + g(\mathbf{x}_2(T)),$$

where

$$\text{(10)} \qquad \begin{cases} \dot{\mathbf{x}}_2(s) = \mathbf{f}(\mathbf{x}_2(s), \boldsymbol{\alpha}_2(s)) & (t+h < s < T) \\ \mathbf{x}_2(t+h) = \mathbf{x}_1(t+h). \end{cases}$$

Now define the control

$$\text{(11)} \qquad \boldsymbol{\alpha}_3(s) := \begin{cases} \boldsymbol{\alpha}_1(s) & \text{if } t \leq s < t+h \\ \boldsymbol{\alpha}_2(s) & \text{if } t+h \leq s \leq T, \end{cases}$$

and let

$$\text{(12)} \qquad \begin{cases} \dot{\mathbf{x}}_3(s) = \mathbf{f}(\mathbf{x}_3(s), \boldsymbol{\alpha}_3(s)) & (t < s < T) \\ \mathbf{x}_3(t) = x. \end{cases}$$

By uniqueness of solutions to the differential equation (1), we have

$$\text{(13)} \qquad \mathbf{x}_3(s) = \begin{cases} \mathbf{x}_1(s) & \text{if } t \leq s \leq t+h \\ \mathbf{x}_2(s) & \text{if } t+h \leq s \leq T. \end{cases}$$

Thus the definition (6) implies

$$\begin{aligned} u(x,t) &\leq C_{x,t}[\boldsymbol{\alpha}_3(\cdot)] \\ &= \int_t^T r(\mathbf{x}_3(s), \boldsymbol{\alpha}_3(s))\, ds + g(\mathbf{x}_3(T)) \\ &= \int_t^{t+h} r(\mathbf{x}_1(s), \boldsymbol{\alpha}_1(s))\, ds + \int_{t+h}^{T} r(\mathbf{x}_2(s), \boldsymbol{\alpha}_2(s))\, ds + g(\mathbf{x}_2(T)) \\ &\leq \int_t^{t+h} r(\mathbf{x}_1(s), \boldsymbol{\alpha}_1(s))\, ds + u(\mathbf{x}_1(t+h), t+h) + \epsilon, \end{aligned}$$

the last inequality resulting from (9). As $\boldsymbol{\alpha}_1(\cdot) \in \mathcal{A}$ was arbitrary, we conclude

$$\text{(14)} \qquad u(x,t) \leq \inf_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + u(\mathbf{x}(t+h), t+h) \right\} + \epsilon,$$

$\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$ solving (1).

2. Fixing again $\epsilon > 0$, select now $\boldsymbol{\alpha}_4(\cdot) \in \mathcal{A}$ so that

$$u(x,t) + \epsilon \geq \int_t^T r(\mathbf{x}_4(s), \boldsymbol{\alpha}_4(s))\, ds + g(\mathbf{x}_4(T)), \tag{15}$$

where

$$\begin{cases} \dot{\mathbf{x}}_4(s) = \mathbf{f}(\mathbf{x}_4(s), \boldsymbol{\alpha}_4(s)) & (t < s < T) \\ \mathbf{x}_4(t) = x. \end{cases}$$

Observe then from (6) that

$$u(\mathbf{x}_4(t+h), t+h) \leq \int_{t+h}^T r(\mathbf{x}_4(s), \boldsymbol{\alpha}_4(s))\, ds + g(\mathbf{x}_4(T)). \tag{16}$$

Therefore

$$u(x,t) + \epsilon \geq \inf_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + u(\mathbf{x}(t+h), t+h) \right\},$$

$\mathbf{x}(\cdot) = \mathbf{x}^{\boldsymbol{\alpha}(\cdot)}(\cdot)$ solving (1). This inequality and (14) complete the proof of (7). □

10.3.3. Hamilton–Jacobi–Bellman equation.

Our eventual goal is writing down as a PDE an "infinitesimal version" of the optimality conditions (7). But first we must check that the value function u is bounded and Lipschitz continuous.

LEMMA (Estimates for value function). *There exists a constant C such that*

$$|u(x,t)| \leq C,$$
$$|u(x,t) - u(\hat{x},\hat{t})| \leq C(|x-\hat{x}| + |t-\hat{t}|)$$

for all $x, \hat{x} \in \mathbb{R}^n$, $0 \leq t, \hat{t} \leq T$.

Proof. 1. Clearly hypothesis (5) implies u is bounded on $\mathbb{R}^n \times [0,T]$.

2. Fix $x, \hat{x} \in \mathbb{R}^n$, $0 \leq t < T$. Let $\epsilon > 0$ and then choose $\hat{\boldsymbol{\alpha}}(\cdot) \in \mathcal{A}$ so that

$$u(\hat{x},t) + \epsilon \geq \int_t^T r(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))\, ds + g(\hat{\mathbf{x}}(T)), \tag{17}$$

where $\hat{\mathbf{x}}(\cdot)$ solves the ODE

$$\begin{cases} \dot{\hat{\mathbf{x}}}(s) = \mathbf{f}(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s)) & (t < s < T) \\ \hat{\mathbf{x}}(t) = \hat{x}. \end{cases} \tag{18}$$

Then

$$
\begin{aligned}
u(x,t) - u(\hat{x},t) \le & \int_t^T r(\mathbf{x}(s), \hat{\boldsymbol{\alpha}}(s))\, ds + g(\mathbf{x}(T)) \\
& - \int_t^T r(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))\, ds - g(\hat{\mathbf{x}}(T)) + \epsilon,
\end{aligned} \tag{19}
$$

where $\mathbf{x}(\cdot)$ solves

$$
\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \hat{\boldsymbol{\alpha}}(s)) & (t < s < T) \\ \mathbf{x}(t) = x. \end{cases} \tag{20}
$$

Since $\mathbf{f}$ is Lipschitz continuous, (18), (20) and Gronwall's inequality (§B.2) imply $|\mathbf{x}(s) - \hat{\mathbf{x}}(s)| \le C|x - \hat{x}|$ $(t \le s \le T)$. Hence we deduce from (5) and (19) that $u(x,t) - u(\hat{x},t) \le C|x - \hat{x}| + \epsilon$. The same argument with the roles of x and $\hat{x}$ reversed implies

$$
|u(x,t) - u(\hat{x},t)| \le C|x - \hat{x}| \quad (x, \hat{x} \in \mathbb{R}^n,\ 0 \le t \le T).
$$

3. Now let $x \in \mathbb{R}^n$, $0 \le t < \hat{t} \le T$. Take $\epsilon > 0$ and choose $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ so that

$$
u(x,t) + \epsilon \ge \int_t^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)),
$$

$\mathbf{x}(\cdot)$ solving the ODE (1). Define

$$
\hat{\boldsymbol{\alpha}}(s) := \boldsymbol{\alpha}(s + t - \hat{t}) \ \text{ for } \hat{t} \le s \le T
$$

and let $\hat{\mathbf{x}}(\cdot)$ solve

$$
\begin{cases} \dot{\hat{\mathbf{x}}}(s) = \mathbf{f}(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s)) & (\hat{t} < s < T) \\ \hat{\mathbf{x}}(\hat{t}) = x. \end{cases}
$$

Then $\hat{\mathbf{x}}(s) = \mathbf{x}(s + t - \hat{t})$. Hence

$$
\begin{aligned}
u(x,\hat{t}) - u(x,t) \le & \int_{\hat{t}}^T r(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))\, ds + g(\hat{\mathbf{x}}(T)) \\
& - \int_t^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds - g(\mathbf{x}(T)) + \epsilon \\
= & - \int_{T+t-\hat{t}}^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T + t - \hat{t})) - g(\mathbf{x}(T)) + \epsilon \\
\le & \ C|t - \hat{t}| + \epsilon.
\end{aligned} \tag{21}
$$

Next pick $\hat{\boldsymbol{\alpha}}(\cdot)$ so that

$$u(x,\hat{t}) + \epsilon \geq \int_{\hat{t}}^{T} r(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s))\, ds + g(\hat{\mathbf{x}}(T)),$$

where

$$\begin{cases} \dot{\hat{\mathbf{x}}}(s) = \mathbf{f}(\hat{\mathbf{x}}(s), \hat{\boldsymbol{\alpha}}(s)) & (\hat{t} < s < T) \\ \hat{\mathbf{x}}(\hat{t}) = x. \end{cases}$$

Define

$$\boldsymbol{\alpha}(s) := \begin{cases} \hat{\boldsymbol{\alpha}}(s + \hat{t} - t) & \text{if } t \leq s \leq T + t - \hat{t} \\ \hat{\boldsymbol{\alpha}}(T) & \text{if } T + t - \hat{t} \leq s \leq T, \end{cases}$$

and let $\mathbf{x}(\cdot)$ solve (1). Then $\boldsymbol{\alpha}(s) = \hat{\boldsymbol{\alpha}}(s + \hat{t} - t)$, $\mathbf{x}(s) = \hat{\mathbf{x}}(s + \hat{t} - t)$ for $t \leq s \leq T + t - \hat{t}$. Consequently

$$\begin{aligned} u(x,t) - u(x,\hat{t}) &\leq \int_{t}^{T} r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)) \\ &\quad - \int_{\hat{t}}^{T} r(\hat{\mathbf{x}}(s), \boldsymbol{\alpha}(s))\, ds - g(\hat{\mathbf{x}}(T)) + \epsilon \\ &= \int_{T+t-\hat{t}}^{T} r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)) - g(\mathbf{x}(T + t - \hat{t})) + \epsilon \\ &\leq C|t - \hat{t}| + \epsilon. \end{aligned}$$

This inequality and (21) prove

$$|u(x,t) - u(x,\hat{t})| \leq C|t - \hat{t}| \quad (0 \leq t \leq \hat{t} \leq T,\ x \in \mathbb{R}^n). \qquad \square$$

We prove next that the value function solves a Hamilton–Jacobi type partial differential equation.

THEOREM 2 (A PDE for the value function). *The value function u is the unique viscosity solution of this terminal-value problem for the* Hamilton–Jacobi–Bellman equation*:*

$$(22) \qquad \begin{cases} u_t + \min_{a \in A} \{\mathbf{f}(x,a) \cdot Du + r(x,a)\} = 0 & \textit{in } \mathbb{R}^n \times (0,T) \\ u = g & \textit{on } \mathbb{R}^n \times \{t = T\}. \end{cases}$$

Remarks. (i) The Hamilton–Jacobi–Bellman PDE has the form

$$u_t + H(Du, x) = 0 \quad \text{in } \mathbb{R}^n \times (0,T),$$

for the Hamiltonian

$$(23) \qquad H(p,x) := \min_{a \in A} \{\mathbf{f}(x,a) \cdot p + r(x,a)\} \quad (p, x \in \mathbb{R}^n).$$

From the inequalities (5), we deduce that H satisfies the estimates (3) in §10.2.

(ii) Since (22) is a *terminal-value problem*, we must specify what we mean by a solution. Let us say that a bounded, uniformly continuous function u is a *viscosity solution* of (22) provided

(a) $u = g$ on $\mathbb{R}^n \times \{t = T\}$,

and

(b) for each $v \in C^\infty(\mathbb{R}^n \times (0,T))$

$$\left\{\begin{array}{l}\text{if } u - v \text{ has a local maximum at a point } (x_0, t_0) \in \mathbb{R}^n \times (0,T),\\ \text{then}\\ \qquad v_t(x_0,t_0) + H(Dv(x_0,t_0),x_0) \geq 0,\end{array}\right. \tag{24}$$

and

$$\left\{\begin{array}{l}\text{if } u - v \text{ has a local minimum at a point } (x_0, t_0) \in \mathbb{R}^n \times (0,T),\\ \text{then}\\ \qquad v_t(x_0,t_0) + H(Dv(x_0,t_0),x_0) \leq 0.\end{array}\right. \tag{25}$$

Observe that for our terminal-value problem (22) we *reverse* the sense of the inequalities from those for the initial-value problem.

(iii) The reader should check that if u is the viscosity solution of (22), then $w(x,t) := u(x, T-t)$ ($x \in \mathbb{R}^n, 0 \leq t \leq T$) is the viscosity solution of the initial-value problem

$$\left\{\begin{aligned} w_t - H(Dw, x) &= 0 \quad \text{in } \mathbb{R}^n \times (0,T)\\ w &= g \quad \text{on } \mathbb{R}^n \times \{t = 0\}.\end{aligned}\right.$$

Proof. 1. In view of the lemma, u is bounded and Lipschitz continuous. In addition, we see directly from (4) and (6) that

$$u(x,T) = \inf_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} C_{x,T}[\boldsymbol{\alpha}(\cdot)] = g(x) \quad (x \in \mathbb{R}^n).$$

2. Now let $v \in C^\infty(\mathbb{R}^n \times (0,T))$, and assume

$$u - v \text{ has a local maximum at a point } (x_0,t_0) \in \mathbb{R}^n \times (0,T).$$

We must prove

$$v_t(x_0,t_0) + \min_{a\in A}\{\mathbf{f}(x_0,a)\cdot Dv(x_0,t_0) + r(x_0,a)\} \geq 0. \tag{26}$$