

**DYNAMIC PROGRAMMING FOR STOCHASTIC TARGET PROBLEMS,
VISCOSITY SOLUTIONS AND HEDGING IN MARKETS WITH
PORTFOLIO CONSTRAINTS AND LARGE INVESTORS**

Rafael Serrano

SERIE DOCUMENTOS DE TRABAJO

No. 170

Octubre de 2014

Dynamic programming for stochastic target problems, viscosity solutions and hedging in markets with portfolio constraints and large investors

Rafael Serrano*

UNIVERSIDAD DEL ROSARIO
Calle 12C No. 4-69
Bogotá, Colombia

Abstract

The purpose of this expository article is to present a self-contained overview of some results on the characterization of the optimal value function of a stochastic target problem as (discontinuous) viscosity solution of a certain dynamic programming PDE and its application to the problem of hedging contingent claims in the presence of portfolio constraints and large investors.

Keywords: Stochastic target problem, dynamic programming principle, viscosity solution, Hamilton-Jacobi-Bellman equation, super-replication, large investor, portfolio constraints

JEL: C61, C02, G13

1 Introduction

Stochastic target problems are a new class of stochastic optimal control problems in which the main goal is to minimize the initial data from which a controlled continuous-time stochastic process can be driven into a given target at a pre-specified future time, by choosing an appropriate control process.

Although this problem does not fit into the class of standard control problems as presented in the usual literature related to stochastic control theory, H. M Soner and N. Touzi proved in [SO/TO 02] that the value function of the stochastic target problem still satisfies a (non-classical) dynamic programming principle (DPP). In [SO/TO2 02], they used this to characterize the optimal value function of a stochastic target problem as a discontinuous viscosity solution of an associated Hamilton-Jacobi-Bellman (HJB) second order partial differential equation with suitable boundary conditions. It should be pointed out that in this case, unlike HJB equations associated with standard stochastic optimal control problems, the use of viscosity solutions seems also necessary in order to derive the sub-solution property from the dynamic programming principle, even if the value function turned out to be differentiable.

Stochastic target problems were originally motivated by the super-replication problem in finance, in which the objective is to find the minimal initial investment that is needed, in the presence of portfolio constraints, to super-replicate (i.e. hedge without risk) a European contingent claim by

*rafael.serrano@urosario.edu.co

means of an admissible portfolio strategy. Here the control is the portfolio, the controlled stochastic process is related to the spot stock prices and the value of the portfolio, and the target is the set of all stock prices and portfolio values at maturity such that the portfolio dominates a nonlinear function of the stock prices given by the contingent claim.

The idea of super-replication (or super-hedging) was first suggested by El Karoui and Quenez [EK/QU 95], and solved by means of convex duality (dual formulation of the constraints). In general, when this approach is available, the dual problem turns out to be a standard stochastic control problem which can be solved via the classical Hamilton-Jacobi-Bellman equations.

However, to this date, there is no general convex duality approach which applies to the ‘large’ investor framework. Roughly speaking, this means that an investor could be influential enough so that his/her investment strategy, or wealth, once exposed, might affect the market prices. Mathematically speaking, this means that the coefficients of the stochastic differential equations that characterize the prices of the underlying securities could depend on the portfolio of any investor. Although the ‘small’ investor model has long been viewed as standard assumption, it has been also noted recently that some investors can affect the prices by holding and trading large amounts of securities or commodities available in the market. A probably indisputable evidence, for example, is the ‘Hedge Fund’ crisis of 1998 in the global financial market, in which the so-called large investors obviously played some important roles.

The purpose of this expository paper is to give a self-contained overview of the results on stochastic target problems mentioned above. The contents of this document are largely based on the articles by H. M. Soner and N. Touzi [SO/TO 02, SO/TO2 02] and organized as follows: In Section 2 the stochastic target problem is formulated. Section 3 presents the formulation and proof of the dynamic programming principle with help of a measurable selection result. Section 4 recalls the notion of viscosity solution of second order partial differential equations and introduces the HJB equation satisfied by the optimal value function in the discontinuous viscosity sense. The characterization of the value function is completed by means of a terminal condition given by first order variational inequality, again in the discontinuous viscosity sense. Finally, in Section 5 all these results are applied to the problem of super-replication of a contingent claim under portfolio constraints in a large investor financial market.

2 Stochastic target problems

A general stochastic target problem is a non-classical optimal stochastic control problem in which the controller tries to steer a controlled stochastic process into a given target at a terminal time, by appropriately choosing a control process.

We will be particularly interested in diffusion stochastic processes of the form $Z_{t,x,y}^\nu = (X_{t,x}^\nu, Y_{t,x,y}^\nu)$ with values in $\mathbb{R}^d \times \mathbb{R}$, and in finding the minimal initial data y such that $Y_{t,x,y}^\nu(T) \geq g(X_{t,x}^\nu(T))$ for some admissible control ν , where g is a measurable function.

2.1 Notation

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be complete probability space and let $T > 0$ be a finite time horizon. Let $\{W(t)\}_{t \in [0, T]}$ an d -dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{P})$ and $\mathbb{F} = \{\mathcal{F}(t)\}_{t \in [0, T]}$ its \mathbf{P} -completed natural filtration. For $t \in [0, T]$, $\Sigma_{t, T}$ will denote the set of all stopping times with values in the interval $[t, T]$.

Let \mathbb{H}_d^0 be the set of all càd-làg processes $X : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ progressively measurable with respect

to the filtration \mathbb{F} , and \mathbb{H}_d^p the subset of \mathbb{H}_d^0 whose elements satisfy

$$\|X\|_{\mathbb{H}_d^p}^p := E \left[\int_0^T |X(t)|^p dt \right] < \infty.$$

For a topological space A , \mathcal{B}_A will denote the set of all Borel subsets of A .

2.2 Admissible controls

Denote by \mathcal{U} the set of all progressively measurable processes $\nu = \{\nu(t), t \in [0, T]\}$ with values in a control set $U \subseteq \mathbb{R}^d$.

Definition 1. Given $\nu_1, \nu_2 \in \mathcal{U}$ and $\theta \in \Sigma_{0, T}$, we define θ -concatenation of (ν_1, ν_2) by

$$\nu_1 \overset{\theta}{\oplus} \nu_2 := \nu_1 \mathbf{1}_{[0, \theta]} + \nu_2 \mathbf{1}_{[\theta, T]}.$$

Definition 2. The set of *admissible controls* is any Borel subset \mathcal{A} of \mathcal{U} which satisfies the following conditions

A1. *Stability under concatenation:* for all $\nu_1, \nu_2 \in \mathcal{A}$ and $\theta \in \Sigma_{0, T}$, $\nu_1 \overset{\theta}{\oplus} \nu_2 \in \mathcal{A}$,

A2. *Stability under measurable selection:* For any $\theta \in \Sigma_{0, T}$ and any measurable map $\phi : (\Omega, \mathcal{F}(\theta)) \rightarrow (\mathcal{A}, \mathcal{B}_{\mathcal{A}})$ there exists $\nu \in \mathcal{A}$ such that

$$\phi = \nu \quad \text{on } [\theta, T] \times \Omega, \quad \text{Leb} \times \mathbf{P} - \text{a.e.}$$

The first condition is crucial in dynamic programming. It essentially states that the set of admissible controls has an additive structure. The second assumption is a technical condition and in many instances it follows from the topological structure imposed on \mathcal{A} , in particular, it holds if \mathcal{A} is a separable metric space:

Lemma 3. *Suppose that \mathcal{A} is a separable metric space. Then the condition A2 holds.*

Proof. We first prove that the result is true for simple functions, then the result follows by density. First suppose that ϕ is a simple function, i.e.,

$$\phi = \sum_{k=1}^{\infty} \nu_k \mathbf{1}_{B_k},$$

for some $\nu_k \in \mathcal{A}$ and pairwise disjoint sets $B_k \in \mathcal{F}(\theta)$ whose union is the whole set Ω . Define

$$\nu(t, \omega) := (\phi(\omega))(t, \omega) \mathbf{1}_{\{t \geq \theta\}}(\omega) + \tilde{\nu}(t, \omega) \mathbf{1}_{\{t < \theta\}}(\omega), \quad (1)$$

for some $\tilde{\nu} \in \mathcal{A}$. We need to show that ν so defined is progressively measurable, i.e. that for any $t \in [0, T]$ and any Borel set $A \in \mathcal{B}_U$ we have $\nu^{-1}(A) \in \mathcal{B}_{[0, t]} \otimes \mathcal{F}(t)$. Indeed, since $\tilde{\nu}$ is progressively measurable,

$$O^* := \{(s, \omega) \in [0, t] \times \Omega : s < \theta(\omega)\} \cap \tilde{\nu}^{-1}(A) \in \mathcal{B}_{[0, t]} \otimes \mathcal{F}(t).$$

Also for each k , $B_k \in \mathcal{F}(\theta)$. Then, by the definition of the σ -algebra $\mathcal{F}(\theta)$,

$$\begin{aligned} O_k &:= \{(s, \omega) \in [0, t] \times \Omega : \theta(\omega) \leq s\} \cap ([0, t] \times B_k) \\ &= \{(s, \omega) \in [0, t] \times \Omega : \theta(\omega) \leq s\} \cap ([0, t] \times \{\omega \in \Omega : \theta(\omega) \leq t\} \cap B_k) \in \mathcal{B}_{[0, t]} \otimes \mathcal{F}(t) \end{aligned}$$

Hence $\nu^{-1}(A) = O^* \cup (\cup_{k \geq 1} O_k) \in \mathcal{B}_{[0,t]} \otimes \mathcal{F}(t)$.

Now, since \mathcal{A} is separable, there exists a sequence of maps $\phi_n : \Omega \rightarrow \mathcal{A}$ which are simple functions as in Step 1, and $\lim_n \phi_n = \phi$. Let ν_n be as in (1) with ϕ_n . Then, by Step 1, ν_n is \mathbb{F} -progressively measurable and moreover ν_n converges to ν everywhere. Hence ν is \mathbb{F} -progressively measurable as well. \square

Our choice for the set of admissible controls \mathcal{A} is the collection of all adapted processes in $L^p([0, T] \times \Omega; Leb \otimes \mathbf{P})$ with values in some closed subset $U \subset \mathbb{R}^d$ and $p \geq 1$. In this case, property **A1** is trivially satisfied.

In view of lemma 3, for the property **A2** to hold, we would like \mathcal{A} to be separable: indeed, since the set of progressively measurable processes is a closed subset of $L^p([0, T] \times \Omega; Leb \otimes \mathbf{P})$, the separability of \mathcal{A} follows from the separability of L^p . According to classical results on separability (see for instance [DOOB 94], page 92), any L^p space is separable if the underlying σ -algebra is countably generated upto null sets, and since the Brownian paths are continuous, \mathbb{F} is countably generated. Therefore, this choice of \mathcal{A} is separable, and Assumption **A2** follows from Lemma 3.

2.3 The state process

Given an initial data $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}$, an initial time $t \in [0, T]$ and a control process $\nu \in \mathcal{A}$, the **state process** will be the pair of processes $Z_{t,z}^\nu = (X_{t,x}^\nu, Y_{t,x,y}^\nu)$ solution of the controlled SDE

$$\begin{aligned} dX_{t,x}^\nu(s) &= \mu(s, X_{t,x}^\nu(s), \nu(s)) ds + \sigma(s, X_{t,x}^\nu(s), \nu(s))^* dW(s), \\ dY_{t,x,y}^\nu(s) &= b(s, Z_{t,z}^\nu(s), \nu(s)) du + a(s, Z_{t,z}^\nu(s), \nu(s))^* dW(s), \end{aligned} \quad s \in (t, T) \quad (2)$$

with initial data

$$X_{t,x}^\nu(t) = x, \quad Y_{t,x,y}^\nu(t) = y.$$

We set $Z_{t,z}^\nu(s) = 0$ for $0 \leq s < t$. The functions

$$\begin{aligned} \mu : [0, T] \times \mathbb{R}^d \times U &\longrightarrow \mathbb{R}^d & \sigma : [0, T] \times \mathbb{R}^d \times U &\longrightarrow \mathbb{R}^{d \times d} \\ b : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times U &\longrightarrow \mathbb{R} & a : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times U &\longrightarrow \mathbb{R}^d \end{aligned}$$

are assumed to be bounded and globally *Lispchitz* in $(x, y, \nu) \in \mathbb{R}^d \times \mathbb{R} \times U$ uniformly in $s \in [0, T]$, i.e.

$$\begin{aligned} |\mu(s, x, \nu) - \mu(s, x', \nu')| + \|\sigma(s, x, \nu) - \sigma(s, x', \nu')\| &\leq K (|x - x'| + |\nu - \nu'|) \\ |b(s, x, y, \nu) - b(s, x', y', \nu')| + |a(s, x, y, \nu) - a(s, x', y', \nu')| &\leq K (|x - x'| + |y - y'| + |\nu - \nu'|) \end{aligned}$$

for all $x, x' \in \mathbb{R}^d$, $y, y' \in \mathbb{R}$, $\nu, \nu' \in U$ and $s \in [0, T]$, so that the state process $Z_{t,z}^\nu = (X_{t,x}^\nu, Y_{t,x,y}^\nu)$ is well defined and satisfies the following properties:

Z1. Pathwise uniqueness: let $\theta, \tau \in \Sigma_{0,T}$ and $\xi \in L^2(\Omega, \mathcal{F}(\theta), \mathbf{P}; \mathbb{R}^d \times \mathbb{R})$. If $\theta \leq \tau$ \mathbf{P} -a.s., then

$$Z_{\theta,\xi}^\nu = Z_{\tau,\zeta}^\nu \quad \text{on } [\tau, T], \quad \text{where } \zeta := Z_{\theta,\xi}^\nu(\tau).$$

Z2. Casuality: if $\nu_1 = \nu_2$ on $[\theta, \tau]$, then

$$Z_{\theta,\xi}^{\nu_1} = Z_{\theta,\xi}^{\nu_2} \quad \text{on } [\theta, \tau].$$

Z3. Measurability: the map

$$(t, z, \nu) \in [0, T] \times \mathbb{R}^d \times \mathcal{A} \longmapsto Z_{t,z}^\nu(T) \in L^2(\Omega, \mathcal{F}(T), \mathbf{P}; \mathbb{R}^d \times \mathbb{R})$$

is Borel measurable.

Properties **Z1** and **Z2** are standard results for solutions of SDEs (see e.g. [GI/SK 72]). We know also from classical estimates of such solutions that for each $\nu \in \mathcal{A}$, the map $(t, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \mapsto Z_{t,z}^\nu \in L^2(\Omega, \mathcal{F}(T), \mathbf{P}; \mathbb{R}^d \times \mathbb{R})$ is continuous. So it remains only to prove that for any fixed data $(t, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$, the map

$$\nu \in \mathcal{U} \longmapsto Z_{t,z}^\nu \in L^2(\Omega, \mathcal{F}(T), \mathbf{P}; \mathbb{R}^d \times \mathbb{R})$$

is continuous uniformly in (t, z) : indeed, if we set

$$\gamma(t, x, y, r) := \begin{pmatrix} \mu(t, x, r) \\ b(t, x, y, r) \end{pmatrix} \quad \text{and} \quad \alpha(t, x, y, r) := \begin{pmatrix} \sigma^*(t, x, r) \\ a^*(t, x, y, r) \end{pmatrix}$$

for $\nu_1, \nu_2 \in \mathcal{A}$ we can then directly estimate

$$\begin{aligned} |Z_{t,z}^{\nu_1}(T) - Z_{t,z}^{\nu_2}(T)| &\leq \int_t^T |\gamma(s, Z_{t,z}^{\nu_1}(s), \nu_1(s)) - \gamma(s, Z_{t,z}^{\nu_2}(s), \nu_2(s))| ds \\ &\quad + \left| \int_t^T [\alpha(s, Z_{t,z}^{\nu_1}(s), \nu_1(s)) - \alpha(s, Z_{t,z}^{\nu_2}(s), \nu_2(s))] dW(s) \right| \end{aligned}$$

The global Lipschitz property, Fubini's theorem and Itô's Isometry yield

$$\begin{aligned} E [|Z_{t,z}^{\nu_1}(T) - Z_{t,z}^{\nu_2}(T)|^2] &\leq C \int_t^T E [|\gamma(s, Z_{t,z}^{\nu_1}(s), \nu_1(s)) - \gamma(s, Z_{t,z}^{\nu_2}(s), \nu_2(s))|^2] ds \\ &\quad + C \int_t^T E [|\alpha(s, Z_{t,z}^{\nu_1}(s), \nu_1(s)) - \alpha(s, Z_{t,z}^{\nu_2}(s), \nu_2(s))|^2] ds \\ &\leq C \left(\|\nu_1 - \nu_2\|_{\mathbb{H}_d^2}^2 + \int_t^T E [|Z_{t,z}^{\nu_1}(s) - Z_{t,z}^{\nu_2}(s)|^2] ds \right) \end{aligned}$$

where C is a generic constant whose value may vary. By Gronwall's inequality,

$$E [|Z_{t,z}^{\nu_1}(T) - Z_{t,z}^{\nu_2}(T)|^2] \leq C e^{C(T-t)} \|\nu_1 - \nu_2\|_{\mathbb{H}_d^2}^2,$$

proving that the map $\nu \in \mathcal{U} \mapsto Z_{t,z}^\nu(T) \in L^2(\Omega, \mathcal{F}(T), \mathbf{P}; \mathbb{R}^d \times \mathbb{R})$ is Lipschitz uniformly in $(t, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$.

2.4 Formulation of the problem: the value function

For a given real-valued measurable function g defined on \mathbb{R}^d , the stochastic target control problem consists in finding the minimal initial data y for which the random constraint $Y_{t,x,y}^\nu(T) \geq g(X_{t,x}^\nu(T))$ holds almost surely. Then, the value function of the stochastic target problem is given by

$$v(t, x) := \inf\{y \in \mathbb{R} : \exists \nu \in \mathcal{A} \text{ s.t. } Y_{t,x,y}^\nu(T) \geq g(X_{t,x}^\nu(T)) \text{ } \mathbf{P} - a.s.\}. \quad (3)$$

In some cases, it is possible to find an initial datum and a control so that $Y_{t,x,y}^\nu(T) = g(X_{t,x}^\nu(T))$. In this case the problem is equivalent to a backward-forward SDE, see e.g. [PARD 98] and [MA/YO 99].

In particular, when $U = \mathbb{R}^d$, the corresponding backward-stochastic SDE has a solution (see e.g. [PA/TA 99]), and it is equal to the value function v . However, when the control set U is bounded, in general there is no solution for this equation, and v is the natural generalization of the backward-forward SDE.

An alternative generalization can be formulated by involving a nondecreasing process as follows: find a triple of \mathbb{F} -adapted processes (X, Y, ν) satisfying

$$(X, Y) \text{ solves (2) with } \nu \in \mathcal{A}, X(0) \text{ fixed, and } Y(T) + A(T) = g(X(T)) \quad (4)$$

for some nondecreasing \mathbb{F} -adapted process with $A(0) = 0$, as well as the minimality condition

$$(\tilde{X}, \tilde{Y}, \tilde{\nu}, \tilde{A}) \text{ satisfies (4)} \implies Y(\cdot) \leq \tilde{Y}(\cdot) \quad \mathbf{P} - \text{a.s.}$$

Notice that the nondecreasing process A is involved in the above definition to account for possible constraints on the control ν , see e.g. [C/K/S 98].

Let us simplify the notation by defining the following sets:

$$\begin{aligned} \mathcal{E}pi(g) &:= \{(x, y) \in \mathbb{R}^d \times \mathbb{R} : y \geq g(x)\} \\ \mathcal{G}(t, x, y) &:= \{\nu \in \mathcal{A} : Z_{t,x,y}^\nu(T) \in \mathcal{E}pi(g) \quad \mathbf{P} - \text{a.s.}\} \end{aligned}$$

Note that $\mathcal{G}(t, x, y)$ may be empty for some initial data (t, x, y) . Finally, we define

$$\mathcal{Y}(t, x) := \{y \in \mathbb{R} : \mathcal{G}(t, x, y) \neq \emptyset\}$$

Thus the stochastic target problem can be then written as

$$v(t, x) = \inf \mathcal{Y}(t, x).$$

We conclude this section with the following remark

Remark 4. The process $Y_{t,x,y}^\nu$ is strictly increasing in the initial condition y . Indeed, for given initial data $(t, x) \in [0, T] \times \mathbb{R}^d$ and a given control $\nu \in \mathcal{A}$, as $X_{t,x}^\nu$ is independent of y , the process $Y_{t,x,y}^\nu$ is solution of the one-dimensional SDE with coefficients

$$(s, y) \longmapsto b(s, X_{t,x}^\nu(s), y) \quad \text{and} \quad (s, y) \longmapsto a(s, X_{t,x}^\nu(s), y)^*.$$

Then, if $y \geq y'$, from standard comparison results for solutions of one-dimensional SDEs (e.g. [KA/SH 91], Proposition 5.2.18, pp. 293) follows that $Y_{t,x,y'}^\nu(s) \geq Y_{t,x,y}^\nu(s)$ \mathbf{P} -a.s. for all $s \in [t, T]$.

Therefore, the set $\mathcal{Y}(t, x)$ satisfies the following important property

$$\text{for all } y \in \mathbb{R}, \text{ if } y \in \mathcal{Y}(t, x) \implies [y, \infty) \subseteq \mathcal{Y}(t, x). \quad (5)$$

3 Dynamic programming principle for stochastic target problems

Dynamic Programming is usually an approach developed to solve sequential, or multi-stage, decision problems; hence, the name “dynamic” programming. The idea is to decompose a hard-to-solve problem into equivalent formats easier to solve. The essence of dynamic programming is Richard Bellman’s Principle of Optimality. This principle, even without rigorously defining the terms, is intuitive:

“An optimal policy has the property that whatever the initial state and the initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision” [BELL 57].

This is a self-evident principle in the sense that a proof by contradiction is immediate.

The stochastic target problem, as formulated in Section 1, does not have the standard form of stochastic optimal control problems for which the classical dynamic programming principle holds, and although it can be transformed in some cases into that form by using convex duality, our approach here is to prove a (non-classical) ‘geometric’ dynamic programming principle for the value function v .

We start with a measurable selection result which is mainly based on the following theorem and is the key step in the proof of our dynamic programming principle (see [BE/SH 78] for the definition of *analytic set* and *analytically measurable function*):

Proposition 5 (Jankov-von Neumann Theorem). *Let S and A be Borel Spaces and B an analytic subset of $S \times A$. Then, there exists an analytically measurable function $\phi : \text{proj}_S(B) \rightarrow A$ such that $\text{Gr}(\phi) \subseteq B$.*

Proof. see [BE/SH 78], Proposition 7.49. □

Set $S := [0, T] \times \mathbb{R}^d \times \mathbb{R}$ and $D := \{(t, z) \in S : \mathcal{G}(t, z) \neq \emptyset\}$.

Lemma 6. *For any probability measure μ on S , there exists a Borel measurable function $\phi_\mu : (D, \mathcal{B}_D) \rightarrow (\mathcal{A}, \mathcal{B}_\mathcal{A})$ such that*

$$\phi_\mu(t, z) \in \mathcal{G}(t, z) \quad \text{for } \mu\text{-almost every } (t, z) \in D$$

Proof. By assumption, S and \mathcal{A} are Borel Spaces. Set

$$B := \{(t, z, \nu) \in S \times \mathcal{A} : \nu \in \mathcal{G}(t, z)\}.$$

First, we claim that B is a Borel subset of $S \times \mathcal{A}$. Indeed, in view of **Z3**, the map $(t, z, \nu) \in S \times \mathcal{A} \mapsto Z_{t,z}^\nu(T) \in L^2(\Omega, \mathbf{P}, \mathcal{F}_T; \mathbb{R}^{d+1})$ is Borel measurable. Therefore, for any bounded continuous real-valued function f , the map

$$\begin{aligned} \Psi_f : S \times \mathcal{A} &\longrightarrow \mathbb{R} \\ (t, z, \nu) &\longmapsto E[f(Z_{t,z}^\nu(T))] \end{aligned}$$

is Borel measurable. If G is a closed subset of \mathbb{R}^{d+1} , then there exists a sequence of continuous functions f^n such that $f^n(z) \rightarrow \mathbf{1}_G(z)$ when $n \rightarrow \infty$ for all $z \in \mathbb{R}^{d+1}$, $f^n = 1$ on G , and $0 < f^n \leq 1$ outside G .

If G is open, $\Psi_{\mathbf{1}_G} = 1 - \Psi_{\mathbf{1}_{G^c}}$ is Borel measurable by the last step. This property extends to any countable union $\cup_n G_n$ of open or closed disjoint subsets G_n of \mathbb{R}^{d+1} since $\mathbf{1}_{\cup_n G_n} = \sum_n \mathbf{1}_{G_n}$. Hence, $\Psi_{\mathbf{1}_G}$ is Borel measurable for any Borel subset G of \mathbb{R}^{d+1} , in particular, for $G = \text{Epi}(g)$.

Since $\mathbf{P}(Z_{t,z}^\nu(T) \in G) = E[\mathbf{1}_G(Z_{t,z}^\nu(T))] = \Psi_{\mathbf{1}_G}(t, z, \nu)$,

$$B = \{(t, z, \nu) \in S \times \mathcal{A} : \Psi_{\mathbf{1}_G}(t, z, \nu) = 1\}$$

is a Borel subset of $S \times \mathcal{A}$.

Now, since any Borel set is also analytic (see e.g. [BE/SH 78], Proposition 7.36), B is an analytic subset of $S \times \mathcal{A}$. We may now apply the Jankov-von Neumann Theorem to deduce the existence of an analytically measurable function $\phi : D \rightarrow \mathcal{A}$ such that $\text{Gr}(\phi) \subseteq B$, i.e. $\phi(t, z) \in \mathcal{G}(t, z)$ for all $(t, z) \in D$.

Finally, we construct a Borel measurable map ϕ_μ which is equal to ϕ , μ almost everywhere: let $P(S)$ be the set of all probability measures on S . For $\mu \in P(S)$ let

$$\mu^*(E) := \inf \{ \mu(K) : E \subseteq K, K \in \mathcal{B}_S \}$$

be the outer measure w.r.t. μ and let

$$\mathcal{B}_S(\mu) := \{ E \subseteq S : \mu^*(E) + \mu^*(E^c) = 1 \}$$

be the completion of the Borel σ -algebra \mathcal{B}_S under μ . Then $\mathcal{U}_S := \bigcap_{\mu \in P(S)} \mathcal{B}_S(\mu)$ is called the universal σ -algebra. In view of Corollary 7.42.1 in [BE/SH 78], every analytic subset of S is universally measurable. In particular, any analytically measurable map ϕ is universally measurable. Since $\mathcal{U}_S \subseteq \mathcal{B}_S(\mu)$ for any $\mu \in P(S)$, it follows that ϕ is $\mathcal{B}_S(\mu)$ -measurable. Then, the definition of $\mathcal{B}_S(\mu)$ implies that there exists a Borel measurable ϕ_μ which is equal to ϕ for μ almost every $(t, z) \in D$. \square

Lemma 7. *Let $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$, $\theta \in \Sigma_{t, T}$ and $\nu \in \mathcal{A}$ be such that*

$$Y_{t,x,y}^\nu(\theta) \geq v(\theta, X_{t,x}^\nu(\theta)), \quad \mathbf{P} - a.s. \quad (6)$$

Then $\mathcal{G}(t, x, y) \neq \emptyset$.

Proof. Let μ the probability measure on $[0, T] \times \mathbb{R}^d \times \mathbb{R}$ induced by $(\theta, Z_{t,x,y}^\nu(\theta))$, i.e.

$$\mu(A \times B) = \mathbf{P}(\theta \in A, Z_{t,x,y}^\nu \in B), \quad A \in \mathcal{B}_{[0,T]}, \quad B \in \mathcal{B}_{\mathbb{R}^{d+1}},$$

and let ϕ_μ be the Borel measurable map constructed in Lemma 6 for which

$$\phi_\mu(t', z') \in \mathcal{G}(t', z') \quad \text{for } \mu\text{-a.e. } (t', z') \in D$$

i.e. $Z_{t',z'}^{\phi_\mu(t',z')}(T) \in \mathcal{Epi}(g)$, for μ - a.e. $(t', z') \in D$. In view of (6), the remark 4 and the definition of v , we have $(\theta, Z_{t,x,y}^\nu(\theta)) \in D$ a.s. Therefore, the map $\phi_\mu \circ (\theta, Z_{t,x,y}^\nu(\theta))$ is $\mathcal{F}(\theta)$ -measurable, and by condition **A2**, there exists $\nu_1 \in \mathcal{A}$ such that

$$\nu_1 = \phi_\mu \circ (\theta, Z_{t,x,y}^\nu(\theta)), \quad \text{on } [\theta, T] \times \Omega, \quad \text{Leb} \times \mathbf{P} - \text{almost everywhere.}$$

It follows that

$$Z_{t',z'}^{\nu_1}(T) = Z_{t',z'}^{\phi_\mu(t',z')}(T) \in \mathcal{Epi}(g), \quad \text{on the event } \{(\theta, Z_{t,x,y}^\nu(\theta)) = (t', z')\}, \quad (7)$$

for μ -almost every $(t', z') \in D$. Define $\tilde{\nu} := \nu \oplus \nu_1$. According to the property **A1**, stability under concatenation, $\tilde{\nu}$ is an admissible control in \mathcal{A} . Finally, we get

$$\begin{aligned} Z_{t,x,y}^{\tilde{\nu}}(T) &= Z_{\theta, Z_{t,x,y}^{\tilde{\nu}}(\theta)}^{\tilde{\nu}}(T) && \text{by } \mathbf{Z1} \\ &= Z_{\theta, Z_{t,x,y}^\nu(\theta)}^{\tilde{\nu}}(T) && \text{by } \mathbf{Z2}, \text{ since } \tilde{\nu} = \nu \text{ on } [t, \theta] \\ &= Z_{\theta, Z_{t,x,y}^\nu(\theta)}^{\nu_1}(T) && \text{by } \mathbf{Z2}, \text{ since } \tilde{\nu} = \nu \text{ on } [\theta, T] \\ &\in \mathcal{Epi}(g) && \text{by (7).} \end{aligned}$$

Hence $\tilde{\nu} \in \mathcal{G}(t, x, y)$. \square

We are now in position to state our ‘geometric’ dynamic programming principle:

Theorem 8 (Soner, Touzi (2002)). *For all $(t, x) \in [0, T] \times \mathbb{R}^d$ and $\theta \in \Sigma_{t,T}$, we have*

$$v(t, x) = \inf\{y \in \mathbb{R} : \exists \nu \in \mathcal{A} \text{ s.t. } Y_{t,x,y}^\nu(\theta) \geq v(\theta, X_{t,x}^\nu(\theta)) \text{ } \mathbf{P} - a.s.\}. \quad (8)$$

Proof. Let $w(t, x)$ denote the right-hand side of (8). In view of Remark 4, for all $y > w(t, x)$ we have $Y_{t,x,y}^\nu(\theta) \geq v(\theta, X_{t,x}^\nu(\theta))$. From Lemma 7 follows then that $\mathcal{G}(t, x, y) \neq \emptyset$, and therefore $y \geq v(t, x)$. Letting y converge to $w(t, x)$ we get $w(t, x) \geq v(t, x)$.

Conversely, for all $y > v(t, x)$, from condition **Z1** we have

$$Z_{\theta, Z_{t,x,y}^\nu(\theta)}^\nu(T) = Z_{t,x,y}^\nu(T) \in \mathcal{Epi}(g)$$

for some $\nu \in \mathcal{A}$ and therefore $Y_{t,x,y}^\nu(\theta) \geq v(\theta, X_{t,x}^\nu(\theta))$. Hence $y \geq w(t, x)$ and the required inequality follows by letting y converge to $v(t, x)$. \square

Remark 9. The above dynamic programming principle (DPP) can be interpreted as follows: at a given initial time t , the stochastic target problem with terminal time T and target $\mathcal{Epi}(g)$ is equivalent to the stochastic target problem with terminal time θ and target $\mathcal{Epi}(v(\theta, \cdot))$.

The following DPP is a direct consequence of Theorem 8 and will be the main tool to characterize the value function of the stochastic target problem as a viscosity solution of a nonlinear second order partial differential equation:

Corollary 10. *Let $(t, x) \in [0, T] \times \mathbb{R}^d$.*

(DP1) *Let $y \in \mathbb{R}$ be such that $\mathcal{G}(t, x, y) \neq \emptyset$. Then, for all $\nu \in \mathcal{G}(t, x, y)$ and $\theta \in \Sigma_{t,T}$,*

$$Y_{t,x,y}^\nu(\theta) \geq v(\theta, X_{t,x}^\nu(\theta)), \quad \mathbf{P} - a.s.$$

(DP2) *Set $y^* := v(t, x)$ and let $\theta \in \Sigma_{t,T}$. Then for all $\nu \in \mathcal{A}$ and $\eta > 0$,*

$$\mathbf{P} [\mathcal{Y}_{t,x,y^*-\eta}^\nu(\theta) > v(\theta, X_{t,x}^\nu(\theta))] < 1.$$

Remark 11. The part (DP2) says that $y^* = v(t, x)$ is precisely the minimal (optimal!) value of y for which the optimality property of the value function (DP1) holds at time $\theta \in \Sigma_{t,T}$.

4 Dynamic programming PDE and viscosity solution property

The aim of this section is to prove that the value function of the stochastic target problem solves a second order partial differential equation (PDE) in the viscosity sense. First, we motivate and introduce the notion of viscosity solution, and then give the proof of the viscosity sub- and super-solution properties by means of the dynamic programming principle stated in Corollary 10. Then, we present a characterization of the value function as the unique viscosity solution by specifying a terminal condition.

4.1 Viscosity solutions of second order PDEs: introduction and definition

The primary virtue of the theory of viscosity solutions is that it allows non-differentiable functions to be solutions of fully non-linear second order partial differential equations of the form

$$F(x, u(x), Du(x), D^2u(x)) = 0, \quad x \in \mathcal{O} \tag{9}$$

where \mathcal{O} is an open subset of \mathbb{R}^N ,

$$F : \mathcal{O} \times \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}(N) \longrightarrow \mathbb{R},$$

and $\mathcal{S}(N)$ is the set of real symmetric $N \times N$ matrices. $Du(x)$ and $D^2u(x)$ correspond respectively to the gradient and the Hessian of the unknown $u : \mathcal{O} \rightarrow \mathbb{R}$ at $x \in \mathcal{O}$.

The main assumptions on the function F will be the following

Definition 12. F is said to be *degenerate elliptic* if it is nonincreasing in its matrix argument:

$$F(x, r, p, X) \leq F(x, r, p, Y) \quad \text{whenever } Y \leq X.$$

If F is degenerate elliptic, we say that it is *proper* if it is also nondecreasing in r , i.e.

$$F(x, r, p, X) \leq F(x, s, p, Y) \quad \text{whenever } Y \leq X \text{ and } r \leq s.$$

Recall the usual ordering in $\mathcal{S}(N)$: $Y \leq X$ iff $\langle X\eta, \eta \rangle \leq \langle Y\eta, \eta \rangle$ for all $\eta \in \mathbb{R}^N$.

Since it is a bit difficult to find examples of second order PDEs which cannot be solved in the classical sense unless the equation is very degenerate, we will consider first the case

$$F(x, u(x), Du(x)) = 0, \quad x \in \mathcal{O} \tag{10}$$

i.e. PDEs of first order (so very degenerate!), and then provide a precise definition for the second order case.

Example 13. On the need for non-smooth solutions. Consider the boundary problem given by the first order PDE

$$\begin{aligned} |Du(x)|^2 - 1 &= 0, & x \in \mathcal{O} \subset \mathbb{R}^2 \\ u(x) &= 0, & x \in \partial\mathcal{O}, \end{aligned} \tag{11}$$

which corresponds to (10) with $F(x, u, p) = p_1^2 + p_2^2 - 1$. If the boundary $\partial\mathcal{O}$ is smooth, the *distance function* $u(x) = \text{dist}(x, \partial\mathcal{O})$ is smooth in a neighborhood of the boundary and solves the equation in such neighborhood.

To establish the desirability of allowing non-differentiable solutions, let us assume u smooth on \mathcal{O} and use the method of characteristics towards a contradiction: let $x(t) \in \mathbb{R}^2$ be the solution of the initial value problem

$$x'(t) = \frac{d}{dt}x(t) = \frac{\partial F}{\partial p}(x(t), u(x(t)), Du(x(t))) = 2Du(x(t)), \quad x(0) = y \in \partial\mathcal{O}$$

over the largest interval for which this solution exists. A computation yields

$$\frac{d}{dt}Du(x(t)) = D^2u(x(t))x'(t) = 2D^2u(x(t))Du(x(t)) = 0$$

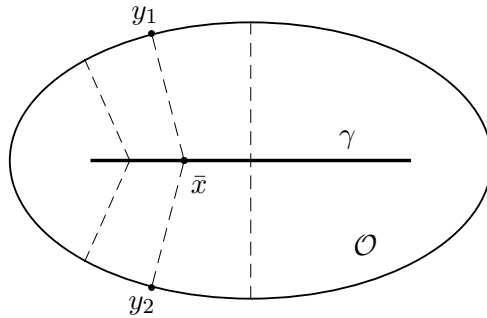


Figure 1: crossing characteristics

where the last equality arises from differentiating $F(x, u(x), Du(x)) = 0$ with respect to x . Hence, $Du(x)$ is constant on the curve $t \rightarrow x(t)$. It would then follow that u is constructed along the ray

$$x(t) = y + 2t\mathbf{n},$$

where $\mathbf{n} = Du(y)$ is the interior unit normal to the set \mathcal{O} at the point y , and along this ray, one has $u(x) = |x - y|$. However, the resulting equality

$$Du(y + 2tDu(y)) = Du(y)$$

yields a contradiction as soon as \mathcal{O} is bounded, since in this case one could find a set γ of interior points \bar{x} with crossing characteristics (Figure 1), that is, points of the form

$$\bar{x} = y_1 + 2tDu(y_1) = y_2 + 2tDu(y_2)$$

with $t > 0$ but $y_1 \neq y_2$, for which $Du(\bar{x}) = Du(y_1) = Du(y_2)$, which is not necessarily true as $Du(y_1)$ and $Du(y_2)$ are not necessarily equal. In conclusion, the distance function is not differentiable at the points \bar{x} such that

$$\text{dist}(x, \partial\mathcal{O}) = |\bar{x} - y_1| = |\bar{x} - y_2|$$

for two distinct $y_1, y_2 \in \partial\mathcal{O}$, and therefore, the boundary value problem for the first order PDE (11) does not admit a global \mathcal{C}^1 solution.

The previous example suggests that, in order to overcome the problem of defining a non-regular solution to (10), we should relax our requirements and consider solutions in a more general sense. As, by Rademacher's theorem, every Lipschitz continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is differentiable almost everywhere, Kruzkow introduced in the 60's the notion of generalized solutions, i.e. solutions which satisfy the equation almost everywhere. This is a powerful idea and a lot of results have been obtained under different set of hypothesis (for a complete description see [LIONS 83] and the references therein).

Unfortunately, this concept of solution is far too weak, and does not lead to any useful uniqueness result:

Example 14. Consider the 1-dimensional equation $|u'(x)| = 1$ in $(-1, 1)$ with boundary conditions $u(-1) = u(1) = 0$. Clearly there are not classical solutions, but one can build infinitely many generalized solutions: is it enough to alternate segments with slope 1 and segments with slope -1 (Figure 2). Furthermore one can construct a sequence of generalized solutions which converge to $u \equiv 0$ that is not a generalized solution. From the application point of view this lack of uniqueness and stability is an important problem.

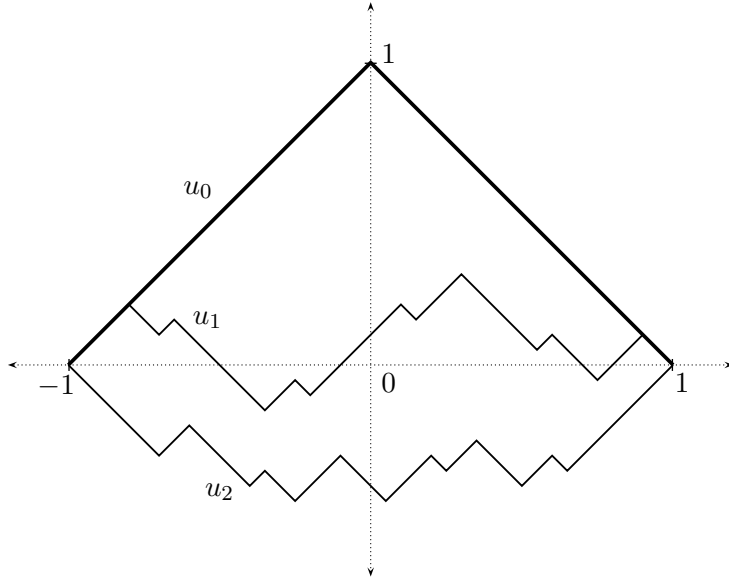


Figure 2: generalized solutions of $|u'(x)| = 1$.

So, how could the new concept of solution be defined? Our approach is to consider the approximate problem

$$F(x, u^\epsilon(x), Du^\epsilon(x)) = \epsilon \Delta u^\epsilon(x), \quad x \in \mathcal{O} \quad (12)$$

where $\Delta u^\epsilon = \sum_{i=1}^N \frac{\partial^2 u^\epsilon}{\partial x_i^2}$ is the Laplacian operator and $\epsilon > 0$. The idea is that whereas (10) involves a fully nonlinear first order PDE, (12) is a quasilinear PDE which, under suitable conditions, turns out to have a smooth solution. Indeed, the term $\epsilon \Delta$ in (12) regularizes the Hamilton-Jacobi equation. Then of course we hope that as $\epsilon \rightarrow 0$ the solutions u^ϵ of (12) will converge to some sort of weak solution of (10). This technique is known as the method of *vanishing viscosity*, since the term $\epsilon \Delta u^\epsilon$ is used to model fluid viscosity.

Unfortunately, as $\epsilon \rightarrow 0$ we can expect to lose control over the various estimates of the function u^ϵ and its derivatives: these estimates depend strongly on the regularizing effect of $\epsilon \Delta$ and blow up as $\epsilon \rightarrow 0$. However, it turns out that we can often in practice at least be sure that the family $\{u^\epsilon\}_{\epsilon > 0}$ is bounded and equicontinuous on compact subsets of \mathcal{O} . Consequently the Arzela-Ascoli compactness criterion ensures that

$$u^{\epsilon_j} \rightarrow u, \quad \text{locally uniformly on } \mathcal{O}, \quad (13)$$

for some subsequence $\{u^{\epsilon_j}\}_{j=0}^\infty$ and some continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$. Now we can surely expect that u is some kind of solution of (10), but as we only know u is continuous, and have absolutely no information whether Du exists in any sense, such an interpretation is difficult.

We will call the solution we build a *viscosity solution*, in honor to the vanishing viscosity technique. Our main goal now is to discover an intrinsic characterization of such generalized solution of (10).

Motivation for the definition of viscosity solutions. Let us assume henceforth that F is continuous. The technique alluded above works as follows: fix any smooth test function $\varphi \in \mathcal{C}^\infty(\mathcal{O})$ and suppose

$$u - \varphi \text{ has a } \textit{strict} \text{ local maximum at } x_0 \in \mathcal{O} \quad (14)$$

This means

$$(u - \varphi)(x_0) > (u - \varphi)(x)$$

for all points x sufficiently close to x_0 but with $x \neq x_0$. We claim that for each sufficiently small $\epsilon_j > 0$, there exists a point x_{ϵ_j} such that

$$u^{\epsilon_j} - \varphi \text{ has a local maximum at } x_{\epsilon_j} \quad (15)$$

and

$$x_{\epsilon_j} \rightarrow x_0 \text{ as } j \rightarrow \infty. \quad (16)$$

Indeed, note that for each sufficiently small $r > 0$, (14) implies

$$\max_{|x-x_0|=r} (u - \varphi)(x) < (u - \varphi)(x_0).$$

In view of (13), $u^{\epsilon_j} \rightarrow u$ uniformly on $\overline{B_r(x_0)}$, so

$$\max_{|x-x_0|=r} (u^{\epsilon_j} - \varphi)(x) < (u^{\epsilon_j} - \varphi)(x_0)$$

provided that ϵ_j is small enough. Consequently $u^{\epsilon_j} - \varphi$ attains a local maximum at some point on $B_r(x_0)$. We can next replace r by a sequence $\{r_n\}_{n \geq 0}$ tending to zero to obtain (15) and (16).

Now, owing to (15), we see that the equation

$$D(u^{\epsilon_j} - \varphi)(x_{\epsilon_j}) = 0 \quad (17)$$

and the inequality

$$\Delta(u^{\epsilon_j} - \varphi)(x_{\epsilon_j}) \leq 0, \quad (18)$$

hold. We can consequently calculate

$$\begin{aligned} F(x_{\epsilon_j}, u^{\epsilon_j}(x_{\epsilon_j}), D\varphi(x_{\epsilon_j})) &= F(x_{\epsilon_j}, u^{\epsilon_j}(x_{\epsilon_j}), Du^{\epsilon_j}(x_{\epsilon_j})), \quad \text{by (17)} \\ &= \epsilon_j \Delta u^{\epsilon_j}(x_{\epsilon_j}), \quad \text{by (12)} \\ &\leq \epsilon_j \Delta \varphi(x_{\epsilon_j}), \quad \text{by (18)}. \end{aligned} \quad (19)$$

Now let $\epsilon_j \rightarrow 0$ and remember (16). Since φ is smooth and F is continuous, we deduce

$$F(x_0, u(x_0), D\varphi(x_0)) \leq 0. \quad (20)$$

Suppose now, instead of (14), that

$$u - \varphi \text{ has a local maximum at } x_0 \in \mathcal{O} \quad (21)$$

but that this maximum is not necessarily strict. Then $u - \tilde{\varphi}$ has a strict local maximum at x_0 where $\tilde{\varphi}(x') := \varphi(x_0) + \delta|x' - x_0|^2$, $\delta > 0$. We thus conclude as above that

$$F(x_0, u(x_0), D\varphi(x_0)) \leq 0,$$

whereupon (20) again follows, since since φ and $\tilde{\varphi}$ agree at x_0 . Consequently (21) implies the inequality (20). Similarly, the reverse inequality

$$F(x_0, u(x_0), D\varphi(x_0)) \geq 0 \quad (22)$$

can be deduced, provided that

$$u - \varphi \text{ has a local minimum at } x_0 \in \mathcal{O}. \quad (23)$$

The proof is exactly like that above, except that the inequalities in (18), and thus in (19), are reversed.

In summary, we have discovered for any smooth function φ that (20) follows from (21), and (22) from (23), and thus, we managed to “put all the derivatives onto φ ”, at the expense of certain inequalities holding.

A viscosity solution will be then defined precisely as a continuous function u satisfying the corresponding inequalities (20) and (22) in the second order case, provided that (21) and (23) hold:

Definition 15. Let $F : \mathcal{O} \times \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}(N) \rightarrow \mathbb{R}$ be proper and consider the second order PDE

$$F(x, u(x), Du(x), D^2u(x)) = 0, \quad x \in \mathcal{O}. \quad (24)$$

(a) $u : \mathcal{O} \rightarrow \mathbb{R}$ is a *viscosity subsolution* of (24) if it is upper semicontinuous and for each $\varphi \in \mathcal{C}^2(\mathcal{O})$ and $x_0 \in \mathcal{O}$ such that $u - \varphi$ has a local maximum at x_0 we have

$$F(x_0, u(x_0), D\varphi(x_0), D^2\varphi(x_0)) \leq 0.$$

(b) $u : \mathcal{O} \rightarrow \mathbb{R}$ is a *viscosity supersolution* of (24) if it is lower semicontinuous and for each $\varphi \in \mathcal{C}^2(\mathcal{O})$ and $x_0 \in \mathcal{O}$ such that $u - \varphi$ has a local minimum at x_0 we have

$$F(x_0, u(x_0), D\varphi(x_0), D^2\varphi(x_0)) \geq 0.$$

Finally, $u : \mathcal{O} \rightarrow \mathbb{R}$ is a *viscosity solution* of (24) if it is both viscosity sub- and supersolution of (24).

Remark 16. As in the vanishing viscosity method, the above definition does not change if the minimum or maximum are local and/or strict.

Recall that $u : \mathcal{O} \rightarrow \mathbb{R}$ is upper (respectively, lower) semicontinuous if for any $x \in \mathcal{O}$ and $\varepsilon > 0$ there is δ such that $f(y) < f(x) + \varepsilon$ (respectively, $f(y) > f(x) - \varepsilon$) for all $y \in \mathcal{O} \cap B_\delta(x)$. Therefore, according to Definition 15, a viscosity solution is automatically continuous.

Observe furthermore that the notion of viscosity solutions is consistent with that of a classical solution: it is not difficult to check that every classical solution is also a viscosity solution, and moreover, if a viscosity solution is differentiable at some point, it solves the PDE (24) there in the classical sense.

Example 17. In connection with Example 14, the function $u_0(x) = 1 - |x|$ is a viscosity solution of $|u'(x)| = 1$ in $(-1, 1)$ with boundary conditions $u(-1) = u(1) = 0$. Indeed, if $x_0 \neq 0$ is a local extremum of $u - \varphi$, then $u'(x_0) = \varphi'(x_0)$. Therefore, at these points both the sub- and supersolution conditions are trivially satisfied. Also, if 0 is a local minimum of $u - \varphi$, a simple calculation shows that $|\varphi'(0)| \leq 1$ and the supersolution condition holds. Observe that 0 cannot attain a local maximum for $u - \varphi$, as this would imply $-1 \geq \varphi'(0) \geq 1$.

Moreover, u_0 can be shown to be the only one, among those shown in Figure 2, that can be obtained as a vanishing viscosity limit.

As the value function of the stochastic target problem can be discontinuous, we need to extend the definition of viscosity solutions and include solutions that are not necessarily continuous. First, we need the following additional tool:

Definition 18. Let $u : \mathcal{O} \subseteq \mathbb{R}^N \rightarrow \overline{\mathbb{R}} = [-\infty, \infty]$. We denote respectively by u_* and u^* the lower and upper semicontinuous envelope of u , i.e.

$$u_*(x) := \liminf_{y \rightarrow x} u(y) := \lim_{r \rightarrow 0^+} \inf\{u(y) : y \in \mathcal{O}, |y - x| \leq r\}$$

$$u^*(x) := \limsup_{y \rightarrow x} u(y) := \lim_{r \rightarrow 0^+} \sup\{u(y) : y \in \mathcal{O}, |y - x| \leq r\}$$

It is easy to verify that u_* (resp. u^*) is lower (resp. upper) semicontinuous:

Definition 19. A locally bounded function $u : \mathcal{O} \rightarrow \mathbb{R}$ is a (discontinuous) *viscosity solution* of (24) if u_* and u^* are, respectively, viscosity super- and subsolution of (24), according to Definition 15.

Finally, we mention (without proof) a standard comparison result which is general strategy to prove uniqueness of viscosity solutions of second order PDEs

Theorem 20. Let \mathcal{O} be a bounded open subset \mathbb{R}^N and $F : \mathcal{O} \times \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}(N) \rightarrow \mathbb{R}$ be proper and continuous. Assume that there exists $\gamma > 0$ such that

$$\gamma(r - s) \leq F(x, r, p, X) - F(x, s, p, X), \quad \text{for } r \geq s, (x, p, X) \in \overline{\mathcal{O}} \times \mathbb{R}^N \times \mathcal{S}(N).$$

Assume further that there exists a function $\omega : [0, \infty] \rightarrow [0, \infty]$ such that $\omega(0+) = 0$ and

$$F(y, r, \alpha(x - y), Y) - F(x, r, \alpha(x - y), X) \leq \omega(\alpha|x - y|^2 + |x - y|)$$

whenever $x, y \in \mathcal{O}$, $r \in \mathbb{R}$, $X, Y \in \mathcal{S}(N)$ and

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq 3\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}$$

holds. Let u and v be, respectively, sub- and supersolution of $F = 0$ in \mathcal{O} and $u \leq v$ on $\partial\mathcal{O}$. Then $u \leq v$ on $\overline{\mathcal{O}}$.

For a proof of the previous theorem and a much more complete description of the theory of viscosity solutions see [C/I/L 92] and the references therein.

4.2 The dynamic programming PDE for the stochastic target problem

Recall that the value function of our stochastic target problem is given by

$$v(t, x) := \inf\{y \in \mathbb{R} : Y_{t,x,y}^\nu(T) \geq g(X_{t,x}^\nu(T)) \text{ } \mathbf{P} - \text{a.s.}, \text{ for some } \nu \in \mathcal{A}\}. \quad (25)$$

where $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is some measurable function and $Z_{t,z}^\nu = (X_{t,x}^\nu, Y_{t,x,y}^\nu)$ is the $\mathbb{R}^d \times \mathbb{R}$ -valued process solution of the controlled SDE

$$\begin{aligned} dX_{t,x}^\nu(s) &= \mu(s, X_{t,x}^\nu(s), \nu(s)) ds + \sigma(s, X_{t,x}^\nu(s), \nu(s))^* dW(s), \\ dY_{t,x,y}^\nu(s) &= b(s, Z_{t,z}^\nu(s), \nu(s)) ds + a(s, Z_{t,z}^\nu(s), \nu(s))^* dW(s), \end{aligned} \quad s \in (t, T)$$

with initial data $(X_{t,x}^\nu(t), Y_{t,x,y}^\nu(t)) = (x, y) \in \mathbb{R}^d \times \mathbb{R}$. The set of admissible controls \mathcal{A} is the collection of all adapted processes in $L^p([0, T] \times \Omega; \text{Leb} \otimes \mathbf{P})$ with values in the control set $U \subset \mathbb{R}^d$, which we assume to be convex and compact.

Definition 21. Let δ_U be the *support function* of U ,

$$\delta_U(\zeta) := \sup_{\nu \in U} (\nu^* \zeta), \quad \zeta \in \mathbb{R}^d.$$

We shall denote by \tilde{U} the *effective domain* of δ_U and by \tilde{U}_1 the restriction of \tilde{U} to the unit circle:

$$\tilde{U} = \{\zeta \in \mathbb{R}^d : \delta_U(\zeta) \in \mathbb{R}\} \quad \text{and} \quad \tilde{U}_1 = \{\zeta \in \tilde{U} : |\zeta| = 1\}$$

so that \tilde{U} is the closed cone generated by \tilde{U}_1 .

One can think as \tilde{U}_1 as the set of allowable directions in which a control can act. However, under our assumptions, since U is a bounded subset of \mathbb{R}^d ,

$$\tilde{U} = \mathbb{R}^d \quad \text{and} \quad \tilde{U}_1 = \{\zeta \in \mathbb{R}^d : |\zeta| = 1\}$$

Remark 22. The compactness of U is only needed to establish some results which require us to extract convergent subsequences from sequences in U . Therefore, many results contained in this section hold for a general closed convex subset U . For this reason, we shall keep using the notation \tilde{U} and \tilde{U}_1 .

Remark 23. For later reference, note that the closed convex set U can be characterized in terms of \tilde{U} (see, e.g., [ROCK 70]):

$$\begin{aligned} \nu \in U \text{ iff } \inf_{\zeta \in \tilde{U}} (\delta_U(\zeta) - \zeta^* \nu) &\geq 0, \\ \text{iff } \inf_{\zeta \in \tilde{U}_1} (\delta_U(\zeta) - \zeta^* \nu) &\geq 0; \end{aligned}$$

the second characterization follows from the facts that \tilde{U} is the closed cone generated by \tilde{U}_1 and δ_U is positively homogeneous.

Remark 24. We shall also use the following characterization of $\text{int}(U)$ in terms of \tilde{U}_1 :

$$\nu \in \text{int}(U) \text{ iff } \inf_{\zeta \in \tilde{U}_1} (\delta_U(\zeta) - \zeta^* \nu) > 0.$$

To see this, suppose that the right-hand side infimum is zero. Then, for all $\varepsilon > 0$, there exists some $\zeta_0 \in \tilde{U}_1$ such that $0 \leq \delta_U(\zeta_0) - \zeta_0^* \nu \leq \varepsilon/2$. Then $\delta_U(\zeta_0) - \zeta_0^* (\nu + \varepsilon \zeta_0) < 0$, and therefore $\nu + \varepsilon \zeta_0 \notin U$ by the previous remark. Since $\varepsilon > 0$ is arbitrary, this proves that $\nu \notin \text{int}(U)$. Conversely, suppose that $l := \inf_{\zeta \in \tilde{U}_1} (\delta_U(\zeta) - \zeta^* \nu) > 0$. Then, by the Cauchy-Schwartz inequality and the characterization of the previous remark, it is easily checked that the ball around ν with radius l is included in U .

Remark 25. Let χ_U be the function defined on \mathbb{R}^d by

$$\chi_U(\nu) := \inf_{\zeta \in \tilde{U}_1} (\delta_U(\zeta) - \zeta^* \nu).$$

Then χ_U is continuous. Indeed, since \tilde{U}_1 is a compact subset of \mathbb{R}^d , the infimum in the above definition of $\chi_U(\nu)$ is attained, say, at $\hat{\zeta}(\nu) \in \tilde{U}_1$. Then, for all $\nu, \nu' \in \mathbb{R}^d$,

$$\chi_U(\nu') \leq \delta_U(\hat{\zeta}(\nu)) - \hat{\zeta}(\nu)^* \nu' + \hat{\zeta}(\nu)^* (\nu - \nu') \leq \chi_U(\nu) + |\nu - \nu'|$$

by the Cauchy-Schwarz inequality. By symmetry, this proves that χ_U is a contracting mapping.

The crucial assumption in this section will be the following: the matrix $\sigma(t, x, r)$ is invertible for every $(t, x, r) \in [0, T] \times \mathbb{R}^d \times U$, and the function

$$r \mapsto \sigma(t, x, r)^{-1} a(t, x, y, r)$$

is one-to-one for all $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$. We will denote with ψ its inverse, i.e.

$$\sigma(t, x, r)^{-1} a(t, x, y, r) = p \iff r = \psi(t, x, y, p), \quad (26)$$

for all $(t, x, y, r) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \times U$, $p \in \mathbb{R}^d$. Since we wish to hit the deterministic target $\mathcal{E}pi(g)$ with probability one, the diffusion process has to degenerate along certain directions and the function ψ captures this fact: the equation (26) will enable us to match the stochastic parts of X and Y by a judicious choice of the control process ν . Similar assumptions were also utilized in the four step scheme used to solve forward-backward SDEs.

Finally, we introduce the *second order differential operator* associated to the process X^ν :

$$\mathcal{L}^\nu u(t, x) := \frac{\partial u}{\partial t}(t, x) + \mu^*(t, x, \nu) Du(t, x) + \frac{1}{2} \text{Tr}(\sigma(t, x, \nu)^* \sigma(t, x, \nu) D^2 u(t, x)),$$

where Du and D^2u denote, respectively, the gradient and the Hessian matrix of u with respect to the x variable.

The next theorem characterizes the value function v as (discontinuous) viscosity solution of an associated second order PDE:

Theorem 26. *Assume that μ, σ, a and b are all bounded and satisfy the usual Lipschitz conditions (1.2) and that v^*, v_* are finite everywhere. Further assume that U has non-empty interior. Then the value function v of the stochastic target problem is a discontinuous viscosity solution of the equation on $[0, T] \times \mathbb{R}^d$,*

$$\min \{ -\mathcal{L}^{\nu_0} u(t, x) + b(t, x, u(t, x), \nu_0(t, x)); H(t, x, u(t, x), Du(t, x)) \} = 0, \quad (27)$$

where

$$\nu_0(t, x) := \psi(t, x, u(t, x), Du(t, x)), \quad (28)$$

$$H(t, x, u(t, x), Du(t, x)) := \chi_U(\psi(t, x, u(t, x), Du(t, x))), \quad (29)$$

i.e., v_* and v^* are, respectively, viscosity supersolution and subsolution of (27).

Remark 27. In view of Remark 23, $H \geq 0$ iff $\nu_0 \in U$. Since U has a nonempty interior, it follows from Remark 24 that $H > 0$ iff $\nu_0 \in \text{int}(U)$.

Remark 28. Although $[0, T] \times \mathbb{R}^d$ is not an open domain, in our setting the time variable moves forward so that the zero boundary is not relevant and the general theory of viscosity solutions is still valid.

The proof of the Theorem 26 will be completed in the following two subsections. The supersolution part of the claim follows from (DP1) in Section 3 and a classical argument in the viscosity theory which is due to P.L. Lions. We shall take advance of the fact that the inequality (DP1) is in the a.s. sense. This allows for a suitable change of measure before taking expectations. The subsolution part is obtained from (DP2) by means of a contraposition argument.

The above result will be completed in Theorem 31 by the description of the boundary condition. The reader who is not interested in the technical proof of the Theorem 26 can go directly to the Section 5.

4.3 Proof of the viscosity supersolution property

Fix $(t_0, x_0) \in [0, T) \times \mathbb{R}^d$, and let φ be a $\mathcal{C}^2([0, T] \times \mathbb{R}^d)$ -function satisfying

$$(v_* - \varphi)(t_0, x_0) = \min_{(t,x) \in [0,T] \times \mathbb{R}^d} (v_* - \varphi).$$

We can assume w.l.o.g. that $v_*(t_0, x_0) = \varphi(t_0, x_0)$ (just take $\varphi - \varphi(t_0, x_0) + v_*(t_0, x_0)$ instead of φ , which does not affect the derivatives on φ). Observe furthermore that $v \geq v_* \geq \varphi$ on $[0, T) \times \mathbb{R}^d$.

Step 1. Let $(t_n, x_n)_{n \geq 1}$ be a sequence in $[0, T) \times \mathbb{R}^d$ such that

$$(t_n, x_n) \rightarrow (t_0, x_0) \quad \text{and} \quad v(t_n, x_n) \rightarrow v_*(t_0, x_0).$$

Set $y_n := v(t_n, x_n) + (1/n)$ and $z_n := (x_n, y_n)$. Then, by definition of the stochastic target control problem, the set $\mathcal{G}(t_n, z_n)$ is not empty. Let ν_n be any element of $\mathcal{G}(t_n, z_n)$.

For any $[0, T - t_n]$ -valued stopping time θ_n (to be chosen later), the dynamic programming principle (DP1) yields

$$Y_{t_n, z_n}^{\nu_n}(t_n + \theta_n) \geq v(t_n + \theta_n, X_{t_n, x_n}^{\nu_n}(t_n + \theta_n)) \quad \mathbf{P} - \text{a.s.}$$

Set $\beta_n := y_n - \varphi(t_n, x_n)$. Since $y_n \rightarrow v_*(t_0, x_0)$ and $\varphi(t_n, x_n) \rightarrow \varphi(t_0, x_0) = v_*(t_0, x_0)$ as n tends to infinity, we get $\beta_n \rightarrow 0$. Further, since $v \geq v_* \geq \varphi$, we have

$$v(t_n + \theta_n, X_{t_n, x_n}^{\nu_n}(t_n + \theta_n)) \geq \varphi(t_n + \theta_n, X_{t_n, x_n}^{\nu_n}(t_n + \theta_n)) \quad \mathbf{P} - \text{a.s.}$$

Then

$$\beta_n + [Y_{t_n, z_n}^{\nu_n}(t_n + \theta_n) - y_n] - [\varphi(t_n + \theta_n, X_{t_n, x_n}^{\nu_n}(t_n + \theta_n)) - \varphi(t_n, x_n)] \geq 0 \quad \mathbf{P} - \text{a.s.}$$

By Itô's Lemma,

$$\begin{aligned} 0 \leq & \beta_n + \int_{t_n}^{t_n + \theta_n} [b(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \mathcal{L}^{\nu_n(s)} \varphi(s, X_{t_n, x_n}^{\nu_n}(s))] ds \\ & + \int_{t_n}^{t_n + \theta_n} [a(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \sigma(s, X_{t_n, x_n}^{\nu_n}(s), \nu_n(s)) D\varphi(s, X_{t_n, x_n}^{\nu_n}(s))]^* dW(s). \end{aligned} \quad (30)$$

Step 2. For some large constant C , set

$$\theta_n := \inf\{s > t_n : |X_{t_n, x_n}^{\nu_n}(s)| \geq C\}.$$

Since U is bounded in \mathbb{R}^d and $(t_n, x_n) \rightarrow (t_0, x_0)$, one can easily show that

$$\liminf_{n \rightarrow \infty} (t \wedge \theta_n) > t_0, \quad \text{for all } t > t_0. \quad (31)$$

For $\xi \in \mathbb{R}$, we introduce the probability measure \mathbf{P}_n^ξ equivalent to \mathbf{P} defined by the density process

$$M_n^\xi(t) := \mathcal{E} \left(-\xi \int_{t_n}^{t \wedge \theta_n} (a - \sigma D\varphi)(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s))^* dW(s) \right), \quad t \geq t_n,$$

where $\mathcal{E}(\cdot)$ is the Doléans-Dade exponential operator:

$$\mathcal{E} \left(- \int_0^t \lambda(s)^* dW(s) \right) = \exp \left(- \int_0^t \lambda(s)^* dW(s) - \frac{1}{2} \int_0^t |\lambda(s)|^2 ds \right)$$

for a \mathbb{R}^d -valued adapted process $\lambda(s)$, $0 \leq s \leq T$. We shall denote by E_n^ξ the conditional expectation with respect to $\mathcal{F}(t_n)$ under \mathbf{P}_n^ξ .

We take the conditional expectation with respect to $\mathcal{F}(t_n)$ under \mathbf{P}_n^ξ in (30). The result is

$$\begin{aligned} 0 \leq & \beta_n + E_n^\xi \left[\int_{t_n}^{t_n+h \wedge \theta_n} [b(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \mathcal{L}^{\nu_n(s)} \varphi(s, X_{t_n, x_n}^{\nu_n}(s))] ds \right] \\ & - \xi E_n^\xi \left[\int_{t_n}^{t_n+h \wedge \theta_n} |a(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \sigma(s, X_{t_n, x_n}^{\nu_n}(s), \nu_n(s)) D\varphi(s, X_{t_n, x_n}^{\nu_n}(s))|^2 ds \right] \end{aligned}$$

for all $h > 0$. We now consider two cases:

- Suppose that the set $\{n \geq 1 : \beta_n = 0\}$ is finite. Then there exists a subsequence, renamed $(\beta_n)_{n \geq 1}$, such that $\beta_n \neq 0$ for all $n \geq 1$. Set $h_n = \sqrt{|\beta_n|}$ and $k_n := \theta \wedge (t_n + h_n)$.
- If the set $\{n \geq 1 : \beta_n = 0\}$ is not finite, then there exists a subsequence, renamed $(\beta_n)_{n \geq 1}$, such that $\beta_n = 0$ for all $n \geq 1$. Set $h_n := 1/n$ and $k_n := \theta \wedge (t_n + h_n)$.

The last inequality still holds if we replace $t \wedge \theta_n$ by k_n . We then divide this inequality by h_n and send n to infinity by using (31), the dominated convergence theorem, and the right continuity of the filtration. The result is

$$\begin{aligned} 0 \leq \liminf_{n \rightarrow \infty} \frac{1}{h_n} \int_{t_n}^{t_n+h_n} & \left[b(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \mathcal{L}^{\nu_n(s)} \varphi(s, X_{t_n, x_n}^{\nu_n}(s)) \right. \\ & \left. - \xi |a(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \sigma(s, X_{t_n, x_n}^{\nu_n}(s), \nu_n(s)) D\varphi(s, X_{t_n, x_n}^{\nu_n}(s))|^2 \right] ds. \end{aligned}$$

We continue by using the following lemma, whose proof is given after the proof of the supersolution property at the end of this section:

Lemma 29. *Let $z_0 := (x_0, \varphi(t_0, x_0))$ and $\psi : [0, T] \times \mathbb{R}^{d+1} \times U \rightarrow \mathbb{R}$ be locally Lipschitz in $(t, z) \in [0, T] \times \mathbb{R}^{d+1}$ uniformly in $r \in U$. Then*

$$\frac{1}{h_n} \int_{t_n}^{t_n+h_n} [\psi(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \psi(t_0, z_0, \nu_n(s))] ds \rightarrow 0 \quad \mathbf{P} - a.s.$$

along some subsequence.

In view of this lemma,

$$\begin{aligned} 0 \leq \liminf_{n \rightarrow \infty} \frac{1}{h_n} \int_{t_n}^{t_n+h_n} & \left[b(t_0, z_0, \nu_n(s)) - \mathcal{L}^{\nu_n(s)} \varphi(t_0, x_0) \right. \\ & \left. - \xi |a(t_0, z_0, \nu_n(s)) - \sigma(t_0, x_0, \nu_n(s)) D\varphi(t_0, x_0)|^2 \right] ds. \end{aligned}$$

Then, since $h_n^{-1} \int_{t_n}^{t_n+h_n} ds = 1$,

$$\begin{aligned} \frac{1}{h_n} \int_{t_n}^{t_n+h_n} & \left[b(t_0, z_0, \nu_n(s)) - \mathcal{L}^{\nu_n(s)} \varphi(t_0, x_0) \right. \\ & \left. - \xi |a(t_0, z_0, \nu_n(s)) - \sigma(t_0, x_0, \nu_n(s)) D\varphi(t_0, x_0)|^2 \right] ds \in \bar{c} \circ \mathcal{V}(t_0, z_0), \end{aligned} \tag{32}$$

where $\bar{\text{co}} \mathcal{V}(t_0, z_0)$ is the closed convex hull of the set

$$\mathcal{V}(t_0, z_0) := \left\{ b(t_0, z_0, \nu) - \mathcal{L}^\nu \varphi(t_0, x_0) - \xi |a(t_0, z_0, \nu) - \sigma(t_0, x_0, \nu) D\varphi(t_0, x_0)|^2 : \nu \in U \right\}.$$

Therefore, it follows from (32) that

$$\begin{aligned} 0 &\leq \sup_{\phi \in \bar{\text{co}} \mathcal{V}(t_0, z_0)} \phi \\ &= \sup_{\nu \in U} \left\{ \xi | -a(t_0, z_0, \nu) + \sigma(t_0, x_0, \nu) D\varphi(t_0, x_0) |^2 - \mathcal{L}^\nu \varphi(t_0, x_0) + b(t_0, z_0, \nu) \right\} \end{aligned} \quad (33)$$

for all $\xi \in \mathbb{R}$.

Step 3. For a large positive integer n , set $\xi = -n$. Since U is compact, the supremum in (33) is attained at some $\tilde{\nu}_n \in U$, and

$$-n | -a(t_0, z_0, \tilde{\nu}_n) + \sigma(t_0, x_0, \tilde{\nu}_n) D\varphi(t_0, x_0) |^2 - \mathcal{L}^{\tilde{\nu}_n} \varphi(t_0, x_0) + b(t_0, z_0, \tilde{\nu}_n) \geq 0.$$

By passing to a subsequence, we may assume that there exists $\nu_0 \in U$ such that $\tilde{\nu}_n \rightarrow \nu_0$. Now let n tend to infinity in the last inequality to prove that

$$|a(t_0, z_0, \tilde{\nu}_n) - \sigma(t_0, x_0, \tilde{\nu}_n) D\varphi(t_0, x_0)|^2 \rightarrow 0 \quad (34)$$

and

$$-\mathcal{L}^{\nu_0} \varphi(t_0, x_0) + b(t_0, z_0, \nu_0) \geq 0. \quad (35)$$

In view of (34) and (26) we conclude that

$$\nu_0 = \psi(t_0, z_0, D\varphi(t_0, x_0)). \quad (36)$$

Since $\nu_0 \in U$, it follows from Remark 23 that

$$\inf_{\zeta \in \tilde{U}_1} (\delta_U(\zeta) - \zeta^* \nu_0) \geq 0. \quad (37)$$

The supersolution property follows from (35), (36) and (37).

Proof of Lemma 29. Since $\psi(t, z, r)$ is locally Lipschitz in (t, z) uniformly in r ,

$$\begin{aligned} &\frac{1}{h_n} \int_{t_n}^{t_n+h_n} [\psi(s, Z_{t_n, z_n}^{\nu_n}(s), \nu_n(s)) - \psi(t_0, z_0, \nu_n(s))] ds \\ &\leq K \frac{1}{h_n} \int_{t_n}^{t_n+h_n} (|s - t_0| + |Z_{t_n, z_n}^{\nu_n}(s) - z_0|) ds \\ &\leq K \left(h_n + |t_n - t_0| + \sup_{t_n \leq s \leq t_n+h_n} |Z_{t_n, z_n}^{\nu_n}(s) - z_0| \right) \end{aligned}$$

for some constant K . Thus, to complete the proof of this lemma, it suffices to show

$$\sup_{t_n \leq s \leq t_n+h_n} |Z_{t_n, z_n}^{\nu_n}(s) - z_0| \rightarrow 0 \quad \mathbf{P} - \text{a.s.}$$

along a subsequence. Set

$$\gamma(t, x, y, r) := \begin{pmatrix} \mu(t, x, r) \\ b(t, x, y, r) \end{pmatrix} \quad \text{and} \quad \alpha(t, x, y, r) := \begin{pmatrix} \sigma^*(t, x, r) \\ a^*(t, x, y, r) \end{pmatrix}$$

The functions α and γ inherit the pointwise bounds from μ, b, σ and a . We directly calculate that, for $t_n \leq s \leq t_n + h_n$,

$$Z_{t_n, z_n}^{\nu_n}(s) - z_0 \leq |z_n - z_0| + \|\gamma\|_\infty h_n + \left| \int_{t_n}^s \alpha(r, Z_{t_n, z_n}^{\nu_n}(r), \nu_n(r)) dW(r) \right|,$$

and, therefore,

$$\begin{aligned} \sup_{t_n \leq s \leq t_n + h_n} |Z_{t_n, z_n}^{\nu_n}(s) - z_0| &\leq |z_n - z_0| + \|\gamma\|_\infty h_n \\ &+ \sup_{t_n \leq s \leq t_n + h_n} \left| \int_{t_n}^s \alpha(r, Z_{t_n, z_n}^{\nu_n}(r), \nu_n(r)) dW(r) \right|. \end{aligned}$$

The first two on the right-hand side converge to zero. We estimate the third term by Doob's maximal inequality for submartingales. The result is

$$\begin{aligned} E \left[\sup_{t_n \leq s \leq t_n + h_n} \left| \int_{t_n}^s \alpha(r, Z_{t_n, z_n}^{\nu_n}(r), \nu_n(r)) dW(r) \right|^2 \right] \\ \leq 4E \left[\int_{t_n}^{t_n + h_n} \|\alpha(r, Z_{t_n, z_n}^{\nu_n}(r), \nu_n(r))\|^2 dr \right] \\ \leq 4\|\alpha\|_\infty^2 h_n. \end{aligned}$$

This proves that

$$\sup_{t_n \leq s \leq t_n + h_n} |Z_{t_n, z_n}^{\nu_n}(s) - z_0| \rightarrow 0 \quad \text{in } L^2(\mathbf{P}),$$

and, therefore, it also converges \mathbf{P} -a.s. along some subsequence. \square

4.4 Proof of the viscosity subsolution property

We start with a technical lemma which will be used both in the proof of the subsolution property and also in the next subsection on the characterization of the terminal data. We first introduce some notation: given a smooth function $\varphi(t, x)$, we define the open subset of $[0, T] \times \mathbb{R}^d$

$$\begin{aligned} \mathcal{M}_0(\varphi) := \{ (t, x) \in [0, T] \times \mathbb{R}^d : \nu_0(t, x) \in \text{int}(U) \text{ and} \\ -\mathcal{L}^{\nu_0(t, x)} \varphi(t, x) + b(t, x, \varphi(t, x), \nu_0(t, x)) > 0 \}, \end{aligned}$$

where $\nu_0(t, x) := \psi(t, x, \varphi(t, x), D\varphi(t, x))$.

Lemma 30. *Let φ be a smooth test function and suppose that there are $t_1 < t_2 \leq T$ such that*

$$\text{cl}(\mathcal{M}) \subset \mathcal{M}_0(\varphi), \quad \text{where } \mathcal{M} := (t_1, t_2) \times B_R(x_0).$$

Then

$$\sup_{\partial_p \mathcal{M}} (v - \varphi) = \max_{\text{cl}(\mathcal{M})} (v^* - \varphi),$$

where $\partial_p \mathcal{M}$ is the parabolic boundary of \mathcal{M} , i.e., $\partial_p \mathcal{M} = ([t_1, t_2] \times \partial B_R(x_0)) \cup (\{t_2\} \times \overline{B_R(x_0)})$.

Proof. We shall denote $\overline{\mathcal{M}} := \text{cl}(\mathcal{M})$. Suppose, to the contrary, that

$$\max_{\overline{\mathcal{M}}} (v^* - \varphi) - \sup_{\partial_p \mathcal{M}} (v - \varphi) := 2\beta > 0,$$

and let us work toward a contradiction of the dynamic programming principle **(DP2)**: choose $(t_0, x_0) \in \mathcal{M}$ so that $(v - \varphi)(t_0, x_0) \geq -\beta + \max_{\overline{\mathcal{M}}}(v^* - \varphi)$ and

$$(v - \varphi)(t_0, x_0) \geq \beta + \sup_{\partial_p \mathcal{M}}(v - \varphi) \quad (38)$$

Step 1. In view of Remark 27, $\inf_{\zeta \in \tilde{U}_1}(\delta_U(\zeta) - \zeta^* \nu_0)$ is equivalent to $\nu_0 \in \text{int}(U)$. Set

$$\mathcal{N} := \{(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} : \hat{\nu}(t, x, y) \in \text{int}(U) \text{ and} \\ -\mathcal{L}^{\hat{\nu}_0(t, x, y)} \varphi(t, x) + b(t, x, y, \hat{\nu}(t, x, y)) > 0\},$$

where $\hat{\nu}(t, x, y) = \psi(t, x, y, D\varphi(t, x))$ and, for $\eta \geq 0$,

$$\mathcal{M}_\eta := \{(t, x) \in [0, T] \times \mathbb{R}^d : (t, x, \varphi(t, x) - \eta) \in \mathcal{N}\}.$$

Note that this definition of $\mathcal{M}_0 := \mathcal{M}_0(\varphi)$ agrees with the previous definition. Moreover, in view of our hypothesis, for all sufficiently small η we have $\overline{\mathcal{M}} \subset \mathcal{M}_\eta$. Fix $\eta \leq \beta$ satisfying this inclusion.

Step 2. Let η be as in the previous step. Let (X_η, Y_η) be the solution of the state equation with initial data $X_\eta(t_0) = x_0, Y_\eta(t_0) = v(t_0, x_0) - \eta$ and the control ν given in the feedback form

$$\nu(t, x) = \psi(t, x, \varphi(t, x) - \eta, D\varphi(t, x)).$$

Set $\nu(t) := \nu(t, X_\eta(t))$ so that

$$(X_\eta, Y_\eta) = Z_{t_0, x_0, v(t_0, x_0) - \eta}^\nu = (X_{t_0, x_0}^\nu, Y_{t_0, x_0, v(t_0, x_0) - \eta}^\nu).$$

Set

$$\widehat{Y}_\eta(t) := \varphi(t, X_\eta(t)) - \eta + (v - \varphi)(t_0, x_0),$$

and observe that $Y_\eta(0) = \widehat{Y}_\eta(0) = v(t_0, x_0) - \eta$. In the next step, we will compare the processes Y_η and \widehat{Y}_η .

Step 3. By Itô's rule,

$$d\widehat{Y}_\eta(t) = \mathcal{L}^{\nu(t)} \varphi(t, X_\eta(t)) dt + D\varphi(t, X_\eta(t)) \cdot \sigma(t, X_\eta(t), \nu(t))^* dW(t).$$

In view of (26) and the definition of $\nu(t)$,

$$D\varphi(t, X_\eta(t)) \cdot \sigma(t, X_\eta(t), \nu(t))^* = a(t, X_\eta(t), \widehat{Y}_\eta(t), \nu(t))^*.$$

Hence

$$d\widehat{Y}_\eta = \hat{b}(t) dt + a(t, X_\eta(t), Y_\eta(t), \nu(t))^* dW(t),$$

where $\hat{b}(t) := \mathcal{L}^{\nu(t)} \varphi(t, X_\eta(t))$. Recall that Y_η solves the same SDE with a different drift term:

$$dY_\eta(t) = b(t) dt + a(t, X_\eta(t), Y_\eta(t), \nu(t))^* dW(t),$$

where $b(t) := b(t, X_\eta(t), Y_\eta(t), \nu(t))$. Let θ be the stopping time

$$\theta := \inf\{s > 0 : (t_0 + s, X_\eta(t_0 + s)) \notin \mathcal{M}\}.$$

Since \mathcal{M} is an open set containing (t_0, x_0) , the stopping time θ is positive \mathbf{P} -a.s.

Now, from the definition of η , we have $\mathcal{M} \subset \mathcal{M}_\eta$. It follows that, for $t \in [t_0, t_0 + \theta)$, $(t, X_\eta(t)) \in \mathcal{M}_\eta$ \mathbf{P} -a.s., i.e. $(t, X_\eta(t), \widehat{Y}_\eta(t)) \in \mathcal{N}$ \mathbf{P} -a.s. by definition of \mathcal{M}_η . Hence

$$b(t) > \mathcal{L}^{\nu(t)}\varphi(t, X_\eta(t)) = \hat{b}(t), \quad t \in [t_0, t_0 + \theta), \quad \mathbf{P} - \text{a.s.}$$

Since $Y_\eta(0) = \widehat{Y}_\eta(0) = v(t_0, x_0) - \eta$, it follows from stochastic comparison (see, for instance, [KA/SH 91], Proposition 5.2.18) that

$$\widehat{Y}_\eta(t) \leq Y_\eta(t), \quad t \in [t_0, t_0 + \theta), \quad \mathbf{P} - \text{a.s.}$$

Step 4. We now proceed to contradict (DP2). First, observe that by continuity of the process X_η , $(t_0 + \theta, X_\eta(t_0 + \theta)) \in \partial_p \mathcal{M}$ \mathbf{P} -a.s. Also, from inequality (38), we have $v \leq \varphi - \beta + (v - \varphi)(t_0, x_0)$ on $\partial_p \mathcal{M}$. Therefore,

$$\begin{aligned} Y_\eta(t_0 + \theta) - v(t_0 + \theta, X_\eta(t_0 + \theta)) &\geq \beta + Y_\eta(t_0 + \theta) - \varphi(t_0 + \theta, X_\eta(t_0 + \theta)) \\ &\quad + (v - \varphi)(t_0, x_0) \\ &= (\beta - \eta) + Y_\eta(t_0 + \theta) - \widehat{Y}_\eta(t_0 + \theta) \\ &= \beta - \eta \geq 0 \end{aligned}$$

from Step 3. By (38) and the definition of (X_η, Y_η) , we have $Y_\eta = Y_{t_0, x_0, v(t_0, x_0) - \eta}^\nu$ and $X_\eta = X_{t_0, x_0}^\nu$. Then the previous inequality contradicts (DP2). \square

Proof of the viscosity subsolution property. Fix $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$, and let φ be a $\mathcal{C}^2([0, T] \times \mathbb{R}^d)$ function satisfying

$$(v^* - \varphi)(t_0, x_0) = (\text{strict}) \max_{(t, x) \in [0, T] \times \mathbb{R}^d} (v^* - \varphi).$$

Set $z_0 := (x_0, \varphi(t_0, x_0))$. Let $\mathcal{M}_0 := \mathcal{M}_0(\varphi)$ be as in the previous lemma. Since (t_0, x_0) is a strict maximizer of $(v^* - \varphi)$ and since \mathcal{M}_0 is an open set, by the previous lemma we conclude that $(x_0, y_0) \notin \mathcal{M}_0$. Then, by the definition of \mathcal{M}_0 ,

$$\min \left\{ \inf_{\zeta \in \bar{U}_1} (\delta_U(\zeta) - \zeta^* \hat{\nu}_0(t_0, z_0)), -\mathcal{L}^{\hat{\nu}_0(t_0, z_0)}\varphi(t_0, x_0) + b(t_0, z_0, \hat{\nu}_0(t_0, z_0)) \right\} \leq 0,$$

and therefore v^* is a viscosity subsolution of (27). \square

4.5 Terminal condition

To characterize the value function as the unique solution of the dynamic programming equation, we need to specify the terminal data. The definition of the value function implies that

$$v(T, x) = g(x), \quad x \in \mathbb{R}^d.$$

However, it is known that

$$\underline{G}(x) := \liminf_{t \uparrow T, x' \rightarrow x} v(t, x')$$

may be strictly larger than $g(x)$ (see, for instance, [B/C/S 98] and lemma 33 below).

In this section we will characterize \underline{G} as the viscosity supersolution of a first order PDE. We will also study

$$\overline{G}(x) := \limsup_{t \uparrow T, x' \rightarrow x} v(t, x')$$

and prove that \overline{G} is a viscosity subsolution of the same equation. More precisely, we have the following theorem.

Theorem 31. *Let the assumptions of Theorem 26 hold, and assume that \underline{G} and \overline{G} are finite for every $x \in \mathbb{R}^d$. Suppose, further that $(g_*)^* \geq g$. Then \underline{G} is viscosity supersolution of the first order PDE on \mathbb{R}^d*

$$\min \{G(x) - g_*(x); H(T, x, G(x), DG(x))\} = 0$$

and \overline{G} is viscosity subsolution of the first order PDE

$$\min \{G(x) - g^*(x); H(T, x, G(x), DG(x))\} = 0.$$

In most cases, since a subsolution is not greater than the supersolution, this implies that $\overline{G} \leq \underline{G}$ and therefore that $\overline{G} = \underline{G}$. In the next section, we provide examples for which this holds, and we will also compute $G := \overline{G} = \underline{G}$ explicitly in those examples.

Remark 32. In the definition of \overline{G} , we may replace v by v^* :

$$\overline{G}(x) = \limsup_{t \uparrow T, x' \rightarrow x} v^*(t, x').$$

Similarly,

$$\underline{G}(x) := \liminf_{t \uparrow T, x' \rightarrow x} v_*(t, x').$$

The rest of this section is devoted to the proof of Theorem 31. We need first the following lemma:

Lemma 33. *Suppose that $\underline{G}(x)$ and $\overline{G}(x)$ are finite for every $x \in \mathbb{R}^d$. Then*

$$\underline{G}(x) \geq g_*(x) \quad \text{for all } x \in \mathbb{R}^d.$$

Proof. Take a sequence $(t_n, x_n) \rightarrow (T, x)$ with $t_n < T$. Set $y_n := v(t_n, x_n) + (1/n)$. Then, for each n there exists a control $\nu_n \in \mathcal{A}$ satisfying

$$Y_{t_n, x_n, y_n}^{\nu_n}(T) \geq g(X_{t_n, x_n}^{\nu_n}(T)) \quad \mathbf{P} - \text{a.s.}$$

Since a and b are bounded,

$$E [Y_{t_n, x_n, y_n}^{\nu_n}(T)] \leq y_n + \|b\|_\infty(T - t_n) = v(t_n, x_n) + \frac{1}{n} + \|b\|_\infty(T - t_n).$$

We continue by using the following claim, whose proof will be provided later:

$$\{Y_{t_n, x_n, y_n}^{\nu_n}(T), n \geq 0\} \quad \text{is uniformly integrable.} \tag{39}$$

Then, by Fatou's lemma

$$\begin{aligned} \liminf_{n \rightarrow \infty} v(t_n, x_n) &\geq \liminf_{n \rightarrow \infty} E [Y_{t_n, x_n, y_n}^{\nu_n}(T)] \\ &= E \left[\liminf_{n \rightarrow \infty} Y_{t_n, x_n, y_n}^{\nu_n}(T) \right] \\ &\geq E \left[\liminf_{n \rightarrow \infty} g(X_{t_n, x_n}^{\nu_n}(T)) \right]. \end{aligned}$$

Since U is compact and (t_n, x_n) converges to (T, x) , $X_{t_n, x_n}^{\nu_n}(T)$ approaches x as n tends to infinity. The required result then follows from the definition of the lower semicontinuous envelope g_* of g .

It remains to prove the claim (39). Since b is bounded,

$$\begin{aligned} |Y_{t_n, x_n, y_n}^{\nu_n}(T)| &\leq |y_n| + (T - t_n) \|b\|_\infty + \left| \int_{t_n}^T a(u, Z_{t_n, x_n, y_n}^{\nu_n}(u), \nu_n(u))^* dW(u) \right| \\ &\leq |v(t_n, x_n)| + T \|b\|_\infty + \left| \int_{t_n}^T a(u, Z_{t_n, x_n, y_n}^{\nu_n}(u), \nu_n(u))^* dW(u) \right|. \end{aligned}$$

Now observe that

$$\limsup_{n \rightarrow \infty} v(t_n, x_n) \leq \limsup_{n \rightarrow \infty} v^*(t_n, x_n) \leq \overline{G}(x)$$

and

$$\liminf_{n \rightarrow \infty} v(t_n, x_n) \geq \liminf_{n \rightarrow \infty} v_*(t_n, x_n) \leq \underline{G}(x).$$

This proves that the sequence $v(t_n, x_n)$ is bounded. In order to complete the proof, it suffices to show that the sequence

$$U_n := \int_{t_n}^T a(u, Z_{t_n, x_n, y_n}^{\nu_n}(u), \nu_n(u))^* dW(u), \quad n \geq 0$$

is uniformly integrable, Since a is bounded,

$$\sup_{n \geq 0} E[U_n^2] \leq \sup_{n \geq 0} (T - t_n) \|a^* a\|_\infty \leq T \|a^* a\|_\infty.$$

Hence $\{U_n, n \geq 0\}$ is bounded in L^2 , and, therefore, it is uniformly integrable. \square

Next, we will show that \underline{G} is a viscosity supersolution of $H = 0$, where H is as in (29):

Lemma 34. *Suppose that $\underline{G}(x)$ is finite for every $x \in \mathbb{R}^d$. Then \underline{G} is a viscosity supersolution of*

$$H(T, x, G(x), DG(x)) = 0.$$

Proof. By definition, \underline{G} is lower semicontinuous. Let f be a $\mathcal{C}^2(\mathbb{R}^d)$ -function satisfying

$$0 = (\underline{G} - f)(x_0) = \min_{x \in \mathbb{R}^d} (\underline{G} - f)$$

at some $x_0 \in \mathbb{R}^d$. Observe that $\underline{G} \geq f$ on \mathbb{R}^d .

Step 1. In view of Remark 32, there exists a sequence (s_n, ξ_n) converging to (T, x_0) such that $s_n < T$ and

$$\lim_{n \rightarrow \infty} v_*(s_n, \xi_n) = \underline{G}(x_0).$$

For a positive integer n , consider the auxiliary test function

$$\varphi_n(t, x) := f(x) - \frac{1}{2}|x - x_0|^2 + \frac{T - t}{(T - s_n)^2}.$$

Let $B := B_1(x_0)$ be the unit ball in \mathbb{R}^d centered at x_0 . Choose $(t_n, x_n) \in [s_n, T] \times \overline{B}$, which maximizes the difference $v_* - \varphi_n$ on $[s_n, T] \times \overline{B}$.

Step 2. We claim that, for sufficiently large n , $t_n < T$, and x_n converges to x_0 : indeed, for sufficiently large n ,

$$(v_* - \varphi_n)(s_n, \xi_n) \leq -\frac{1}{2(T - s_n)}.$$

On the other hand, for any $x \in \overline{B}$,

$$(v_* - \varphi_n)(T, x) = \underline{G}(x) - f(x) + \frac{1}{2}|x - x_0|^2 \geq \underline{G}(x) - f(x) \geq 0.$$

Comparing the two inequalities leads us to conclude that $t_n < T$ for large n . Suppose that, on a subsequence, x_n converges to x^* . Since $t_n \geq s_n$ and (t_n, x_n) minimizes the difference $(v_* - \varphi)$,

$$\begin{aligned} & (\underline{G} - f)(x^*) - (\underline{G} - f)(x_0) \\ & \leq \liminf_{n \rightarrow \infty} (v_* - \varphi_n)(t_n, x_n) - (v_* - \varphi_n)(s_n, \xi_n) - \frac{1}{2}|x_n - x_0|^2 \\ & \leq \limsup_{n \rightarrow \infty} (v_* - \varphi_n)(t_n, x_n) - (v_* - \varphi_n)(s_n, \xi_n) - \frac{1}{2}|x_n - x_0|^2 \\ & \leq -\frac{1}{2}|x^* - x_0|^2. \end{aligned}$$

Since x_0 minimizes the difference $\underline{G} - f$,

$$0 \leq (\underline{G} - f)(x^*) - (\underline{G} - f)(x_0) \leq -\frac{1}{2}|x^* - x_0|^2.$$

Hence $x^* = x_0$. The above argument also proves that

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} (v_* - \varphi_n)(t_n, x_n) - (v_* - \varphi_n)(s_n, \xi_n) \\ &= -\underline{G}(x_0) + \lim_{n \rightarrow \infty} v_*(t_n, x_n) + \frac{(T - s_n) - (T - t_n)}{(T - s_n)^2} \\ &\geq -\underline{G}(x_0) + \limsup_{n \rightarrow \infty} v_*(t_n, x_n). \end{aligned}$$

This proves that $\limsup_{n \rightarrow \infty} v_*(t_n, x_n) \leq \underline{G}(x_0)$. Since

$$\limsup_{n \rightarrow \infty} v_*(t_n, x_n) \geq \liminf_{n \rightarrow \infty} v_*(t_n, x_n) \geq \underline{G}(x_0),$$

by definition of \underline{G} , we obtain that

$$\lim_{n \rightarrow \infty} v_*(t_n, x_n) = \underline{G}(x_0). \tag{40}$$

This implies that, for all sufficiently large n , (t_n, x_n) is a local minimizer of the difference $(v_* - \varphi_n)$. In view of the general theory of viscosity solutions (see, for instance, [FL/SO 93]), the viscosity property of v_* holds at (t_n, x_n) .

Step 3. We now use the viscosity property of v_* in $[0, T) \times \mathbb{R}^d$: for every n ,

$$H(t_n, x_n, v_*(t_n, x_n), D\varphi(t_n, x_n)) \geq 0.$$

Note that $D\varphi_n(t_n, x_n) = Df(t_n, x_n) - (x_n - x_0)$, and recall that H is continuous; see Remark 25. Since (t_n, x_n) tends to (T, x_0) , (40) implies that

$$H(T, x_0, \underline{G}(x_0), Df(x_0)) \geq 0.$$

□

These results imply that \underline{G} is a viscosity supersolution of

$$\min \{G(x) - g_*(x); H(T, x, G(x), DG(x))\} = 0, \quad (41)$$

proving the first part of Theorem 31. The following result concludes the proof of the theorem,

Lemma 35. *Suppose that $\underline{G}(x)$ and $\overline{G}(x)$ are finite for every $x \in \mathbb{R}^d$ and that $(g_*)^* \geq g$. Then \overline{G} is a viscosity subsolution on \mathbb{R}^d of*

$$\min \{G(x) - g_*(x); H(T, x, G(x), DG(x))\} = 0.$$

Proof. By definition, \overline{G} is upper semicontinuous. Let $x_0 \in \mathbb{R}^d$ and $f \in C^2(\mathbb{R}^d)$ satisfy

$$0 = (\overline{G} - f)(x_0) = \max_{x \in \mathbb{R}^d} (\overline{G} - f).$$

We need to show that, if $\overline{G}(x_0) > g^*(x_0)$, then

$$H(T, x_0, \underline{G}(x_0), D\underline{G}(x_0)) \leq 0. \quad (42)$$

So we assume that

$$\overline{G}(x_0) > g^*(x_0). \quad (43)$$

For a positive integer n , set $s_n := T - \frac{1}{n^2}$, and consider the auxiliary test function

$$\varphi_n(t, x) := f(x) + \frac{1}{2}|x - x_0|^2 + n(T - t), \quad (t, x) \in [s_n, T] \times \mathbb{R}^d.$$

In order to obtain the required result, we will first prove that the test function φ_n does not satisfy the condition of Lemma 23 on $[s_n, T] \times B_R(x_0)$ for some $R > 0$, and then we will pass to the limit as $n \rightarrow \infty$.

Step 1. By definition, $\overline{G} \geq \underline{G}$. From lemma 33, this provides $\overline{G} \geq g_*$ and then $\overline{G} \geq (g_*)^*$ by uppersemicontinuity of \overline{G} . Hence, by assumption of the lemma,

$$\overline{G} \geq g. \quad (44)$$

This proves that $(v - \varphi_n)(T, x) = (g - f)(x) - |x - x_0|^2/2 \leq (\overline{G} - f)(x) \leq 0$ by definition of the test function f . Then, for all $R > 0$,

$$\sup_{B_R(x_0)} (v - \varphi_n)(T, \cdot) \leq 0.$$

Now suppose that there exists a subsequence of (φ_n) , still denoted by (φ_n) , such that

$$\lim_{n \rightarrow \infty} \sup_{B_R(x_0)} (v - \varphi_n)(T, \cdot) = 0,$$

and let us work toward a contradiction. For each n , let $(x_n^k)_k$ be a maximizing sequence of $(v - \varphi_n)(T, \cdot)$ on $B_R(x_0)$, i.e.,

$$\lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} (v - \varphi_n)(T, x_n^k) = 0.$$

Then it follows from (44) that $(v - \varphi_n)(T, x_n^k) \leq -|x - x_0|^2/2$, which provides

$$\lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} x_n^k = x_0.$$

Therefore,

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} (v - \varphi_n)(T, x_n^k) = \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} g(x_n^k) - f(x_0) \\ &\leq \limsup_{x \rightarrow x_0} g(x) - f(x_0) = (g^* - f)(x_0) < (\overline{G} - f)(x_0) \end{aligned}$$

by (43), but this cannot happen since $(\overline{G} - f)(x_0) = 0$. The consequence of this is

$$\limsup_{n \rightarrow \infty} \sup_{B_R(x_0)} (v - \varphi_n)(T, \cdot) < 0, \quad \text{for all } R > 0. \quad (45)$$

Step 2. Let (t_n, x_n) be a maximizing sequence of $(v_* - \varphi_n)$ on $[s_n, T] \times \partial B_R(x_0)$. Then, since $T - t_n \leq T - s_n = n^{-2}$,

$$\limsup_{n \rightarrow \infty} \sup_{[s_n, T] \times \partial B_R(x_0)} (v^* - \varphi_n) \leq \limsup_{n \rightarrow \infty} (v^*(t_n, x_n) - f(x_n)) - \frac{1}{2}R^2.$$

Since $t_n \rightarrow T$ and, after passing to a subsequence, $x_n \rightarrow x^*$ for some $x^* \in \partial B_R(x_0)$, we get

$$\limsup_{n \rightarrow \infty} \sup_{[s_n, T] \times \partial B_R(x_0)} (v^* - \varphi_n) \leq (\overline{G} - f)(x^*) - \frac{1}{2}R^2 \leq -\frac{1}{2}R^2.$$

This, together with (45), implies that for all $R > 0$ there exists $n(R)$ such that, for all $n > n(R)$,

$$\max\{(v - \varphi_n) : \partial_p((s_n, T) \times B_R(x_0))\} < 0 = (v^* - \varphi_n)(T, x_0).$$

Hence, it follows from Lemma 30 that

$$(s_n, T) \times B_R(x_0) \text{ is not a subset of } \mathcal{M}_0(\varphi_n) \quad \text{for all } n > n(R). \quad (46)$$

Step 3. Observe that, for all $\nu \in U$ and (t, x, y) ,

$$-\mathcal{L}^\nu \varphi_n(t, x) = n - \mathcal{L}^\nu f(x) - \mu(t, x, \nu)^*(x - x_0) - \frac{1}{2} \text{Tr}[\sigma^* \sigma](t, x, \nu) > b(t, x, y, \nu),$$

provided that n is sufficiently large. Then, for large n ,

$$\begin{aligned} &\mathcal{M}_0(\varphi_n) \cap ((s_n, T) \times B_R(x_0)) \\ &= \{(t, x) \in (s_n, T) \times B_R(x_0) : H(t, x, \varphi_n(t, x), D\varphi_n(t, x)) > 0\}. \end{aligned}$$

In view of this, it follows from (46) that there exists a sequence (t_n, x_n) converging to (T, x_0) such that

$$H(t_n, x_n, \varphi_n(t_n, x_n), D\varphi_n(t_n, x_n)) \leq 0.$$

We now let n tend to infinity to obtain (42). □

5 Hedging with portfolio constraints and large investors

The celebrated papers of Black and Scholes [BL/SCH 73] and Merton [MERT 73] paved the way for pricing options on stocks, based on the following principle: In a *complete market* every contingent claim can be exactly replicated at the terminal time by investing wisely in the market and starting with a large enough initial capital. Thus, the "fair price" of the claim is taken to be the minimal

such capital, which coincides with the expectation of the claim's discounted value under the unique, "risk-neutral" equivalent probability measure. The argument that leads to this result, and to the associated "valuation formulae", is based on the martingale representation and Girsanov theorems from stochastic analysis.

The foregoing argument fails, unfortunately, in the presence of *constraints* on portfolio choice, e.g. constraints on borrowing, on short-selling of stocks, even on accessing certain stocks at all, as in the case of "incomplete markets". However, in such markets it is often the case that, with sufficient initial wealth, a hedging agent can construct a portfolio which respects the constraints and still leads to a final wealth that *super-replicates* -dominates almost surely- the payoff of the contingent claim.

The idea of super-replication (or super-hedging) was first suggested by El Karoui and Quenez [EK/QU 95], and in this case, there is no risk (for the hedger) associated with the contingent claim, as the super-replicating price is the smallest initial capital that allows the seller to construct a portfolio which dominates almost surely the payoff at the terminal time.

Another fundamental assumption that it is removed from the usual continuous-time model of the stock market prices is the so-called 'small' investor assumption: the classical Black-Scholes model considered in mathematical finance assumes perfect elasticity for the supply and demand of traded assets so that orders of arbitrary size do not affect asset prices. This assumption is justified as long as one considers 'small' investors whose trading volume is easily covered by market liquidity. However, if there is a 'large' investor in the market, whose orders involve a significant part of the available shares, market prices will no longer evolve independently of the trading strategies chosen by this investor

Mathematically speaking, under the classical small large investor framework the coefficients of the price equations are independent of the wealth and portfolio process of the investor, but here we will consider also the case in which the influence of the investor's financial behavior is *not* a priori known to be irrelevant and the price model is *not* necessarily linear. In other words, the mean rate of return and volatility coefficients can both be nonlinear in the price process and also depend on the portfolio process ν of the investor.

The financial market. We will consider a financial market consisting of

- a non-risky asset (bond) with price process S^0 normalized to the unity, i.e. $S^0 \equiv 1$.
- d risky assets (stocks) with positive prices S^i , $i = 1, \dots, d$.

The normalization of the non-risky asset to the unity is, as usual, obtained by discounting, i.e. taking the non-risky asset as a *numéraire*.

A portfolio strategy is an \mathbb{F} -adapted process $\nu = \{\nu(t), t \in [0, T]\}$ with values in a closed and convex set $U \subset \mathbb{R}^d$, which represents the constraints on portfolio choice. At each time $t \in [0, T]$, $\nu^i(t)$ is the fraction of wealth invested in the risky asset S^i . The set of all portfolio strategies is denoted by \mathcal{A} .

The so-called *self-financing* condition states that the variation of the wealth process is only affected by the variation of the price process. So, under this condition, given an initial capital $\tilde{y} > 0$ and a portfolio strategy ν , the wealth process \tilde{Y} is defined by

$$d\tilde{Y}_{\tilde{y}}^{\nu}(t) = \sum_{i=1}^d \tilde{Y}_{\tilde{y}}^{\nu}(t) \nu^i(t) \frac{dS^i(t)}{S^i(t)}, \quad \text{with } \tilde{Y}_{\tilde{y}}^{\nu}(0) = \tilde{y}.$$

As we consider a “large investor” model for the stock prices, which are furthermore assumed to be positive, we will write the stocks prices in the “exponential” form $S^i = \exp(X^\nu)^i$, where $(X^\nu)^i$ is the i -th component of the d -dimensional process X^ν solution of the SDE

$$dX^\nu(t) = \mu(t, X^\nu(t), \nu(t)) dt + \sigma(t, X^\nu(t), \nu(t))^* dW(t), \quad X^\nu(0) = x.$$

We will also consider the log-wealth process

$$Y_y^\nu(t) := \ln \tilde{Y}_y^\nu(t), \quad \text{with } Y_y^\nu(0) = y := \ln \tilde{y}.$$

Then, a direct application of Itô’s lemma provides

$$dY_y^\nu(t) = b(t, X^\nu(t), \nu(t)) dt + \nu(t)^* \sigma(t, X^\nu(t), \nu(t)) dW(t),$$

where

$$b(t, x, r) = r^* \mu(t, x, r) + \frac{1}{2} \text{Tr}[(\sigma^* \sigma)(t, x, r) \text{diag}[r]] - \frac{1}{2} |\sigma(t, x, r) r|^2$$

and $\text{diag}[r]$ is the $d \times d$ -matrix with diagonal $(r^1, \dots, r^d)^*$. Let $f : \mathbb{R}^d \rightarrow [0, \infty)$ be a measurable function. The super-replication price is then defined by

$$\tilde{v}(0, S(0)) := \inf \left\{ \tilde{y} > 0 : \exists \nu \in \mathcal{A}, \tilde{Y}_y^\nu(T) \geq f(S(T)) \text{ } \mathbf{P} - a.s. \right\}.$$

Here $f(S(T))$ is a contingent claim. The value function is then the minimal initial capital which allows the seller of the contingent claim to face the promised payoff $f(S(T))$ through some clever portfolio strategy $\nu \in \mathcal{A}$.

To see that the super-replication problem belongs to the general class of stochastic target problems studied in the previous sections, we introduce

$$v(t, x) := \ln \tilde{v}(t, s) \quad \text{and} \quad g(x) := \ln f(s),$$

where $s := (e^{x^1}, \dots, e^{x^d})$. Here the s -variable stands for the stocks price process S and x for the process X^ν . With this change of variable we get

$$v(0, X^\nu(0)) := \inf \left\{ y \in \mathbb{R} : \exists \nu \in \mathcal{A}, Y_y^\nu(T) \geq g(X^\nu(T)) \text{ } \mathbf{P} - a.s. \right\}.$$

In the small investor framework, when σ and μ do not depend on the portfolio strategy, the super-replication problem is usually reduced via convex duality to a stochastic control problem in standard form. Indeed, if we define

$$\mathcal{D} := \{ \text{bounded } \mathbb{F} - \text{adapted processes with values in } \tilde{U} \},$$

and for each $\nu \in \mathcal{D}$, the equivalent probability measure \mathbf{P}^ν with density

$$\begin{aligned} \frac{d\mathbf{P}^\nu}{d\mathbf{P}} \Big|_{\mathcal{F}(T)} &:= \exp \left\{ \int_0^T [\nu(t) - \tilde{\mu}(t, S(t))]^* \tilde{\sigma}(t, S(t))^{-1} dW(t) \right. \\ &\quad \left. - \frac{1}{2} \int_0^T |[\nu(t) - \tilde{\mu}(t, S(t))]^* \tilde{\sigma}(t, S(t))^{-1}|^2 dt \right\} \end{aligned}$$

where

$$\begin{aligned} \tilde{\mu}_i(t, s) &:= \mu_i(t, x) + \frac{1}{2} \sum_{j=1}^d \sigma_{ij}(t, x)^2, \\ \tilde{\sigma}(t, s) &:= \sigma(t, x) \end{aligned}$$

then, by Girsanov's theorem, the value function of the super-replication problem can be written as

$$\tilde{v}(0, S(0)) = \sup_{\nu \in \mathcal{D}} E^{\mathbf{P}^\nu} \left[f(\tilde{S}^\nu(T)) e^{-\int_0^T \delta_U(\nu(t)) dt} \right]$$

where δ_U is the support function of U ,

$$\tilde{S}^\nu(0) = S^\nu(0) \quad \text{and} \quad d\tilde{S}^\nu(t) = \text{diag}[\tilde{S}^\nu(t)] \left(\nu(t) dt + \tilde{\sigma}(t, \tilde{S}^\nu(t)) dW(t) \right)$$

(for a proof see e.g. [EK/QU 95], [CV/KA 93], [FÖ/KR 97]). From the general theory of stochastic optimal control, if \tilde{v} is locally bounded, then \tilde{v}_* is a viscosity supersolution of

$$\begin{aligned} -\frac{\partial u}{\partial t}(t, s) - \frac{1}{2} \text{Tr}[\text{diag}[s](\tilde{\sigma}^* \tilde{\sigma})(t, s) \text{diag}[s] D^2 u(t, s)] \\ - y^* \text{diag}[s] Du(t, s) + \delta_U(y) u(t, s) = 0, \end{aligned} \quad (47)$$

for all $y \in \tilde{U}$, where $\tilde{\sigma}(t, s) = \sigma(t, \ln s^1, \dots, \ln s^d)$. Using the notation of the previous section, since \tilde{U} is the cone generated by $\tilde{U}_1 = \{y \in \tilde{U} : |y| = 1\}$, (47) is equivalent to

$$\begin{aligned} \min \left\{ -\frac{\partial u}{\partial t}(t, s) - \frac{1}{2} \text{Tr}[\text{diag}[s](\tilde{\sigma}^* \tilde{\sigma})(t, s) \text{diag}[s] D^2 u(t, s)]; \right. \\ \left. \inf_{y \in \tilde{U}_1} (\delta_U(y) u(t, s) - y^* \text{diag}[s] Du(t, s)) \right\} = 0. \end{aligned} \quad (48)$$

The analysis developed above can also be extended, under mild conditions, to the case when the drift coefficient in the dynamics of S is influenced by the portfolio ν . Unfortunately, this dual formulation of the constraints does not extend to the general large investor framework. This is due to the fact that there is no way to get rid of the dependence of σ on ν by proceeding to some equivalent change of measure: it is well-known that the measures induced by diffusions with different diffusion coefficients are singular.

The methodology developed in Section 3 allows to avoid this step and to obtain the PDE characterization directly from the nonclassical formulation of the problem without using convex duality.

Remark 36. Assume that the function g is bounded. Then the value function v is bounded. Using the notation of the previous section, we also have that v_* , v^* , \underline{G} and \overline{G} are bounded functions.

Denote by \overline{F} and \underline{F} the functions

$$\overline{F}(s) := \limsup_{t \uparrow T, s' \rightarrow s} \tilde{v}(t, s') \quad \text{and} \quad \underline{F}(s) := \liminf_{t \uparrow T, s' \rightarrow s} \tilde{v}(t, s')$$

Applying Theorems 26 and 31, we obtain the following characterization of the value function \tilde{v} of the super-replication problem by a simple change of variable, which is clearly a generalization of (48) to the large investor model:

Theorem 37. *Let μ and σ be bounded Lipschitz functions uniformly in the t -variable, and $\sigma > 0$. Suppose further that $(g_*)^* \geq g$. Then*

(i) \tilde{v} is a (discontinuous) viscosity solution of the second order PDE

$$\min \left\{ -\frac{\partial \tilde{v}}{\partial t}(t, s) - \frac{1}{2} \text{Tr}[\text{diag}[s](\tilde{\sigma}^* \tilde{\sigma})(t, s, \hat{\nu}_0(t, s)) \text{diag}[s] D^2 \tilde{v}(t, s)]; \chi_U(\hat{\nu}(t, s)) \right\} = 0$$

on $[0, T) \times [0, \infty)^d$, where

$$\hat{\nu}_0(t, s) := \frac{\text{diag}[s] D \tilde{v}(t, s)}{\tilde{v}(t, s)}, \quad \text{and} \quad \tilde{\sigma}(t, s, \nu) := \sigma(t, \ln s^1, \dots, \ln s^d, \nu).$$

(ii) \underline{F} is viscosity supersolution of the first order PDE on $[0, \infty)^d$

$$\min \left\{ F(s) - f_*(s); \chi_U \left(\frac{\text{diag}[s]DF(s)}{F(s)} \right) \right\} = 0$$

and \overline{F} is viscosity subsolution of the first order PDE on $[0, \infty)^d$

$$\min \left\{ F(s) - f^*(s); \chi_U \left(\frac{\text{diag}[s]DF(s)}{F(s)} \right) \right\} = 0.$$

5.1 The case $U = [-l, u]$ with $l, u \geq 0$: constraints on borrowing and short-selling

As an example, we conclude this section by considering the special case $d = 1$ (only one stock) and $U = [-l, u]$ where $l, u \geq 0$ and $l + u > 0$. Then the agent has to adhere to the following constraints on borrowing and short-selling: the agent can not borrow more than u times the agent's current wealth, and can not short-sell more than $-l$ times the the agent's current wealth.

For simplicity, to avoid the change of variable $s = e^x$, we will assume that the payoff function f , and consequently the value function \tilde{v} , depend directly on the X^ν -process (and not on the stock price S), that is, the value function of the super-replication problem is given by

$$\tilde{v}(0, X(0)) := \inf \left\{ \tilde{y} > 0 : \exists \nu \in \mathcal{A}, \tilde{Y}_{\tilde{y}}^\nu(T) \geq f(X^\nu(T)) \mathbf{P} - a.s. \right\}.$$

Let us introduce the support function of the interval $[-\frac{1}{l}, \frac{1}{u}]$:

$$h(p) := \frac{1}{u}p^+ + \frac{1}{l}p^-,$$

with the convention $1/0 = +\infty$, and the usual notation $p^+ := p \vee 0$ and $p^- := (-p)^+$. Observe that h is a mapping from \mathbb{R} into $\mathbb{R} \cup \infty$.

Since for every $\varphi \in C^1(\mathbb{R})$, $(\varphi'/\varphi)(x) \in [-l, u]$ iff $\varphi(x) - h(\varphi'(x)) \geq 0$, for $U = [-l, u]$ the theorem 37 can be rewritten as

Theorem 38. *Let μ and σ be bounded Lipschitz functions uniformly in the t -variable, and $\sigma > 0$. Suppose further that $(g_*)^* \geq g$. Then*

(i) \tilde{v} is a (discontinuous) viscosity solution of

$$\min \left\{ -\tilde{v}_t(t, x) - \frac{1}{2}\sigma(t, x, \tilde{v}_x(t, x))\tilde{v}_{xx}(t, x); \tilde{v}(t, x) - h(\tilde{v}_x(t, x)) \right\} = 0$$

on $[0, T) \times \mathbb{R}$.

(ii) \underline{F} is viscosity supersolution of

$$\min \{ F(x) - f_*(x); F(x) - h(F_x(x)) \} = 0$$

and \overline{F} is viscosity subsolution of

$$\min \{ F(x) - f^*(x); F(x) - h(F_x(x)) \} = 0$$

The rest of this section is devoted to the characterization of the terminal functions \overline{F} and \underline{F} . It is known that the first order variational inequality appearing in part (ii) of the above theorem could fail to have a unique bounded discontinuous viscosity solution. However, under the condition $(f_*)^* \geq f$, all viscosity discontinuous bounded solutions have the same lower semicontinuous envelope, see e.g. [BARL 93]. Therefore, not much can be said in the case where the payoff function f is not continuous.

The following is a characterization of the terminal condition of the super-replication problem in the case of Lipschitz payoff function f .

Proposition 39. *Let the conditions of Theorem 38 hold. Assume, further, that the payoff function f is Lipschitz on \mathbb{R} . Then*

$$\overline{F}(x) = \underline{F}(x) = \hat{f}(x) := \sup_{y \in \mathbb{R}^d} f(x+y)e^{-\delta_U(y)}$$

where δ_U is the support function of the interval $U = [-l, u]$.

Proof. From Theorem 38, the functions \overline{F} and \underline{F} are, respectively, upper and lower semicontinuous viscosity sub- and supersolutions of

$$\min \{F(x) - f(x); F(x) - h(F_x(x))\} = 0 \quad (49)$$

on \mathbb{R} . In order to obtain the required result, we prove first that \hat{f} is a (continuous) viscosity supersolution of (49) (Step 1). Then we will prove that $\underline{F} \geq \hat{f}$ (Step 2). The proof is then concluded by means of a comparison theorem (Theorem 4.3 in [BARL 94], p.93); since f is Lipschitz, conditions (H1), (H4) and (H11) of this theorem are easily seen to hold. Since $\overline{F} \geq \underline{F}$ by definition, the above claims provide $\hat{f} \geq \overline{F} \geq \underline{F} \geq \hat{f}$.

Step 1. Let us prove that \hat{f} is a continuous viscosity supersolution of (49):

(i) \hat{f} is a Lipschitz function. To see this, observe that, since δ_U is a sublinear function, it follows that \hat{f} . Then, since \hat{f} and δ_U are nonnegative,

$$\begin{aligned} \hat{f}(x+y) - \hat{f}(x) &\leq \hat{f}(x+y)(1 - e^{-\delta_U(y)}), \quad \text{for all } y \in \mathbb{R} \\ &\leq \hat{f}(x+y)\delta_U(y) \leq \|f\|_\infty \max(u, l)|y|. \end{aligned}$$

(ii) \hat{f} is a supersolution of (49). To see this, let $x_0 \in \mathbb{R}$ and $\varphi \in \mathcal{C}^1(\mathbb{R})$ be such that

$$0 = (\hat{f} - \varphi)(x_0) = \min(\hat{f} - \varphi).$$

Observe that $\hat{f} \geq \varphi$. Since $\hat{f} > 0$, we can assume without loss of generality that $\varphi > 0$. By definition, we have $\hat{f}(x_0) \geq \varphi(x_0)$.

It remains to prove that $(\varphi'/\varphi)(x_0) \in [-l, u]$. Since $\hat{f} = \hat{f}$, we have

$$\varphi(x_0) = \hat{f}(x_0) \geq \hat{f}(x_0 + h)e^{\delta_U(y)} \geq \varphi(x_0 + h)e^{\delta_U(y)}$$

for all $h \in \mathbb{R}$. Now let h be an arbitrary positive constant. Then

$$\frac{\varphi(x_0 + h) - \varphi(x_0)}{h} \leq \varphi(x_0 + h) \frac{1 - e^{-uh}}{h},$$

and, by sending h to zero, we get $\varphi'(x_0) \leq u\varphi(x_0)$. Similarly, by considering an arbitrary constant $h \leq 0$, we see that $\varphi'(x_0) \geq -l\varphi(x_0)$.

Step 2. We now prove that $\underline{F} \geq \hat{f}$. From the supersolution property of \underline{F} , we have that $\underline{F} \geq f$, and for all $y \in \mathbb{R}$, \underline{F} is viscosity supersolution of

$$\delta_U(y)F(x) - yF_x(x) = 0.$$

By an easy change of variable, we see that $\underline{G} := \ln \underline{F}$ is viscosity supersolution of

$$\delta_U(y) - yG_x(x) = 0.$$

This proves that the function $x \rightarrow \delta_U(y)x - y\underline{G}(x)$ is nondecreasing (see, e.g. [C/P/T 99]) and therefore

$$\begin{aligned} \delta_U(y)(x+y) - y\underline{G}(x+y) &\geq \delta(y)x - y\underline{G}(x), \quad \text{for all } y > 0 \\ \delta_U(y)(x+y) - y\underline{G}(x+y) &\geq \delta(y)x - y\underline{G}(x), \quad \text{for all } y < 0. \end{aligned}$$

Recalling that $\underline{F} \geq f$, this provides

$$\underline{F}(x) \geq \sup_{y \in \mathbb{R}^d} \underline{F}(x+y)e^{-\delta_U(y)} \geq \sup_{y \in \mathbb{R}^d} f(x+y)e^{-\delta_U(y)} = \hat{f}(x).$$

□

References

- [BARL 93] G. BARLES. Discontinuous viscosity solutions of first-order Hamilton-Jacobi equations: A guided visit, *Nonlinear Anal.* **20** (1993) 1123-1134.
- [BARL 94] G. BARLES. Solutions de viscosité des équations de Hamilton-Jacobi, *Math. Appl.* **17** (1994) Springer-Verlag, Paris.
- [BE/SH 78] D.P. BERTSEKAS, S.E. SHREVE. *Stochastic Optimal Control: The Discrete Time Case*. Mathematics in Science and Engineering **139** (1978) Academic Press.
- [BELL 57] R. BELLMAN. *Dynamic Programming*. (1957) Princenton Univ. Press, Princenton, New Jersey.
- [BL/SCH 73] F. BLACK, M. SCHOLES. The pricing of options and corporate liabilities. *J. Political Economy* **81** (1973) 637-659.
- [B/C/S 98] M. BROADIE, J. CVITANIĆ, H.M. SONER. Optimal replication of contingent claims under portfolio constraints. *The Review of Financial Studies* **11** (2002) 59-79.
- [C/E/L 84] M.G. CRANDALL, L.C. EVANS, ISHII, P.L. LIONS. Some properties of viscosity solutions of Hamilton-Jacobi Equations. *Transactions of the American Mathematical Society* **282** (1984) 487-502.
- [C/I/L 92] M.G. CRANDALL, H. ISHII, P.L. LIONS. User's guide to viscosity solutions of second order partial differential equations. *Bulletin (New Series) of the American Mathematical Society*, vol **27**, No.1 (1992) 1-67.
- [CR/LI 83] M.G. CRANDALL, P.L. LIONS. Viscosity solutions of Hamilton-Jacobi Equations. *Transactions of the American Mathematical Society* **277** (1983) 1-42.

- [CV/KA 93] J. CVITANIĆ, I. KARATZAS. Hedging contingent claims with constrained portfolios. *Ann. Appl. Probab.*, **3**, (1993) 652-681.
- [C/K/S 98] J. CVITANIĆ, I. KARATZAS, H.M. SONER Backward SDE's with constraints on the gains process. *Ann. Appl. Probab.*, **26**, (1998) 1522-1551.
- [C/P/T 99] J. CVITANIĆ, H. PHAM, N. TOUZI. Super-replication in stochastic volatility models under portfolio constraints. *J. Appl. Probab.*, **36**, (1999) 523-545.
- [DOOB 94] J.L. DOOB. *Measure Theory*. (1994) Springer-verlag.
- [EK/QU 95] N. EL KAROUI, M.-C. QUENEZ Dynamic programming and pricing of contingent claims in a incomplete market. *SIAM J. Control Optim.*, **33** (1995) 29-66.
- [FL/SO 93] W.H. FLEMING, H.M. SONER. *Controlled Markov Processes and Viscosity Solutions* (1993) Springer-Verlag, Berlin.
- [FÖ/KR 97] H. FÖLLMER, D. KRAMKOV. Optional decomposition under constraints. *Probab. Theory and Related Fields*, **109**, (1997) 1-25.
- [GI/SK 72] I.I. GIHMAN, A.V. SKOROHOD. *Stochastic Differential Equations* (1972) Springer-Verlag, New York, Heidelberg, Berlin.
- [KA/SH 91] I. KARATZAS, S.E. SHREVE. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, **113** (1991) Springer-Verlag, New York.
- [KA/SH 98] I. KARATZAS, S.E. SHREVE. *Methods of Mathematical Finance*. Applications of Mathematics **39** (1998) Springer-Verlag, New York.
- [LIONS 83] P.L. LIONS. *Generalized solutions of Hamilton-Jacobi equations* (1983) Pitman.
- [LIONS1 83] P.L. LIONS. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part I: The dynamic programming principle and applications. *Comm. P.D.E.* **8** (1983) 1101-1174.
- [LIONS2 83] P.L. LIONS. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part II: Viscosity solutions and uniqueness. *Comm. P.D.E.* **8** (1983) 1229-1276.
- [MA/YO 99] J. MA, J. YONG. *Forward-Backward stochastic differential equations and their applications*. Lecture Notes in Mathematics, **1702** (1999) Springer-Verlag, New York.
- [MERT 73] R. MERTON. Theory of rational option pricing. *Bell Journal of Economics and Management*, **4** (1973) 141-183.
- [LIONS 85] P.L. LIONS. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part III. *Nonlinear PDE and Appl., Séminaire du Collège de France*, vol V (1985) Pitman.
- [PARD 98] E. PARDOUX. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. *Stochastic analysis and related topics VI (The Geilo Workshop, 1996)*, Editado por: L. Decreusefond, J. Gjerde, B. Øksendal; A.S. Üstünel. *Progr. Probab.*, vol **42**, Birkhäuser Boston, Boston MA (1998) 79-127.

- [PA/TA 99] E. PARDOUX, S. TANG. Forward-backward stochastic differential equations and quasi-linear PDEs' of second order, *Probab. Theory and Related Fields* **114** (1999) 123-150.
- [ROCK 70] R.T. ROCKAFELLAR. *Convex Analysis*. (1970) Princenton University Press., Princenton, NJ.
- [SO/TO 02] H. M. SONER, N. TOUZI. Stochastic target problems, dynamic programming and viscosity solutions. *SIAM J. Control Optim.*, **Vol 41**, No. 2 (2002) pp. 404-424.
- [SO/TO2 02] H. M. SONER, N. TOUZI. Dynamic programming for stochastic target problems and geometric flows, *J. Eur. Math. Soc.* **4**, No. 2 (2002) 201-236.