WENDELL H. FLEMING  
PANAGIOTIS E. SOUGANIDIS  
PDE-viscosity solution approach to some problems of large deviations 


<http://www.numdam.org/item?id=ASNSP_1986_4_13_2_171_0>
0. - Introduction.

The theory of nonlinear, first order PDE of Hamilton-Jacobi type has been substantially developed with the introduction by M. G. Crandall and P.-L Lions [3] of the class of viscosity solutions. This turns out to be the correct class of generalized solutions for such equations. M. G. Crandall, L. C. Evans and P.-L. Lions [2] provide a simpler introduction to the subject while the book by P.-L. Lions [16] and the paper by M. G. Crandall and P. E. Souganidis [5] provide a view of the scope of the theory and references to much of the recent literature.

Recently, L. C. Evans and H. Ishii [9] illustrated the usefulness of the viscosity solution methods in studying various asymptotic problems concerning stochastic differential equations with small noise intensities. They gave new proofs based on PDE-viscosity solution methods for results of Ventcel-Freidlin type which were previously treated by quite different (probabilistic and stochastic control) techniques. In the present work we use a similar PDE-viscosity solution method to give a new proof and extend a result of W. H. Fleming and C.-P. Tsai [15] concerning optimal exit probabilities and differential games. The problem considered is to control the drift of a Markov diffusion process in such a way that the probability that the process exits from a given region $D$ during a given finite time interval is minimum. An asymptotic formula for the minimum exit probability when the process is nearly deterministic is given. This

(*) This research was partially supported by NSF under Grant No. MCS 8121940, by ONR under Grant No. N00014-83-K-0542 and by AFOSR under Grant No. AF-AFOSR 81-0116.

(**) This research was partially supported by ONR under Grant No. N00014-83-K-0542 and by NSF under Grant No. DMS-8401725.

formula involves a crucial quantity $I$, which turns out to be the lower value of an associated differential game.

More precisely, following [15] let $\xi(\cdot)$ be an $N$-dimensional stochastic process with continuous sample paths defined for times $t > s$. Let $D \subseteq \mathbb{R}^N$ be open and bounded with smooth boundary $\partial D$. For initial time $s$ and state $x \in D$, let $\tau_{x,s}$ denote the exit time from $D$ (i.e., the first $t$ such that $\xi(t) \in \partial D$). For fixed $T > 0$, $P(\tau_{x,s} < T)$ is the exit probability.

We assume that $\xi(\cdot)$ is a controlled Markov diffusion process satisfying in the Itô-sense the stochastic differential equation

$$d\xi(t) = b(\xi(t), y(t)) dt + \varepsilon \sigma(\xi(t)) dw(t),$$

where $y(t)$ is a control applied at time $t$, $\varepsilon > 0$ is a parameter, $\sigma$ is an $N \times N$ matrix and $w(\cdot)$ is an $N$-dimensional Brownian motion. In [15], it was assumed that $\sigma = \text{identity matrix}$. We assume that $y(t) \in Y$, where $Y \subseteq \mathbb{R}^N$ is compact. Moreover, the control processes $y(\cdot)$ admitted in (0.1) have the feedback form

$$y(t) = y(t, \xi(t))$$

where $y : [s, T] \times \mathbb{R}^N \to Y$ is Borel measurable. As far as $b$ and $\sigma$ are concerned we assume that $b(\cdot, \cdot)$ is bounded Lipschitz continuous, $\sigma(\cdot)$ is of class $C^2$ and bounded together with its derivatives, and there exists $\theta > 0$ such that

$$a_c(x) \xi, \xi, \geq \theta |\xi|^2$$

for all $x, \xi \in \mathbb{R}^N$

where $a = \sigma\sigma'$. Let us note here that we have made no attempt to discover minimal assumptions; in particular, it will be clear later that by using obvious approximations we really need only assume $\sigma$ to be continuously differentiable with bounded derivatives.

Let

$$q^s_x(x, s) = P(\tau_{x,s} < T).$$

Of course, the exit probability depends on $\varepsilon$ and $y$ in view of (0.1), namely $\tau_{x,s} = \tau_{x,s}^y$. The minimum exit probability is

$$q^y(x, s) = \min_y q^s_x(x, s).$$

The function $q^y(x, s)$ satisfies the dynamic programming equation (for details see [15])

$$q^y_t + \frac{\varepsilon}{2} \text{tr} a(x) D^2 q^y + \min_{y \in Y} \{b(x, y) \cdot Dq^y \} = 0 \quad \text{in } [0, T) \times D,$$
where $Dq^e$ denotes the gradient vector and $\text{tr} a(x) D^2 q^e = \sum_{ij} a_{ij} q_x x_i x_j$. The boundary conditions are

$$
\begin{align*}
q^e(x, s) &= 1 \quad \text{for } s < T, \ x \in \partial D \\
q^e(x, T) &= 0 \quad \text{for } x \in \bar{D}.
\end{align*}
$$

In general, it is difficult to get effective information about $q^e$ and the optimal control in this way. Instead, we seek an asymptotic formula for $q^e$, valid for small $\varepsilon > 0$, of the form

$$
- \lim_{\varepsilon \to 0} \varepsilon \log q^e = I,
$$

where $I$ turns out to be the lower value of a certain differential game. Equation (0.6) can be written as

$$
q^e = \exp \left\{ - \frac{I + o(1)}{\varepsilon} \right\},
$$

which is a weaker result than a WKB expansion

$$
q^e = \exp \left\{ - \frac{I}{\varepsilon} \right\} \text{asymptotic series in powers of } \varepsilon.
$$

The expansion (0.7) cannot be expected to hold except in certain regions where $I(x, s)$ is a smooth function of $(x, s)$. We have no results concerning (0.7), although an expansion up to terms involving $\varepsilon, \varepsilon^2$ for a similar problem arising in stochastic control was given in [12, Sec. 6] and [14].

The differential game, which has $I$ as its lower value, is formally described as follows. (For more details and motivation see [15].) There are two players, a maximizing player who chooses $y(t) \in Y$ and a minimizing player who chooses $z(t) \in \mathbb{R}$. The state $x(t)$ of the game at time $t$ satisfies

$$
x(t) = x + \int_{s}^{t} z(r) \, dr.
$$

Let $\tau_x$ denote the exit time of $x(t)$ from $D$, and $\tau_x \wedge T = \min (\tau_x, T)$. Let

$$
\begin{align*}
L(x, y, z) &= \frac{1}{2} (b(x, y) - z)^T a(x)^{-1} (b(x, y) - z) \\
\chi(x) &= \begin{cases} 0, & x \in \partial D \\
+ \infty, & x \in \bar{D} \end{cases}
\end{align*}
$$

and

$$
(0.9)
$$
The game payoff is

\[ \int_{s}^{T \wedge \tau_z} L(x(t), y(t), z(t)) \, dt + \xi(\tau(T \wedge \tau_z)) . \]

We consider the lower game in which (formally speaking) the minimizing player has the information advantage of knowing both \( y(t) \) and \( x(t) \) before \( z(t) \) is chosen, while his opponent knows only \( x(t) \) before choosing \( y(t) \). This formal description can be made precise in one of several possible ways, each of which involves concepts of game strategy. The Elliott-Kalton formulation, which is made precise in Section 1 ([6], [7], [10]), is convenient here.

Let \( I = I(x, s) \) denote the lower value of the game, in the Elliott-Kalton sense. Let

\[ H(x, p) = \max_{y \in Y} \min_{z \in \mathbb{R}^n} [L(x, y, z) + p \cdot z] . \]

The Isaacs or dynamic programming equation associated with this lower game is

\[ I_x + H(x, DI) = 0 \quad \text{in } D \times [0, T) . \]

The main result is:

**Theorem.** (a) \( I(x, s) \) is the unique viscosity solution of (0.12) in \( D \times [0, T) \) with the boundary conditions

\[
\begin{cases}
I(x, s) = 0 & \text{for } x \in \partial D, \ 0 < s < T \\
I(x, s) \uparrow + \infty & \text{as } s \uparrow T, \ \text{for } x \in D .
\end{cases}
\]

(b) Let \( I^\varepsilon = -\varepsilon \log q^\varepsilon \). Then:

\[ \lim_{\varepsilon \to 0} I^\varepsilon = I . \]

As candidates for viscosity solution of (0.12) we admit functions \( I \in C^{0,1}(\bar{D} \times [0, T')) \), \( \forall T' < T \), where \( C^{0,1}(\bar{D} \times [0, T']) \) is the space of continuous functions defined on \( \bar{D} \times [0, T'] \), which are Lipschitz continuous with respect to \( x \).

We continue with the basic plan for the proof of the theorem, thus explaining the previously mentioned PDE-viscosity solution method. The
fact that $I$ is a viscosity solution follows by standard arguments concerning
the relation between the Elliot-Kalton formulation of the value of dif-
ferential games and viscosity solutions (L. C. Evans and P. E. Souga-
 nidis [10]) as well as some rescaling similar to the one used by W. H.
Fleming and C.-P. Tsai [15]. The uniqueness part was pointed out to us
by P.-L. Lions and a simpler argument was given by M.G. Crandall. (For
more details see M.G. Grandall, P.-L. Lions and P. E. Songanidis [4].)
For part (b) following L. C. Evans and H. Ishii [9] we obtain estimates,
independent of $\epsilon$, on $I^\epsilon$ and its first derivatives, using the fact that $I^\epsilon$ sat-
fies the equation

$$I^\epsilon_t + \frac{\epsilon}{2} \text{tr}(a(x))D^2I^\epsilon + H(x, DI^\epsilon) = 0 \quad \text{in } [0, T) \times D$$

with boundary conditions (0.13). These estimates together with part (a)
and some standard results concerning viscosity solutions conclude the proof.

There are several advantages of this method over the original proof
of [15]. Principally, we considerably simplify the proof of [15], which
involves differential game theoretic and probabilistic arguments as well
as several limit arguments. As a result of this simplification, we are able
to extend the results of [15]. For a summary of the above, as well as some
other results concerning the use of PDE-viscosity solution methods in
stochastic control we refer to the expository paper by W. H. Fleming [13]

The paper is organized as follows: Section 1 deals with part (a) of the
theorem. Section 2 is devoted to part (b). Finally, in the Appendix we
reproduce in a simple case the uniqueness result concerning viscosity solu-
tions of (0.12), (0.13).

ACKNOWLEDGEMENT. We would like to thank M. G. Crandall and P.-L.
Lions for several useful discussions.

1. - We begin this section by reviewing and modifying, when necessary,
some basic definitions and concepts concerning the lower value of a dif-
ferential game. We employ here mostly the notation of Elliott-Kalton
[6], [7] (cf. also [10], [17], [18]) and the introduction.

For $s \in [0, T]$ define

$$M(s, T) = \{y \colon [s, T] \to Y \colon y(\cdot) \text{ measurable} \} ;$$

$$N(s, T) = \{z \colon [s, T] \to \mathbb{R}^n \colon z(\cdot) \text{ measurable} \} ;$$

we will hereafter identify any two functions which agree a.e. Any mapping
$\beta \colon N(s, T) \to M(s, T)$ is called a strategy for the minimizing player provided
for each \( r \in (s, T) \),

\[ y(t) = \hat{y}(t) \quad \text{for a.e. } s < t < r \]

implies

\[ \beta(y)(t) = \beta(\hat{y})(t) \quad \text{for a.e. } s < t < r. \]

Denote by \( A(s, T) \) the set of strategies for the minimizing player.

Next, for \( (x, s) \in \bar{D} \times [0, T] \), we define

\[
I(x, s) = \inf_{\beta \in A(x, s)} \sup_{r \in M(x, s)} \left\{ \int_{s}^{T} L(x(t), y(t), \beta(y)(t)) \, dt + \chi(x(t_x \land T)) \right\},
\]

with \( x(\cdot) \) solving (0.8), \( z(\cdot) = \beta(y)(\cdot) \) and \( L, \chi \) as in (0.9). We call \( I \) the lower value of the differential game. Sometimes we write \( I = I(x, s; T) \) to emphasize its dependence on \( T \). Since no upper bound is imposed on control functions \( z(\cdot) \in N(s, T) \), it is easy to see that \( I(x, s) < \infty \) for all \( (x, s) \in \bar{D} \times [0, T] \) despite the penalty function \( \chi \) in the payoff. Moreover, it is immediate that \( I(x, s) = 0 \) for every \( (x, s) \in \partial \bar{D} \times [0, T] \).

Perhaps the most important property of \( I \) is that it satisfies the dynamic programming optimality condition. In particular, we have:

**Theorem 1.1.** For \( (x, s) \in \bar{D} \times [0, T] \) and \( T - s > \sigma > 0 \)

\[
I(x, s) = \inf_{\beta \in A(x, s)} \sup_{r \in M(x, s)} \left\{ \int_{s}^{(s + \sigma) \land T} L(x(t), y(t), \beta(y)(t)) \, dt + I(x((s + \sigma) \land T), (s + \sigma) \land T) \right\}.
\]

The dynamic programming optimality condition can be formulated more generally using the notion of nonanticipating functionals. Since we do not need this general form here, we do not state it. (1.2) is proved as in L. C. Evans and P. E. Souganidis [10] with some minor changes (cf. also [8], [17], [18]). To illustrate that the presence of \( \chi \) in the payoff does not cause any major difficulties we include the proof. The notations are similar to the ones in [10]. Finally, for simplicity we drop the dependence of \( A(s, T) \) and \( M(s, T) \) on \( T \).

**Proof of Theorem 1.1.** Set

\[
W(x, s) = \inf_{\beta \in A(x, s)} \sup_{r \in M(x, s)} \left\{ \int_{s}^{(s + \sigma) \land T} L(x(t), y(t), \beta(y)(t)) \, dt + I(x((s + \sigma) \land T), (s + \sigma) \land T) \right\},
\]

for each \( r \in (s, T) \),

\[ y(t) = \hat{y}(t) \quad \text{for a.e. } s < t < r \]
and fix $\gamma > 0$. Then there exists $\delta \in \Delta(s)$ such that

$$W(x, s) \geq \sup_{\forall \in M(s)} \left\{ \int_{s}^{(s + \sigma) \wedge \tau_x} L(x(t), y(t), \delta[y](t)) dt + \chi(x((s + \sigma) \wedge \tau_x), (s + \sigma) \wedge \tau_x) \right\} - \gamma.$$ 

Also, for each $\forall \in \overline{D}$

$$I(\forall, s + \sigma) = \inf_{\beta \in \Delta(s + \sigma)} \sup_{\forall \in M(s + \sigma)} \left\{ \int_{s + \sigma}^{T \wedge \tau_x} L(x(t), y(t), \delta[y](t)) dt + \chi(x(T \wedge \tau_x)) \right\},$$

where $x(\cdot)$ solves (0.8) on $(s + \sigma, T)$ with the initial condition $x(s + \sigma) = \forall$ and $\tau_x$ is its exit time from $D$. Thus there exists $\delta_x \in \Delta(s + \sigma)$ for which

$$I(\forall, s + \sigma) \geq \sup_{\forall \in M(s + \sigma)} \left\{ \int_{s + \sigma}^{T \wedge \tau_x} L(x(t), y(t), \delta_x[y](t)) dt + \chi(x(T \wedge \tau_x)) \right\} - \gamma.$$ 

Define $\beta \in \Delta(s)$ as follows: for each $\forall \in M(s)$ solve (0.8) on $(s, T)$ for $x(s) = x$ and $z(\cdot) = \delta[y](\cdot)$ and compute $\tau_x$. Set

$$\beta[y](t) = \begin{cases} 
\delta[y](t) & \text{for } s < t < (s + \sigma) \wedge \tau_x \\
0 & \text{for } \tau_x < t < T, \quad \text{if } \tau_x < s + \sigma \\
\delta_x(s + \sigma)[y](t) & \text{for } s + \sigma < t < T, \quad \text{if } s + \sigma < \tau_x.
\end{cases}$$

It is easy to check and we leave it to the reader that $\beta$ is a strategy. Consequently for any $\forall \in M(s)$, (1.4) and (1.5) imply

$$W(x, s) \geq \int_{s}^{T \wedge \tau_x} L(x(t), y(t), \beta[y](t)) dt + \chi(x(T \wedge \tau_x)) - 2\gamma;$$

so that

$$\sup_{\forall \in M(s)} \left\{ \int_{s}^{T \wedge \tau_x} L(x(t), y(t), \beta[y](t)) dt + \chi(x(T \wedge \tau_x)) \right\} < W(x, s) + 2\gamma.$$ 

Hence

$$I(\forall, s) < W(x, s) + 2\gamma.$$
On the other hand there exists \( \beta \in \Lambda(s) \) for which

\[
I(x, s) > \sup_{y \in M(s)} \left\{ \int_2^{T \wedge \tau_x} L(x(t), y(t), \beta(y)(t)) \, dt + \chi(x(T \wedge \tau_x)) \right\} - \gamma.
\]

By (1.3)

\[
W(x, s) \leq \sup_{y \in M(s)} \left\{ \int_2^{(s + \sigma) \wedge \tau_x} L(x(t), y(t), \beta(y)(t)) \, dt + I\left(x((s + \sigma) \wedge \tau_x), (s + \sigma) \wedge \tau_x\right) \right\},
\]

and consequently there exists \( y_1 \in M(s) \) such that

\[
W(x, s) \leq \int_2^{(s + \sigma) \wedge \tau_x} L(x(t), y_1(t), \beta(y_1)(t)) \, dt + I\left(x((s + \sigma) \wedge \tau_x), (s + \sigma) \wedge \tau_x\right) + \gamma.
\]

If \( \tau_x \leq s + \sigma \), in view of (1.7), there is nothing to prove. If \( \tau_x > s + \sigma \) for each \( y \in M(s + \sigma) \) define \( \bar{y} \in M(s) \) and \( \bar{\beta} \in \Lambda(t + \sigma) \) by

\[
\bar{y}(t) = \begin{cases} y_1(t) & \text{for } s < t < s + \sigma \\ y(t) & \text{for } s + \sigma < t < T \end{cases},
\]

\[
\bar{\beta}(y)(t) = \beta(\bar{y})(t) & \text{for } s + \sigma < t < T.
\]

Now

\[
I(x(s + \sigma), s + \sigma) \leq \sup_{y \in M(s + \sigma)} \left\{ \int_2^{(s + \sigma) \wedge \tau_x} L(x(t), y(t), \beta(y)(t)) \, dt + \chi(x(T \wedge \tau_x(s + \sigma))) \right\}
\]

and so there exists \( y_2 \in M(s + \sigma) \) for which

\[
I(x(s + \sigma), s + \sigma) \leq \int_2^{(s + \sigma) \wedge \tau_x} L(x(t), y_2(t), \beta(y_2)(t)) \, dt + \chi(x(T \wedge \tau_x(s + \sigma))) + \gamma.
\]

Define \( y \in M(s) \) by

\[
y(t) = \begin{cases} y_1(t) & \text{for } s < t < s + \sigma \\ y_2(t) & \text{for } s + \sigma < s < T. \end{cases}
\]

Then (1.7) and the definition of \( \bar{\beta} \) imply that \( \tau_x(s + \sigma) < T \) in (1.9). Moreover,
(1.8) and (1.9) yield
\[ W(x, s) < \int_0^{\tau_s} L(x(t), y(t), \beta(y)(t)) \, dt + \chi(x(T \wedge \tau_s)) + 2\gamma, \]
and so (1.7) implies
\[ W(x, s) < I(x, s) + 3\gamma. \]
This and (1.6) complete the proof.

Next we want to establish the joint in \((x, s)\) continuity of the value function \(I\). This will follow from the following proposition concerning the behavior of \(I\) under rescaling (cf. [15], Lemma 4.1). We have:

**Lemma 1.1.** Let \(\frac{1}{2}T \leq T' < T\). There exists a constant \(R_1\) (depending on \(T\)) such that
\[ I(x, 0; T) = \inf_{\beta \in \Delta(0, T)} \sup_{y \in M(0, T)} \int_0^{\tau_s} L(x(t), y(t), \beta(y)(t)) \, dt, \]
where \(\Delta(0, T)\) consists of those \(\beta \in \Delta(0, T)\) such that \(\tau_s < T\) for all \(y \in M(0, T)\). From this it is easy to see that \(I(x, 0; T)\) is a nonincreasing function of \(T\), which is the left hand inequality in (1.11).

We obtain the right hand inequality as follows. Let \(\alpha = T'/T - 1\). For any \(\beta \in \Delta(0, T)\) and \(y \in M(0, T)\) we consider \(\beta' \in \Delta(0, T')\) and \(y' \in M(0, T')\) given by
\[ y'(t) = y(x^{-1}t'), \quad \beta'[y'](t') = x^{-1}\beta[y](x^{-1}t'), \quad \text{for every } t' \in [0, T']. \]
It is immediate that
\[ x(t) = x'(t') \quad \text{for } t' = \alpha t \quad \text{and } \quad t \in [0, T] \]
where \(x(\cdot), x'(\cdot)\) solve (0.8) with the same \(x, s = 0\) and \(z = \beta[y], \beta'[y']\) respectively. We want to estimate the difference of their respective payoffs \(\pi\) and \(\pi'\). If \(x(t) \in D\) for \(t \in [0, T]\), then obviously
\[ \pi = \pi' = +\infty. \]
Hence we may assume that
\[ \tau_s = a^{-1} \tau_s' < T = a^{-1}T' . \]
By (0.10) we have
\[ \pi' - \pi = \int_0^{\tau_s} \left\{ xL(x(t), y(t), x^{-1}\beta[t](t)) - L(x(t), y(t), \beta[t](t)) \right\} dt . \]
By (0.9) and the fact that \(-1 \leq L \leq 1\),
\[ L(x, y, x^{-1}z) - L(x, y, z) = \frac{x^{-1} - 1}{2} b' a^{-1} b + \frac{x^{-1} - 1}{2} z' a^{-1} z < (1 - \alpha) A |z|^2 \]
where \(|a^{-1}(x)| < A\). On the other hand, since
\[ L > C_1 |z|^2 - C_2 \]
for some constants \(C_1, C_2 > 0\), we have
\[ \pi' < \pi + A(1 - \alpha) \int_0^{\tau_s} |\beta[t](t)|^2 dt < \pi + \frac{A(1 - \alpha)}{C_1} (\pi + C_2 T) \]
and therefore
\[ I(0, x; T') < I(0, x; T) + \frac{T - T'}{C_1 T} A(I(0, x; T) + C_2 T) . \]
To conclude the proof we need to show that
\[ I(0, x; T < K \]
for some constant \(K\) which depends only on \(T\). To this end, choose \(z_0 \in \mathbb{R}^n\) such that
\[ |z_0| T > \text{diam } D \]
and define \(\beta \in A(0, T)\) by
\[ \beta[y] = z_0 \quad \text{for every } y \in M(0, T) . \]
Then, \(\tau_s < T\) for every \(y \in M(0, T)\). Moreover, the payoff \(\pi\) satisfies \(\pi < TL^*\), where \(L^*\) is a bound of \(L(x, y, z)\) for \(|z| < |z_0|\). Let
\[ K = TL^* , \quad R_1 = \frac{A(K}{T + C_2} . \]
We then have
\begin{equation}
I(x, 0; T') < I(x, 0; T) + R_4 (T - T') ,
\end{equation}
which proves Lemma 1.1.

As a consequence of this lemma, we prove the following result concerning the continuity of \( I \) with respect to \( s \). We have:

**Proposition 1.1.** For \( x \in \overline{D} \) and \( s \in [0, T) \), \( s \mapsto I(x, s) \) is continuous.

**Proof.** The result is an immediate consequence of Lemma 1.1 given that
\[ I(x, s; T) = I(x, 0; T - s) \]
where the last equality follows easily from the definition of \( I \).

Next we prove the Lipschitz continuity with respect to \( x \) of the value function \( I \). As a preliminary step we need the following lemma.

**Lemma 1.2.** Let \( x, \hat{x} \in D \). There exists a constant \( R_2 \) which depends only on \( T \) such that
\begin{equation}
I(\hat{x}, 0; T + |x - \hat{x}|) < I(x, 0; T) + R_4 |x - \hat{x}|.
\end{equation}

**Proof.** For any \( x' \in \overline{D} \) such that
\[
\text{dist}(x', \partial D) < |x - \hat{x}|
\]
let \( z(x') \in \mathbb{R}^N \) be such that
\[ 1 < |z(x')| < 2 \quad \text{and} \quad x' + |x - \hat{x}|z(x') \in D. \]

For \( \gamma > 0 \) fixed there exists \( \delta \in \mathcal{A}(0, T) \) such that
\[
I(x, 0; T) > \int_0^{\tau_x} L(x(t), y(t), \delta[y](t)) \, dt - \gamma \quad \text{for every } y \in M(0, T)
\]
and
\[ \tau_x < T \quad \text{for every } y \in M(0, T). \]

Next, define \( \hat{\beta} : M(0, T + |x - \hat{x}|) \to N(0, T + |x - \hat{x}|) \) as follows: For \( y \in M(0, T + |x - \hat{x}|) \) let \( \hat{y} \in M(0, T) \) be such that
\[ y = \hat{y}|_{[0, \tau_1]}. \]
If $\hat{x}(\cdot)$ is the solution of (0.8) for $s = 0$, $z(\cdot) = \delta[y](\cdot)$ and $\hat{x}(0) = \hat{x}$, then

$$
\hat{y}(t) = \begin{cases} 
\delta[y](t), & 0 < t < \tau_x \\
z(\hat{x}(\tau_x)), & \tau_x < t < T + |x - \hat{x}|, 
\end{cases}
$$

and

$$
x(t) - \hat{x}(t) = x - \hat{x}, \quad 0 < t < \tau_x.
$$

It is easy to check that $\hat{y} \in A(0, T + |x - \hat{x}|)$ and we leave it up to the reader. In view of the above definition, it is obvious that, for every $\hat{y} \in M(0, T + |x - \hat{x}|)$,

$$
\tau_x < T + |x - \hat{x}|.
$$

We then have

$$
\int_0^{\tau_x} [L(x(t), y(t), \hat{y}(t)) \delta[y](t)] dt < I(x, 0; T) + \gamma + L_1^a|x - \hat{x}| + \Gamma,
$$

where $|L(x, y, z)| < L_1^a$ if $|z| < 2$ and

$$
\Gamma = \int_0^{\tau_x} |L(x(t), y(t), \delta[y](t)) - L(\hat{x}(t), y(t), \delta[y](t))| dt.
$$

From (0.9) and the fact that $a^{-1}, b$ are bounded Lipschitz

$$
|L(x + h, y, z) - L(x, y, z)| < (k_1|x|^2 + k_2)|h|.
$$

Since

$$
|x|^2 < C_1^{-1}(L + C_2)
$$

by taking $h = x - \hat{x}$ we obtain

$$
\Gamma \leq (b_1 I(x, 0; T) + b_2)|x - \hat{x}| < (b_1 K + b_2)|x - \hat{x}|
$$

for suitable $b_1, b_2$, where $K$ depends on $T$. We take $R_2 = L_1^a + b_1 K + b_2$.

We are now ready to prove the following proposition.

**Proposition 1.2.** For $\frac{1}{2} T < T' < T$ there exists a constant $R_3$ which depends on $T'$ such that

$$
(1.14) \quad |I(x, s; T) - I(\hat{x}, s; T)| < R_3|x - \hat{x}| \quad \text{for every } x, \hat{x} \in D \text{ and } s \in [0, T'].
$$
Proof. We have by Lemmas 1.1, 1.2

\[ I(ζ, 0; T) < I(ζ, 0; T + |x - x|) + R_3 |x - x| < I(x, 0; T) + (R_1 + R_2) |x - x|. \]

The same inequality holds if \( x \) and \( x \) are exchanged. We take \( R_3 = R_1 + R_2 \).

The next result is concerned with the limit of \( I(x, s; T) \) as \( s \uparrow T \). In particular, we have

**Proposition 1.3.** For \( x \in D \), \( \lim_{s \uparrow T} I(x, s; T) = + \infty \).

**Proof.** Let \( γ > 0 \) be fixed. There exists \( δ \in A(s, T) \) such that

\[ I(x, s; T) \geq \int_s^{τ_s} L(x(t), y(t), δ[y][t]) \, dt - γ \quad \text{for every } y \in M(s, T). \]

This implies

\[ I(x, s; T) \geq C_1 \int_0^{τ_τ} |δ[y][t]|^2 \, dt - C_2 - γ, \quad \text{with } C_1 > 0, \]

But

\[ \text{dist}(x, \partial D) \leq |x(τ_s) - x| \leq \int_s^{τ_s} |δ[y][t]| \, dt \leq (τ_s - s)^\dagger \left( \int_s^{τ_s} |δ[y][t]|^2 \, dt \right)^\dagger. \]

Combining the above inequalities we obtain

\[ I(x, s; T) \geq C_1 \frac{\text{dist}^2(x, D)}{τ_s - s} - C_2 - γ \geq C_1 \frac{\text{dist}^2(x, D)}{T - s} - C_2 - γ \]

which finishes the proof.

We are now ready to prove part (a) of the Theorem. We have:

**Proof of part (a).** The fact that \( I \) belongs to the class of functions and satisfies the boundary conditions required by the statement of the theorem follows from the previous propositions. Moreover, Theorem 1.1 and the proof of Theorem 4.1 of [10] imply that \( I \) is a viscosity solution of

\[ I_s + H(x, Dl) = 0 \quad \text{in } D \times [0, T). \]

Finally, as explained in the Introduction, the uniqueness is a part of general results of [4] and it is illustrated in a simple case in the Appendix.
2. — We begin this section with some observations concerning the minimum exit probability \( q^e(x, s) \) and \( I^e(x, s) = -\varepsilon \log q^e(x, s) \). In particular, \( q^e \in C^{\beta}_{p,A}(\mathcal{D}) \) for any compact \( A \subset \bar{D} \times [0, T) \), where \( C^{\beta}_{p,A}(\mathcal{D}) \) is the space of functions \( f \) defined on \( \mathcal{D} \) such that \( f, f_t, f_{x_i}, f_{x_ix_j}, i, j = 1, ..., N \) are Hölder continuous with exponent \( \beta \). Since \( q^e \) satisfies (0.4), (0.5) an elementary calculation shows that \( I^e \) satisfies the nonlinear parabolic PDE

\[
I^e_t + \frac{\varepsilon}{2} a_{ij}(x) I^e_{x_ix_j} - \frac{1}{2} a_{ij} I^e_{x_i} I^e_{x_j} \max_{y \in Y} [b(x,y) \cdot D I^e] = 0
\]

with boundary conditions

\[
\begin{aligned}
I^e(x, s) &= 0 \quad \text{for } x \in \partial D \text{ and } s < T \\
I^e(x, T) &= \infty \quad \text{for } x \in D.
\end{aligned}
\]

Moreover, it is an easy exercise to write (2.1) in the form

\[
I^e_t + \frac{\varepsilon}{2} a_{ij}(x) I^e_{x_ix_j} - \frac{1}{2} a_{ij} I^e_{x_i} I^e_{x_j} + H(x, D I^e) = 0
\]

where \( H \) is given by (0.11).

Next, we establish estimates, independent of \( \varepsilon \), on \( I^e \) and its derivatives. We begin with an \( L^n \)-estimate for \( s < T \).

**Lemma 2.1.** There exist positive constants \( a, b \) such that

\[
I^e(x, s) \leq \frac{a(1 + b(T - s))^2}{T - s} \quad \text{for every } (x, s) \in \bar{D} \times [0, T).
\]

**Proof.** — Without any loss of generality we may assume that \( D \) lies in the slab, \( 0 < x_1 < C \) for some \( C > 0 \). Let \( v : \bar{D} \times [0, T) \mapsto \mathbb{R} \) be defined by

\[
v(x, s) = \frac{a(1 + b(T - s))^2}{T - s} x_1
\]

where \( a, b \) are some positive constants to be determined so that \( v \) is a supersolution of (2.1). Since \( v(x, s) > I^e(x, s) \) for \( (x, s) \in \partial D \times [0, T) \), simple maximum principle considerations will imply (2.4). We need

\[
v_t + \frac{\varepsilon}{2} a_{ij} v_{x_i x_j} - \frac{1}{2} a_{ij} v_{x_i} v_{x_j} \max_{y \in Y} [b(x,y) \cdot Dv] < 0.
\]
But

\[
\begin{align*}
    v_s + \varepsilon \frac{a \nu}{2} a_y v_{x_y} - \frac{\nu}{2} a_y v_{x_y} + \max_{y \in \mathbb{R}} [b(x, y) \cdot Dv] \\
    &< \frac{a(1 + b(T - s))^2}{(T - s)^2} x_1 - \frac{2ab(1 + b(T - s))}{T - s} x_1 - \frac{1}{2} \frac{a^2(1 + b(T - s))^4}{(T - s)^2} \\
    &+ B \frac{a(1 + b(T - s))^2}{T - s}
\end{align*}
\]

where \( \theta \) is given by (0.2) and \( B \) is such that

\[
|b(x, y)| < B.
\]

A simple computation shows that if \( a^2 \geq 2C \) and \( 3ab\theta > 2B \) then (2.5) is valid, thus the result.

The next result gives more precise information about the upper bound of \( I^x \) and, moreover, characterizes the way that \( I^x \) assumes its value as \( x \) approaches \( \partial D \). We have:

**Lemma 2.2.** There exist positive constants \( \gamma, \delta > 0 \) such that

\[
I^x(x, s) \leq \gamma \frac{(1 + \delta(T - s))^2}{T - s} \text{dist}(x, \partial D)
\]

for every \( (x, s) \in \overline{D} \times [0, T) \) and for \( 0 < s < 2 \).

**Proof.** By Lemma 2.1 it suffices to verify such an estimate in some neighborhood of \( \partial D \). Moreover, it is well known that \( d \in C^2(\Gamma) \) (cf. J. Serrin [19]) where \( \Gamma = \{x \in D: \text{dist}(x, \partial D) < \varrho\} \) for some appropriate \( \varrho > 0 \). Let \( u: \overline{D} \times [0, T) \rightarrow \mathbb{R} \) be defined by

\[
u(x, s) = \gamma \frac{(1 + \delta(T - s))^2}{T - s} \text{dist}(x, \partial D)
\]

where \( \gamma, \delta \) are to be determined so that

\[
I^x(x, s) \leq u(x, s) \quad \text{for every } (x, s) \in \{x \in D: \text{dist}(x, \partial D) = \varrho\} \times [0, T)
\]

and \( u \) is a supersolution of (2.1) in \( \Gamma \times [0, T) \). We only indicate now how to satisfy the second requirement, since the first one follows immediately from Lemma 2.1.
To this end, observe that for $\varepsilon < 2$, we have

\[
u_s + \frac{\varepsilon}{2} a_{ij} u_{x_i x_j} - \frac{1}{2} a_{ij} u_x u_{x_j} + \max_{y \in \gamma} (b(x, y) \cdot D u)
\leq \left( \frac{\gamma(1 + \delta(T-s))^2}{(T-s)^2} - \frac{2\gamma \delta(1 + \delta(T-s))}{T-s} \right) \text{dist}(x, \partial D)
+ \Theta C \frac{\gamma(1 + \delta(T-s))^2}{T-s} - \frac{2}{(T-s)^2} + \frac{\theta}{2} \gamma^2 (1 + \delta(T-s))^4 + B \frac{\gamma(1 + \delta(T-s))^2}{T-s}
\]

where $\theta$ is as in (0.2) and $N^2 |a_{ij}(x)| \leq \Theta$. Here we used the fact that in $I^t |\partial(x, \partial D)_{x_j}| < C$ and $|D \text{dist}(x, \partial D)| = 1$. The rest of the proof is similar to the one of Lemma 2.1.

To conclude the remarks concerning the uniform $L^\infty$ estimates on $I^s$ as well as on the way that $I^s$ assumes its boundary conditions we need the following lemma.

**Lemma 2.3.** Let $\Phi \in C^2(D)$. For every $M > 0$ there exists a positive constant $C_M$ such that

\[I^s(x, s) > M \Phi(x) - C_M (T-s) \quad \text{for every } (x, s) \in D \times [0, T)
\]

and $0 < \varepsilon < 2$.

**Proof.** The result follows from the maximum principle provided we show that

\[w(x, s) = M \Phi(x) - C_M (T-s)
\]

is a subsolution of (2.1) for an appropriate constant $C_M$. Indeed, we have

\[
w_s + \frac{\varepsilon}{2} a_{ij} w_{x_i x_j} - \frac{1}{2} a_{ij} w_x w_{x_j} + \max_{y \in \gamma} (b(x, y) \cdot D w)
\geq C_M + \frac{\varepsilon}{2} Ma_{ij} \Phi_{x_i x_j} - \frac{M^2}{2} a_{ij} \Phi_x \Phi_{x_j} - BM |D \Phi|
\geq C_M - M \Theta |D^2 \Phi| - \frac{M^2}{2} \theta |D \Phi|^2 - BM |D \Phi| > 0
\]

provided $C_M$ is sufficiently large.

We continue with a result concerning a uniform in $\varepsilon$ estimate on the modulus of continuity of $I^s$ on $\overline{D} \times [0, T')$ for every $T' < T$. In particular, using Bernstein’s trick and Lemma 2.2 we obtain a uniform in $\varepsilon$ estimate
LEMMA 2.4. For every $T' < T$ there exist a constant $C = C(T')$ and $\alpha = \alpha(T')$ with $0 < \alpha < 1$ such that

$$|D^\varepsilon(x, s)| \leq C \quad \text{for every} \quad (x, s) \in \bar{D} \times [0, T']$$

and

$$|I^\varepsilon(x, s) - I^\varepsilon(x, s_\delta)| \leq C|s - s_\delta|^\alpha \quad \text{for every} \quad x \in \bar{D}, \ s, s_\delta \in [0, T'] .$$

PROOF. To simplify notation we now omit writing the superscript $\varepsilon$. We select a smooth cutoff function $\zeta = \zeta(s)$ such that $\zeta = 1$ on $[0, T')$, $\zeta \equiv 0$ near $T$. Set

$$z = \zeta^2 a_{ii} I_{x_i} I_{x_i} - \lambda I ,$$

where $\lambda > 0$ is a constant to be selected below.

Let $(x_0, s_0) \in \bar{D} \times [0, T)$ be such that

$$z(x_0, s_0) < z(x_0, s_0) \quad \text{for every} \quad (x, s) \in \bar{D} \times [0, T] .$$

Then, for every $(x, s) \in \bar{D} \times [0, T']$, we have

$$\theta |D\varepsilon(x, s)|^2 < \lambda I(x, s) + \Theta |D\varepsilon(x_0, s_0)|^2 ,$$

which implies (2.8) provided we have an estimate on $|D\varepsilon(x_0, s_0)|^2$. If $x_0 \in \partial D$ such an estimate follows immediately from Lemma 2.2. So we may assume that $x_0 \in D$. Then at $(s_0, x_0)$

$$z_{x_r} = (\xi^2 a_{ii} I_{x_i} I_{x_i} - \lambda I)_{x_r} = 0 \quad (r = 1, \ldots, N)$$

and

$$z_s + \frac{\varepsilon}{2} a_{ii} z_{x_i} z_{x_i} \leq 0 .$$

In what follows we are going to assume for technical reasons that we are dealing with an equation of the form

$$I_s + \frac{\varepsilon}{2} a_{ii} I_{x_i} - \frac{1}{2} a_{ii} I_{x_i} I_{x_i} + G(x, DI) = 0$$
with $G$ a smooth function such that
\[
\begin{cases}
|G(x, p)| < B|p|, \\
|G_x(x, p)| < B|p| 
\text{ for every } (x, p), \\
|G_{xx}(x, p)| < B
\end{cases}
\]
where $B$ is such that $|b(x, y)| < B$. The result then follows for the original equation by a simple approximation argument and we leave it to the reader to fill in the details. Using now routine calculations we deduce that at $(x_0, s_0)$
\[
0 > \frac{\varepsilon}{2} a_{ij} z_{ij} + 2\xi^2 a_{kl} I_x \left( I_x + \frac{\varepsilon}{2} a_{ij} I_x a_{ij} \right) + \lambda \left( -I_x - \frac{\varepsilon}{2} a_{ij} I_x a_{ij} \right) + 2\xi^2 a_{kl} I_x I_{x_l} + \varepsilon \xi^2 \theta_2 |D^2 I|^2 - 2\xi^2 a_{kl} a_{ij} I_x I_{x_l} I_{x_l} I_{x_l} + 2\xi^2 (a_{kl})_{x_l} I_x I_{x_l} I_{x_l} - C|D^1|^2.
\]
Moreover, 
\[
|2\xi^2 a_{kl} a_{ij} I_x I_{x_l} I_{x_l}| < \alpha C|I_x|^2 + \alpha^{-1}\xi^2 (I_{x_l})^2
\]
and similarly for products $2\xi^2 (a_{kl})_{x_l} a_{ij} I_{x_l} I_{x_l}$. If we take $\alpha$ sufficiently large then (2.12) and the above imply that at $(x_0, s_0)$
\[
0 > \frac{\varepsilon}{4} \xi^2 \theta_2 |D^2 I|^2 + 2\xi^2 a_{kl} I_x \left( \frac{\varepsilon}{2} a_{ij} I_x I_{x_l} - G(x, D I) \right) + \lambda G(x, D I) - C|D^1|^2 > a_{kl} I_x (\xi^2 a_{ij} I_x I_{x_l})_{x_l} - 2\xi^2 a_{kl} I_x (G(x, D I))_{x_l} - \frac{\lambda}{2} a_{ij} I_x I_{x_l} - \lambda C|D^1|^2 - C|D| |I|^2 > 2\xi^2 a_{kl} I_x I_{x_l} G(x, D I)_{x_l} = 2\xi^2 a_{kl} I_x I_{x_l} G_{x_l}
\]
where in the above calculations we repeatedly used $C$ to denote any constant which depends on the data.

Now, recall (2.10) to compute
\[
(2.13) \quad \frac{\lambda}{2} a_{ij} I_x I_{x_l} < \lambda C|D^1| + C|D| |I|^2 \quad \text{ at } (x_0, s_0).
\]
This for $\lambda$ sufficiently large implies
\[
|D| |I| < \frac{\lambda - 2C}{2\lambda C} \quad \text{ at } (x_0, s_0)
\]
and thus the result.
For the estimate on the Hölder norm let us assume $0 \in D$. Define

\[
\begin{align*}
I(x, s) &= I(\sqrt{e} x, s), \quad \tilde{a}_{ij}(x) = a_{ij}(\sqrt{e} x) \\
\tilde{f}(x, s) &= -\frac{1}{2} \tilde{a}_{ij} I_{x_i}(\sqrt{e} x, s) I_{x_j}(\sqrt{e} x, s) + \max_{y \in Y} \left[ b(\sqrt{e} x) \cdot I(\sqrt{e} x, s) \right] \\
\bar{Q} &= \{(x, s) : (e x, s) \in D \times [0, T]\}
\end{align*}
\]

then

\[
I_s + \frac{1}{2} \tilde{a}_{ij} I_{s x_i x_j} = \tilde{f}.
\]

Since $\tilde{f} \in C^{\infty}(\bar{Q})$ standard parabolic estimates give a local in $s$ estimate of the Hölder norm of $I^*$ with respect to $s$.

By combining all the previous results we may now proceed with the proof of part (b) of Theorem 1. Indeed, we have:

**Proof of part (b).** Lemma 2.1 and Lemma 2.4 imply that along every subsequence $\varepsilon_k \to 0$, $I^* \to \bar{I}$ locally uniformly in $D \times [0, T)$, where $I$, which in principle depends on the subsequence, is a continuous function on $D \times [0, T)$. In view of a standard result concerning viscosity solutions ([3], Theorem VI.1), $\bar{I}$ is a viscosity solution of (0.12). Moreover, Lemmas 2.2, 2.3 and 2.4 imply

\[
\begin{align*}
\bar{I}(x, s) &= 0 \quad \text{for } (x, s) \in \partial D \times [0, T) \\
\bar{I} &\in C^{1,1}(\bar{D} \times [0, T]) \\
\bar{I}(x, s) &\uparrow \infty \quad \text{as } s \uparrow T \quad \text{for } x \in D.
\end{align*}
\]

The uniqueness of viscosity solutions of (0.12) with boundary conditions (0.13) together with part (a) imply that

\[
\bar{I} = I,
\]

thus the result.

**Appendix.**

Here we want to reproduce in a simple case a result which is going to appear in M. G. Crandall, P.-L. Lions and P. E. Souganidis [4] concerning uniqueness of viscosity solutions of (0.12), (0.13). For simplicity we assume that $\sigma = 1$ (identity matrix) and, moreover, that we deal with the for-
ward in time problem

\[
\begin{align*}
    u_t + \frac{1}{2} |Du|^2 + \min_{y \in \mathbb{R}^n} [-b(x, y) \cdot Du] &= 0 \quad \text{in } D \times (0, T) \\
    u(x, t) &= 0 \quad \text{for } x \in \partial D, \ t \in (0, T) \\
    u(x, t) &\uparrow \infty \quad \text{as } t \downarrow 0 \quad \text{for } x \in D
\end{align*}
\] (A.1)

with \( u \in C^{1,1}(\overline{D} \times [\varepsilon, T]) \) for every \( \varepsilon > 0 \). The substitution \( t = T - s \) changes this into (0.12)-(0.13), with \( \sigma = 1 \). It follows from the results of G. Barles [1] and M. G. Crandall and P.-L. Lions [3] that for every \( \varphi \in C^{1,1}(\overline{D}) \) with \( \varphi|_{\partial D} = 0 \), the problem

\[
\begin{align*}
    v_t + \frac{1}{2} |Dv|^2 + \min_{y \in \mathbb{R}^n} [-b(x, y) \cdot Dv] &= 0 \quad \text{in } D \times (0, T) \\
    v(x, t) &= 0 \quad \text{for } (x, t) \in \partial D \times [0, T] \\
    v(x, 0) &= \varphi(x) \quad \text{for } x \in D
\end{align*}
\] (A.2)

has a unique viscosity solution which we denote by \( S(t)\varphi(\cdot) \). Moreover, \( S(t) \) has the semigroup property, \( S(t)\varphi \leq S(t)\psi \) if \( \varphi \leq \psi \), and \( S(t)\varphi \) is continuous in \( t \) in the uniform norm.

Let \( u \) be a solution of (A.1). It is not difficult to show that for every \( x \in \overline{D} \), \( u \) is a decreasing function of \( t \). From this, it follows that \( u(x, t) \uparrow \infty \) as \( t \downarrow 0 \) uniformly on compact subsets of \( D \). Now for every \( \varphi \in C^{1,1}(\overline{D}) \) and \( \varepsilon > 0 \) the comparison estimates of [3] imply

\[
S(t)(u(\cdot, \varepsilon) \wedge \varphi) \leq S(t)u(\cdot, \varepsilon) = u(\cdot, t + \varepsilon)
\] (A.3)

where \( w \wedge \varphi \) denotes the minimum of \( w, \varphi \). We claim that as \( \varepsilon \downarrow 0 \)

\[
u(\cdot, \varepsilon) \wedge \varphi \rightarrow \varphi \quad \text{uniformly on } \overline{D}.
\] (A.4)

Indeed, if not, there exist \( \delta > 0 \) and \( \varepsilon_n > 0 \), \( x_n \in D \) such that

\[
x_n \rightarrow x_0 \in \overline{D} \quad \text{and} \quad u(x_n, \varepsilon_n) < \varphi(x_n) - \delta.
\]

Since \( u(\cdot, \varepsilon_n) \rightarrow +\infty \) uniformly on compact subsets of \( D \), \( x_0 \in \partial D \). However,

\[
u(x_n, \varepsilon_n) \leq u(x_n, \varepsilon_n) < \varphi(x_n) - \delta, \quad u(x_0, \varepsilon) = \varphi(x_0) = 0,
\]

which gives a contradiction.
From (A.3) and (A.4)

\[ S(t)\varphi(\cdot) < u(\cdot, t) , \quad \forall \varphi \in C^{\alpha}(\overline{D}) \text{ with } \varphi|_{\partial D} = 0 . \]

If (A.1) has two solutions \( u, w \) the above observations imply that

\[
\begin{cases}
    w(\cdot, t + \varepsilon) < u(\cdot, t) & \text{for every } \varepsilon > 0 \\
    u(\cdot, t + \varepsilon) < w(\cdot, t)
\end{cases}
\]

by taking \( \varphi = w(\cdot, \varepsilon) \), \( \varphi = u(\cdot, \varepsilon) \) respectively, which yields the uniqueness.

REFERENCES


Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02915