

CAPA
Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Matemática
Mestrado em Matemática

Controle Estocástico, Backward SDEs e EDPs Parabólicas

Jorge Alexandre Cardoso do Nascimento

JOÃO PESSOA – PB
MAIO DE 2015

Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Matemática
Mestrado em Matemática

Controle Estocástico, Backward SDEs e EDPs Parabólicas

por

Jorge Alexandre Cardoso do Nascimento

sob a orientação do

Prof. Dr. Alberto Masayoshi Faria Ohashi

João Pessoa – PB
Maio de 2015

Catálogo na publicação
Universidade Federal da Paraíba
Biblioteca Setorial do CCEN

N244c Nascimento, Jorge Alexandre Cardoso do.
Controle Estocástico, Backward SDEs e EDPs Parabólicas /
Jorge Alexandre Cardoso do Nascimento. - João Pessoa, 2015.
95f.
Orientador: Alberto Masayoshi Faria Ohashi
Dissertação (Mestrado) - UFPB/CCEN
1. Matemática. 2. Equações diferenciais estocásticas.
3. Controle estocástico.

UFPB/BC

CDU: 51(043)

Controle Estocástico, Backward SDEs e EDPs Parabólicas

por

Jorge Alexandre Cardoso do Nascimento ¹

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática da Universidade Federal da Paraíba como requisito parcial para a obtenção do título de Mestre em Matemática.

Área de Concentração: Probabilidade

Aprovada em 29 de Maio de 2015.

Banca Examinadora:

Prof. Dr. Alberto Masayoshi Faria Ohashi – UFPB
(Orientador)

Prof. Dr. Paulo Regis Caron Ruffino – UNICAMP
(Examinador Externo)

Prof. Dra. Evelina Shamarova – UFPB
(Examinador Interno)

¹O autor foi bolsista da CAPES durante a elaboração desta dissertação.

*A todos aqueles que me
proporcionaram e proporcionam
momentos felizes nesta
vida.*

Agradecimentos

Acima de tudo, a Deus por me amparar e sempre me fazer acreditar nas minhas capacidades como Ser Humano.

À minha Família, que nunca deixou de acreditar em mim e sempre me apoiou nas horas mais difíceis. Em especial, à minha esposa Suzane Rodrigues e irmã Rita do Nascimento, pela ajuda na elaboração de alguma parte textual, nesta dissertação.

Como é evidente, ao Professor Doutor Alberto Ohashi pela paciência demonstrada, que sempre se disponibilizou em esclarecer minhas dúvidas com respeito à matéria elaborada. Um reconhecimento também, a todos os demais Professores e Colegas que interagiram comigo, me ajudando a adquirir mais conhecimento matemático ao longo destes dois anos de curso.

Em particular, aos amigos Alexandre da Fonseca, António Vale, José Martins, Nuno Cabrita e Raul Neto, que mesmo estando geograficamente longe, não deixaram de ser uma inspiração para a elaboração desta obra.

À CAPES pelo apoio financeiro prestado.

Resumo

A Dissertação aborda algumas relações existentes entre teoria de controle, cálculo estocástico e equações diferenciais parciais parabólicas. O interesse é estudar representações de soluções de viscosidade para equações parabólicas via fórmulas de Feynman-Kac não lineares. Para isso, o ferramental de teoria de controle tem papel importante na conexão entre a abordagem estocástica e determinística.

Palavras-chave: equações diferenciais estocásticas, controle estocástico.

Abstract

The Dissertation study the relations between control theory, stochastic calculus and parabolic partial differential equations. The aim is to study representations of viscosity solutions for parabolic equations via the Feynman-Kac nonlinear formulas. To this end, the control theory plays an important role in the connection between the stochastic and deterministic approaches.

Keywords: stochastic differential equations, stochastic control.

Sumário

Introdução	1
1 Controle Determinístico	3
1.1 Problema Determinístico	3
1.2 Princípio do máximo de Pontryagin	5
1.3 Princípio da Programação Dinâmica	14
1.3.1 Unicidade da solução de viscosidade	23
1.3.2 Solução de viscosidade na diferenciabilidade	28
2 BSDE	33
2.1 Existência e unicidade	33
2.2 Princípio de Pontryagin Estocástico	52
2.2.1 Controle Estocástico	53
2.2.2 BSDE e Controle	56
3 BSDE e PDE	67
3.1 Programação Dinâmica estocástica	67
3.2 Fórmulas de Feynman-Kac	76
3.2.1 Solução de uma PDE parabólica linear	77
3.2.2 Solução de uma PDE parabólica não linear	82
Referências Bibliográficas	95

Notações

A seguir, lista-se algumas notações utilizadas neste trabalho, por ordem de leitura.

- SDE(SDEs) - equação(s) diferencial(s) estocástica(s);
- PDE(PDEs) - equação(s) diferencial(s) parcial(s);
- HJB - Hamilton-Jacobi-Bellman;
- BSDE(BSDEs) - equação(s) diferencial(s) estocástica(s) invertida, fixado o tempo final;
- $\mathcal{V}_{ad}[a, b]$ - conjunto dos controles admissíveis no intervalo $[a, b]$, em problemas determinísticos;
- $L^p([a, b]; (\cdot))$ - conjunto das funções f tomando valores em $(\cdot)$, mensuráveis à Lebesgue tal que $\int_a^b |f(t)|^p dt < +\infty$, com $p \geq 1$;
- $\mathcal{B}(\cdot)$ - sigma álgebra de Borel gerada pelos abertos de $(\cdot)$;
- $m_L(\cdot)$ - medida de Lebesgue no conjunto $(\cdot)$;
- $\mathbb{1}_{E_\varepsilon}(t)$ - função indicadora do conjunto E_ε definida por 1 se $t \in E_\varepsilon$ e por 0 se $t \notin E_\varepsilon$;
- $|\cdot|$ - norma Euclideana;
- $O(\varepsilon)$ - conjunto das sequências $x_n \neq 0$ para o qual existe uma constante c positiva, tal que $|\frac{x_n}{\varepsilon}| < c$;
- $o(\varepsilon)$ - conjunto das sequências $x_n \neq 0$ para o qual o limite de $\frac{x_n}{\varepsilon}$ é zero;

- $\langle \cdot \rangle$ - produto interno usual;
- $C^1(\cdot)$ (ou de classe C^1) - conjunto de todas as funções contínuas em $(\cdot)$ até à primeira derivada, inclusive;
- V - função valor determinística;
- $\exp(\cdot)$ - exponencial de $(\cdot)$;
- $C(\cdot)$ - conjunto de todas as funções contínuas em $(\cdot)$;
- $a \vee b$ - máximo entre a e b ;
- $\overline{\lim}_n c_n$ - limite de c_N tal que $c_N = \sup_{n \geq N} c_n$;
- $\underline{\lim}_n c_n$ - limite de c_N tal que $c_N = \inf_{n \geq N} c_n$;
- Ω - universo de acontecimentos em um espaço de medida de probabilidade;
- $\mathcal{F}$ - sigma álgebra gerada pelo movimento Browniano $W(\cdot)$;
- $\{\mathcal{F}_t\}_{t \geq 0}$ - filtração do tempo na sigma álgebra $\mathcal{F}$, em um espaço de medida de probabilidade;
- $E[\cdot]$ - esperança de $[\cdot]$;
- $L^p_{\mathcal{F}}([a, b]; (\cdot))$ - conjunto dos processos ζ tomando valores em $(\cdot)$, $\{\mathcal{F}\}_{t \geq 0}$ - adaptados, tal que $E\left[\int_a^b |\zeta(t)|^p dt\right] < +\infty$, com $p \geq 1$;
- P a.s - em quase em toda a parte, na medida de probabilidade;
- $\bar{C}[0, T]$ - conjunto das funções tomando valores em $\mathbb{R}^n$, contínuas em $[0, T]$;
- $a \wedge b$ - mínimo entre a e b ;
- $\bar{C}_t[0, T]$ - conjunto $\{\zeta(\cdot \wedge t) \text{ tal que } \zeta(\cdot) \in \bar{C}[0, T]\}$, com $t \in [0, T]$;
- $\mathcal{B}(\bar{C}_{t+}[0, T])$ - $\bigcap_{s \geq t} \mathcal{B}(\bar{C}_s[0, T])$, com $s \in [t, T]$;

- $A(\mathbb{R}^n)$ - conjunto de todos os processos $b : [0, T] \times \bar{C}[0, T] \rightarrow \mathbb{R}^n$, que são $\{\mathcal{B}(\bar{C}_{t+}[0, T])\}_{t \geq 0}$ - progressivamente mensuráveis;
- $A(\mathbb{R}^{n \times m})$ - conjunto de todos os processos $\sigma : [0, T] \times \bar{C}[0, T] \rightarrow \mathbb{R}^{n \times m}$, que são $\{\mathcal{B}(\bar{C}_{t+}[0, T])\}_{t \geq 0}$ - progressivamente mensuráveis;
- $e^{(\cdot)}$ - o mesmo que $\exp(\cdot)$;
- $\mathcal{U}_{ad}^L[a, b]$ - conjunto dos controles admissíveis no intervalo $[a, b]$, em problemas lineares estocásticos;
- $\mathcal{U}_{ad}[a, b]$ - conjunto dos controles admissíveis no intervalo $[a, b]$, em problemas estocásticos não necessariamente lineares;
- $ess \inf$ - ínfimo de uma família, no máximo não enumerável, de variáveis aleatórias;
- $Q \ll P$ - medida Q é absolutamente contínua em relação à medida P ;
- $\mathbb{V}$ - função valor estocástica.

Introdução

Esta dissertação estuda algumas relações entre teoria de controle, equações diferenciais estocásticas (abreviando SDEs) e equações diferenciais parciais (abreviando PDEs). Concretamente, a intenção é provar que sob determinadas condições, uma classe ampla de PDEs parabólicas podem ser solucionadas através de técnicas probabilísticas. Por exemplo, para abordar sistemas termodinâmicos ou certos comportamentos financeiros, recorre-se em algumas situações, a PDEs parabólicas. Nesse sentido, fixado o tempo final T para um determinado sistema, considere a seguinte equação diferencial de segunda ordem

$$\begin{cases} \frac{\partial v}{\partial t} + A \frac{\partial^2 v}{\partial x^2} + B \frac{\partial v}{\partial x} + F(v) = 0 \\ v_T = h, \end{cases} \quad (1)$$

em que v depende do tempo t e do estado x do fenômeno considerado. Vamos assumir que A e B são independentes da solução v de (1) e que em T , esse fenômeno está condicionado a h . Nestas condições, se F for uma função afim dada por $F(v) = cv + g$, então a solução de (1) será representada por

$$v = E \left\{ \int_t^T g(s, X) \exp \left[\int_t^s c(r, X) dr \right] ds + h(X_T) \exp \left[\int_t^T c(r, X) dr \right] \right\}, \quad (2)$$

onde c , g e h dependem de um processo X , que é solução de uma SDE. A expressão (2) é chamada fórmula de Feynman-Kac. O caso não linear é mais delicado. Veremos que nessa situação, a solução v não poderá mais ser representado por (2). Neste caso, teremos uma única representação dada pela esperança de um processo Y , que é solução de uma Backward SDE (abreviando BSDE). Este tipo especial de SDEs teve grande desenvolvimento em sua teoria, através dos trabalhos de Pardoux e Peng.

O objetivo deste trabalho é estudar as relações entre PDEs parabólicas e BSDEs. Para isso, a teoria de controle estocástico terá papel fundamental, juntamente com a noção de soluções de viscosidade introduzidas por Grandall e Lions. Assim, este trabalho será dividido em três capítulos. No *Capítulo 1*, é analisada a teoria de controle em problemas

determinísticos, onde estudamos o Princípio de Pontryagin, bem como, outros resultados relacionados com a programação dinâmica na vertente determinística. Nesse sentido, serão introduzidas as equações Hamilton-Jacobi-Bellman (abreviando HJB). Seguidamente no *Capítulo 2*, são estudadas as equações estocásticas, onde será evidente a componente de difusão nessas equações, intrinsecamente ligadas ao movimento Browniano. Estabelece-se então a principal diferença em relação à abordagem feita no capítulo anterior. Nessa vertente, será abordada uma classe importante de SDEs, ou seja as BSDEs, à qual irão desempenhar um papel decisivo em todo este mecanismo. Aqui, são analisados alguns resultados importantes para esse tipo de equações, onde se destaca o resultado da unicidade da solução de Pardoux-Peng. Aproveitando toda essa ideia, será também analisado o controle não determinístico em analogia com o capítulo anterior, que nos levará ao Princípio de Pontryagin estocástico. Finalmente no *Capítulo 3*, é estudada a programação dinâmica estocástica, onde será tratada a questão da formulação fraca em problemas de controle. Nesta etapa, será evidente a conexão com toda a teoria dos capítulos anteriores. Isto irá permitir então, estudar versões não lineares da representação de Feynman-Kac.

Capítulo 1

Controle Determinístico

Neste capítulo, como motivação para o estudo posterior das BSDEs e consequente aplicação a certas PDEs, é importante o estudo do problema de controle, para já na ótica determinística, bem como algumas das suas estratégias para lidar com este tipo de assunto. As técnicas aqui enunciadas, englobam duas vertentes. São elas, o Princípio do máximo de Pontryagin e o Princípio da programação dinâmica.

1.1 Problema Determinístico

Em meados do Século XX, após o termino da segunda grande guerra surgiu uma nova forma de abordar a otimização de recursos. Os problemas de otimização começaram a ser analisados do ponto de vista matemático com todo o seu formalismo. Apresenta-se assim, em seguida, uma possível formulação de um problema de controle determinístico e que será a alavanca para o trabalho de dissertação.

Para formalizar este tipo de problema, assuma T a variável temporal com $0 < T < +\infty$. Além disso, seja x a variável que representa o estado do sistema ao longo do tempo e u a variável do controle que podemos aplicar nesse sistema, dependendo também do tempo.

Considere $S(t) : [0, T] \rightarrow 2^{\mathbb{R}^n}$ e $U(t) : [0, T] \rightarrow 2^{\Gamma}$ em que $S(t)$ e $U(t)$ são as classes de todas as funções do estado e controle em $[0, T]$, sendo $\mathbb{R}^n$ e Γ o espaço do estado e do controle, respectivamente. Fixe as funções $x(t) \in S(t)$ e $u(t) \in U(t)$. Neste contexto para modelar o sistema, o estado é representado por uma equação diferencial que depende do controle que se aplica ao mesmo, ao longo do tempo. Para cada $t \in [0, T]$, assuma que $U(t)$ é invariante, isto é $U(t) = U$. Por outro lado, vamos considerar $S(t) = \mathbb{R}^n$ e $\Gamma = U$.

Para uma abordagem determinística em que não existe o fator aleatório no comporta-

mento do sistema, este pode ser descrito formalmente como

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt \\ x(0) = x_0 \in \mathbb{R}^n, \end{cases} \quad (1.1)$$

sendo $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$ uma função mensurável, com $t \in [0, T]$.

Definição 1.1. Seja $\mathcal{V}[0, T] := \{u : [0, T] \rightarrow U \text{ tal que } u \text{ é mensurável}\}$. Diz-se que um controle é admissível e escrevemos o seu conjunto como $\mathcal{V}_{ad}[0, T]$, sendo $(x(t), u(t))$ um par admissível se, e somente se,

- (1) $u \in \mathcal{V}[0, T]$,
- (2) x é solução única de (1.1).

Considere $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ mensuráveis, tal que a função objetivo

$$J(u(\cdot)) := \int_0^T f(t, x(t), u(t))dt + h(x(T)), \quad (1.2)$$

seja finita para qualquer par admissível $(x(t), u(t))$. Assim, o desafio é achar um controle ótimo $\bar{u}$ em $\mathcal{V}_{ad}[0, T]$, que minimize (1.2). Chamemos este problema de D , ou seja, encontre $\bar{u}$ verificando a equação (1.1) tal que,

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{V}_{ad}[0, T]} J(u(\cdot)).$$

A função h em (1.2) pode ser vista como uma penalização da minimização no tempo final T , isto é, por exemplo no contexto econômico com $n = 1$, como uma taxa fixa no preço $x(T)$ de um determinado produto.

Definição 1.2. O problema D é finito se, e somente se $\exists c > 0$ tal que $c \leq J(u(\cdot))$, $\forall u(\cdot) \in \mathcal{V}_{ad}[0, T]$. Além disso, esse problema é solúvel se, e somente se $\exists \bar{u} \in \mathcal{V}_{ad}[0, T]$ tal que $J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{V}_{ad}[0, T]} J(u(\cdot))$

Para o problema D , assuma que:

(D1) (U, d) é um espaço métrico completo e separável.

(D2) As funções $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são mensuráveis, existe uma constante $L > 0$ e existe um módulo de continuidade $\bar{w} : [0, +\infty) \rightarrow [0, +\infty)$ tal que para $\varphi(t, x, u) := b(t, x, u)$; $\varphi(t, x, u) := f(t, x, u)$; $\varphi(t, x, u) :=$

$h(x)$, tem-se

$$\left\{ \begin{array}{l} |\varphi(t, x_1, u_1) - \varphi(t, x_2, u_2)| \leq L|x_1 - x_2| + \bar{w}(d(u_1, u_2)) \\ \forall t \in [0, T], \quad \forall x_1, x_2 \in \mathbb{R}^n, \quad \forall u_1, u_2 \in U \\ \\ |\varphi(t, 0, u)| \leq L \\ \forall (t, u) \in [0, T] \times U. \end{array} \right.$$

(D3) $\forall (t, x) \in [0, T] \times \mathbb{R}^n$, o conjunto

$$(b, f)(t, x, U) := \{ (b_i(t, x, u), f(t, x, u)); u \in U, i = 1, \dots, n \},$$

é convexo e fechado em $\mathbb{R}^{n+1}$.

(D4) $S(t) = \mathbb{R}^n$.

Um resultado clássico garante que sob as condições acima, tem-se a existência de um controle ótimo.

Teorema 1.3. Se o problema D é finito com as hipóteses $D1$ a $D4$, então admite um controle ótimo.

Observação 1.4. Para prova, veja Teorema 5.1 em [11], pág 66-67.

1.2 Princípio do máximo de Pontryagin

É imprescindível analisar um teorema bastante importante na versão determinística, que foi enunciado por Lev Pontryagin e que será importante para entender, posteriormente neste trabalho, as BSDEs. Por agora, investigue-se a sua versão determinística.

Considere-se então o seguinte:

(D*1) (U, d) é um espaço métrico completo e separável.

(D*2) As funções $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são mensuráveis, existe uma constante $L > 0$ e existe um módulo de continuidade $\bar{w} : [0, +\infty) \rightarrow [0, +\infty)$ tal que para $\varphi(t, x, u) := b(t, x, u)$; $\varphi(t, x, u) := f(t, x, u)$; $\varphi(t, x, u) :=$

$h(x)$, temos

$$\left\{ \begin{array}{l} |\varphi(t, x_1, u_1) - \varphi(t, x_2, u_2)| \leq L|x_1 - x_2| + \bar{w}(d(u_1, u_2)) \\ \forall t \in [0, T], \quad x_1, x_2 \in \mathbb{R}^n, \quad u_1, u_2 \in U \\ \\ |\varphi(t, 0, u)| \leq L \\ \forall (t, u) \in [0, T] \times U. \end{array} \right.$$

(D*3) As funções $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são de classe C^1 em x e existe um módulo de continuidade $\bar{w} : [0, +\infty) \rightarrow [0, +\infty)$ tal que para $\varphi(t, x, u) := b(t, x, u)$; $\varphi(t, x, u) := f(t, x, u)$; $\varphi(t, x, u) := h(x)$, resulta

$$\left| \frac{\partial \varphi}{\partial x}(t, x_1, u_1) - \frac{\partial \varphi}{\partial x}(t, x_2, u_2) \right| \leq \bar{w}(|x_1 - x_2| + d(u_1, u_2))$$

$$\forall t \in [0, T], \quad x_1, x_2 \in \mathbb{R}^n, \quad u_1, u_2 \in U.$$

Com o objetivo de enunciar e provar o teorema principal desta seção, é importante fazer algumas considerações úteis.

Proposição 1.5. (Desigualdade de Gronwall) Considere $f : [0, T] \rightarrow \mathbb{R}$ e $g : [0, T] \rightarrow \mathbb{R}$ duas funções contínuas, com f não decrescente e g não negativa. Seja $h : [0, T] \rightarrow \mathbb{R}$ uma função contínua não negativa, tal que

$$h(t) \leq f(t) + \int_0^t g(s)h(s)ds, \quad 0 \leq t \leq T.$$

Então,

$$h(t) \leq f(t) + \int_0^t f(s)g(s) \exp \left[\int_s^t g(r)dr \right] ds, \quad 0 \leq t \leq T. \quad (1.3)$$

Observação 1.6. Uma variante deste resultado com a respectiva prova, considerando a função g uma constante K não negativa, pode ser encontrada na Proposição 3.10 em [12], pág 82-83. Neste caso, a desigualdade (1.3), terá a forma

$$h(t) \leq f(t) \exp(Kt), \quad 0 \leq t \leq T. \quad (1.4)$$

Assumindo a existência de um controle ótimo $\bar{u}$, considere o par ótimo $(\bar{x}(\cdot), \bar{u}(\cdot))$ e defina para $\varepsilon > 0$, $E_\varepsilon \in \mathcal{B}[0, T]$ tal que $m_L(E_\varepsilon) = \varepsilon$. Para $u \in \mathcal{V}_{ad}[0, T]$ qualquer, denote

$$u^\varepsilon(t) := \begin{cases} u(t) & \text{se } t \in E_\varepsilon \\ \bar{u}(t) & \text{se } t \in [0, T] \setminus E_\varepsilon. \end{cases} \quad (1.5)$$

Pela definição de função mensurável, é imediato que $u^\varepsilon(t) \in \mathcal{V}_{ad}[0, T]$.

Lema 1.7. Suponha que as hipóteses D^*1 a D^*3 valem. Seja $x^\varepsilon := x(\cdot, u^\varepsilon(\cdot))$ associado ao controle construído anteriormente. Seja y^ε a solução de

$$\begin{cases} \frac{dy^\varepsilon}{dt}(t) = \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t))y^\varepsilon(t) + \{b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t))\} \mathbb{1}_{E_\varepsilon}(t) \\ y^\varepsilon(0) = 0. \end{cases} \quad (1.6)$$

Então,

$$\sup_{0 \leq t \leq T} |x^\varepsilon(t) - \bar{x}(t)| = O(\varepsilon), \quad (1.7)$$

$$\sup_{0 \leq t \leq T} |y^\varepsilon(t)| = O(\varepsilon), \quad (1.8)$$

$$\sup_{0 \leq t \leq T} |x^\varepsilon(t) - \bar{x}(t) - y^\varepsilon(t)| = o(\varepsilon), \quad (1.9)$$

$$\begin{aligned} J(u^\varepsilon) - J(\bar{u}) &= \left\langle \frac{dh}{dx}(\bar{x}(T)), y^\varepsilon(T) \right\rangle + \int_0^T \left\{ \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle \right. \\ &\quad \left. + [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] \mathbb{1}_{E_\varepsilon}(t) \right\} dt + o(\varepsilon). \end{aligned} \quad (1.10)$$

Demonstração. Para provar a primeira condição (1.7), defina $\varphi^\varepsilon(t) := x^\varepsilon(t) - \bar{x}(t)$. Pela hipótese D^*2 , o teorema fundamental do cálculo e sabendo que $\varphi^\varepsilon(0) = 0$, resulta que

$$\begin{aligned} |\varphi^\varepsilon(t)| &= \left| \int_0^t [b(s, x^\varepsilon(s), u^\varepsilon(s)) - b(s, \bar{x}(s), \bar{u}(s))] ds \right| \\ &\leq \int_0^t |b(s, x^\varepsilon(s), u^\varepsilon(s)) - b(s, \bar{x}(s), \bar{u}(s))| ds \\ &\leq L \int_0^t |\varphi^\varepsilon(s)| ds + \int_{E_\varepsilon} \bar{w}[d(u^\varepsilon(s), \bar{u}(s))] ds \\ &\leq L \int_0^t |\varphi^\varepsilon(s)| ds + \sup_{0 \leq s \leq t} \bar{w}[d(u^\varepsilon(s), \bar{u}(s))] \varepsilon. \end{aligned}$$

Portanto, pela desigualdade de Gronwall em (1.4), é imediato que

$$|\varphi^\varepsilon(t)| \leq \sup_{0 \leq s \leq t} \bar{w}[d(u^\varepsilon(s), \bar{u}(s))] \varepsilon \exp(Lt),$$

ou seja, $\sup_{0 \leq t \leq T} |x^\varepsilon(t) - \bar{x}(t)| = O(\varepsilon)$. A segunda condição (1.8), demonstra-se da mesma maneira de (1.7). Assim por (1.6), segue que

$$\begin{aligned} |y^\varepsilon(t)| &\leq \int_0^t \left| \frac{dy^\varepsilon}{ds}(s) \right| ds \leq \int_0^t \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right| |y^\varepsilon(s)| ds + \int_{E_\varepsilon} \bar{w} [d(u^\varepsilon(s), \bar{u}(s))] ds \\ &\leq \int_0^t \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right| |y^\varepsilon(s)| ds + \sup_{0 \leq s \leq t} \bar{w} [d(u^\varepsilon(s), \bar{u}(s))] \varepsilon. \end{aligned}$$

Utilizando agora (1.3), temos para $y^\varepsilon(t)$,

$$\begin{aligned} |y^\varepsilon(t)| &\leq \sup_{0 \leq s \leq t} \bar{w} [d(u^\varepsilon(s), \bar{u}(s))] \varepsilon \\ &\quad \times \left[1 + \int_0^t \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right| \exp \left(\int_s^t \left| \frac{\partial b}{\partial x}(r, \bar{x}(r), \bar{u}(r)) \right| dr \right) ds \right] \\ &\leq \sup_{0 \leq s \leq t} \bar{w} [d(u^\varepsilon(s), \bar{u}(s))] \varepsilon \\ &\quad \times \left[1 + \sup_{0 \leq s \leq t} \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right| \int_0^t \exp \left(\sup_{s \leq r \leq t} \left| \frac{\partial b}{\partial x}(r, \bar{x}(r), \bar{u}(r)) \right| (t-s) \right) ds \right] \\ &\leq K \sup_{0 \leq s \leq t} \bar{w} [d(u^\varepsilon(s), \bar{u}(s))] \varepsilon. \end{aligned}$$

Para provar (1.9), denote $\phi^\varepsilon(t) := \varphi^\varepsilon(t) - y^\varepsilon(t)$. Por (1.5) e aplicando o teorema fundamental do cálculo juntamente com a mudança de variável na integral para $\theta \in [0, 1]$, com $x^\varepsilon(t) = \bar{x}(t) + \theta \varphi^\varepsilon(t)$, é válido escrever

$$\begin{aligned} &b(t, x^\varepsilon(t), u^\varepsilon(t)) - b(t, \bar{x}(t), \bar{u}(t)) - \{b(t, x^\varepsilon(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t))\} \mathbb{1}_{E_\varepsilon}(t) = \\ &= \int_0^1 \frac{\partial b}{\partial x}(t, \bar{x}(t) + \theta \varphi^\varepsilon(t), u^\varepsilon(t)) d\theta \varphi^\varepsilon(t). \end{aligned} \tag{1.11}$$

De (1.11) vem que

$$\begin{aligned} \frac{d\phi^\varepsilon}{dt}(t) &= \frac{d\varphi^\varepsilon}{dt}(t) - \frac{dy^\varepsilon}{dt}(t) = \frac{dx^\varepsilon}{dt}(t) - \frac{d\bar{x}}{dt}(t) - \frac{dy^\varepsilon}{dt}(t) \\ &= \int_0^1 \frac{\partial b}{\partial x}(t, \bar{x}(t) + \theta \varphi^\varepsilon(t), u^\varepsilon(t)) d\theta \varphi^\varepsilon(t) - \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t)) y^\varepsilon(t). \end{aligned}$$

Subtraindo e somando $\frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t)) \varphi^\varepsilon(t)$ no membro da direita da equação anterior,

tem-se com os devidos ajustes

$$\begin{aligned} \frac{d\phi^\varepsilon}{dt}(t) &= \int_0^1 \left\{ \frac{\partial b}{\partial x}(t, \bar{x}(t) + \theta\varphi^\varepsilon(t), u^\varepsilon(t)) - \frac{\partial b}{\partial x}(t, \bar{x}(t), u^\varepsilon(t)) \right\} d\theta \varphi^\varepsilon(t) \\ &\quad + \left\{ \frac{\partial b}{\partial x}(t, \bar{x}(t) + \theta\varphi^\varepsilon(t), u^\varepsilon(t)) - \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t)) \right\} \varphi^\varepsilon(t) \mathbb{1}_{E_\varepsilon}(t) \\ &\quad + \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t)) \phi^\varepsilon(t). \end{aligned}$$

Pelo fato de $\phi^\varepsilon(0) = 0$ e novamente pelo teorema fundamental do cálculo,

$$\begin{aligned} \phi^\varepsilon(t) &= \int_0^t \frac{d\phi^\varepsilon}{ds}(s) ds = \int_0^t \int_0^1 \left\{ \frac{\partial b}{\partial x}(s, \bar{x}(s) + \theta\varphi^\varepsilon(s), u^\varepsilon(s)) - \frac{\partial b}{\partial x}(s, \bar{x}(s), u^\varepsilon(s)) \right\} d\theta \varphi^\varepsilon(s) ds \\ &\quad + \int_0^t \left\{ \frac{\partial b}{\partial x}(s, \bar{x}(s), u(s)) - \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right\} \varphi^\varepsilon(s) \mathbb{1}_{E_\varepsilon}(s) ds \\ &\quad + \int_0^t \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \phi^\varepsilon(s) ds. \end{aligned}$$

Aplicando D^*3 , resulta

$$\begin{aligned} |\phi^\varepsilon(t)| &\leq \int_0^t \int_0^1 \left| \left\{ \frac{\partial b}{\partial x}(s, \bar{x}(s) + \theta\varphi^\varepsilon(s), u^\varepsilon(s)) - \frac{\partial b}{\partial x}(s, \bar{x}(s), u^\varepsilon(s)) \right\} \right| d\theta |\varphi^\varepsilon(s)| ds \\ &\quad + \int_0^t \left| \left\{ \frac{\partial b}{\partial x}(s, \bar{x}(s) + \theta\varphi^\varepsilon(s), u^\varepsilon(s)) - \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right\} \right| |\varphi^\varepsilon(s)| \mathbb{1}_{E_\varepsilon}(s) ds \\ &\quad + \int_0^t \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right| |\phi^\varepsilon(s)| ds \\ &\leq \int_0^t \int_0^1 \bar{w} |\theta\varphi^\varepsilon(s)| d\theta |\varphi^\varepsilon(s)| ds + \int_{E_\varepsilon} \bar{w} [d(\bar{u}(s), u(s))] |\varphi^\varepsilon(s)| ds \\ &\quad + \sup_{0 \leq s \leq t} \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right| \int_0^t |\phi^\varepsilon(s)| ds \\ &\leq K_1 K_2 \varepsilon^2 \int_0^t \bar{w} ds + K_2 \varepsilon^2 \sup_{0 \leq s \leq t} \bar{w} [d(\bar{u}(s), u(s))] \\ &\quad + \sup_{0 \leq s \leq t} \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right| \int_0^t |\phi^\varepsilon(s)| ds. \end{aligned}$$

Logo, outra vez por (1.3) e definindo para $0 \leq t \leq T$

$$f(t) := K_1 K_2 \varepsilon^2 \int_0^t \bar{w} ds + K_2 \varepsilon^2 \sup_{0 \leq s \leq t} \bar{w} [d(\bar{u}(s), u(s))]$$

e

$$g(t) := \sup_{0 \leq s \leq t} \left| \frac{\partial b}{\partial x}(s, \bar{x}(s), \bar{u}(s)) \right|,$$

tem-se o resultado. Finalmente, na prova de (1.10), utiliza-se a expansão de Taylor para funções de classe C^1 . Ora por (D^*3) , as funções h e f são de classe C^1 em $x \in \mathbb{R}^n$, donde fixando o controle ótimo $\bar{u}$, temos para cada $t \in [0, T]$ e com $\varphi^\varepsilon(t) := x^\varepsilon(t) - \bar{x}(t)$,

$$\begin{cases} f(t, x^\varepsilon(t), \bar{u}(t)) - f(t, \bar{x}(t), \bar{u}(t)) = \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), \varphi^\varepsilon(t) \right\rangle \\ h(x^\varepsilon(T)) - h(\bar{x}(T)) = \left\langle \frac{dh}{dx}(\bar{x}(T)), \varphi^\varepsilon(T) \right\rangle. \end{cases} \quad (1.12)$$

Portanto, pela definição de $J(u(\cdot))$ em (1.2), sabendo que $\phi^\varepsilon(t) := \varphi^\varepsilon(t) - y^\varepsilon(t)$ e por (1.12), segue que

$$\begin{aligned} J(u^\varepsilon) - J(\bar{u}) &= \int_0^T f(t, x^\varepsilon(t), u^\varepsilon(t)) dt + h(x^\varepsilon(T)) - \int_0^T f(t, \bar{x}(t), \bar{u}(t)) dt - h(\bar{x}(T)) \\ &= \left\langle \frac{dh}{dx}(\bar{x}(T)), \varphi^\varepsilon(T) \right\rangle + \int_0^T f(t, x^\varepsilon(t), u^\varepsilon(t)) dt - \int_0^T f(t, x^\varepsilon(t), \bar{u}(t)) dt \\ &\quad + \int_0^T f(t, x^\varepsilon(t), \bar{u}(t)) dt - \int_0^T f(t, \bar{x}(t), \bar{u}(t)) dt \\ &= \left\langle \frac{dh}{dx}(\bar{x}(T)), \varphi^\varepsilon(T) \right\rangle + \int_0^T \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), \varphi^\varepsilon(t) \right\rangle dt \\ &\quad + \int_0^T f(t, x^\varepsilon(t), u^\varepsilon(t)) dt - \int_0^T f(t, x^\varepsilon(t), \bar{u}(t)) dt \\ &= \left\langle \frac{dh}{dx}(\bar{x}(T)), y^\varepsilon(T) \right\rangle + \int_0^T \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle dt \\ &\quad + \int_0^T [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] \mathbb{1}_{E_\varepsilon}(t) dt \end{aligned}$$

$$+ \left\langle \frac{dh}{dx}(\bar{x}(T)), \phi^\varepsilon(T) \right\rangle + \int_0^T \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), \phi^\varepsilon(t) \right\rangle dt.$$

Assim, tendo em conta (1.9), resulta (1.10). $\square$

Definição 1.8. Dado um par admissível $(x(\cdot), u(\cdot))$, diz-se que existe um sistema Hamiltoniano quando, para uma função H , tem-se

$$\begin{cases} \frac{dx}{dt}(t) = \frac{\partial H}{\partial p}(t, x(t), u(t), p(t)) \\ \frac{dp}{dt}(t) = -\frac{\partial H}{\partial x}(t, x(t), u(t), p(t)) \\ x(0) = x_0, \quad p(T) = -\frac{dh}{dx}(x(T)), \end{cases} \quad (1.13)$$

em que $H : [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \rightarrow \mathbb{R}$ e $p : [0, T] \rightarrow \mathbb{R}^n$. A função H é chamada de Hamiltoniano.

Teorema 1.9. (Princípio do máximo determinístico, condição necessária) Suponha que valem as hipóteses D^*1 a D^*3 e seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ o par ótimo do problema D . Considere uma função $p : [0, T] \rightarrow \mathbb{R}^n$, satisfazendo

$$\begin{cases} \frac{dp}{dt}(t) = -\frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top p(t) + \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)) \\ p(T) = -\frac{dh}{dx}(\bar{x}(T)). \end{cases} \quad (1.14)$$

Então

$$H(t, \bar{x}(t), \bar{u}(t), p(t)) = \max_{u \in U} H(t, \bar{x}(t), u(t), p(t)), \quad (1.15)$$

onde

$$H(t, x, u, p) := \langle p, b(t, x, u) \rangle - f(t, x, u), \quad 0 \leq t \leq T, \quad x \in \mathbb{R}^n, \quad u \in U. \quad (1.16)$$

Demonstração. Seja p a solução de (1.14) no par ótimo $(\bar{x}(\cdot), \bar{u}(\cdot))$. Tem-se

$$\begin{aligned} \frac{d}{dt} \langle p(t), y^\varepsilon(t) \rangle &= \left\langle \frac{dp}{dt}(t), y^\varepsilon(t) \right\rangle + \left\langle p(t), \frac{dy^\varepsilon}{dt}(t) \right\rangle \\ &= -\left\langle \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top p(t) + \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle \\ &\quad + \left\langle p(t), \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t)) y^\varepsilon(t) \right\rangle \end{aligned}$$

$$\begin{aligned}
 & + \left\langle \left\{ b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t)) \right\} \mathbb{1}_{E_\varepsilon}(t) \right\rangle \\
 & = \left\langle \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top p(t), y^\varepsilon(t) \right\rangle + \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle \\
 & + \left\langle p(t), \frac{\partial b}{\partial x}(t, \bar{x}(t), \bar{u}(t)) y^\varepsilon(t) \right\rangle \\
 & + \langle p(t), \{ b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t)) \} \mathbb{1}_{E_\varepsilon}(t) \rangle. \tag{1.17}
 \end{aligned}$$

Pelo teorema fundamental do cálculo e por (1.17), vem que

$$\begin{aligned}
 \langle p(T), y^\varepsilon(T) \rangle - \langle p(0), y^\varepsilon(0) \rangle & = \int_0^T \frac{d}{dt} \langle p(t), y^\varepsilon(t) \rangle dt \\
 & = \int_0^T \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle dt \\
 & + \int_0^T \langle p(t), \{ b(t, \bar{x}(t), u(t)) \\
 & - b(t, \bar{x}(t), \bar{u}(t)) \} \mathbb{1}_{E_\varepsilon}(t) \rangle dt. \tag{1.18}
 \end{aligned}$$

Fixe $\bar{t} \in [0, T]$. Faça $u(t) = u$ e defina u^ε como em (1.5), onde para $\varepsilon > 0$, $E_\varepsilon \in \mathcal{B}[0, T]$ tal que $m_L(E_\varepsilon) = \varepsilon$ e $E_\varepsilon = [\bar{t}, \bar{t} + \varepsilon]$. Considerando a equação (1.10) no par $(u^\varepsilon, E_\varepsilon)$ e por (1.18), segue que

$$\begin{aligned}
 J(u^\varepsilon) - J(\bar{u}) & = \left\langle \frac{dh}{dx}(\bar{x}(T)), y^\varepsilon(T) \right\rangle \\
 & + \int_0^T \left\{ \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle \right. \\
 & + \left. [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] \mathbb{1}_{E_\varepsilon}(t) \right\} dt + o(\varepsilon) \\
 & = - \int_0^T \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle dt \\
 & - \int_0^T \langle p(t), \{ b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t)) \} \mathbb{1}_{E_\varepsilon}(t) \rangle dt \\
 & + \int_0^T \left\langle \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)), y^\varepsilon(t) \right\rangle dt
 \end{aligned}$$

$$+ \int_0^T [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] \mathbb{1}_{E_\varepsilon}(t) dt + o(\varepsilon).$$

Logo por (1.16),

$$\begin{aligned} J(u^\varepsilon) - J(\bar{u}) &= - \int_0^T \langle p(t), \{b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t))\} \mathbb{1}_{E_\varepsilon}(t) \rangle dt \\ &\quad + \int_0^T [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] \mathbb{1}_{E_\varepsilon}(t) dt + o(\varepsilon) \\ &= - \int_{\bar{t}}^{\bar{t}+\varepsilon} \langle p(t), \{b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t))\} \rangle dt \\ &\quad + \int_{\bar{t}}^{\bar{t}+\varepsilon} [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] dt + o(\varepsilon) \\ &= - \int_{\bar{t}}^{\bar{t}+\varepsilon} [H(t, \bar{x}(t), u(t), p(t)) - H(t, \bar{x}(t), \bar{u}(t), p(t))] dt + o(\varepsilon) \\ &= - \frac{1}{\varepsilon} \int_{\bar{t}}^{\bar{t}+\varepsilon} [H(t, \bar{x}(t), u(t), p(t)) - H(t, \bar{x}(t), \bar{u}(t), p(t))] dt. \end{aligned}$$

Como $J(u^\varepsilon) - J(\bar{u}) \geq 0$ e aplicando o teorema da diferenciação de Lebesgue, vem que $H(t, \bar{x}(t), \bar{u}(t), p(t)) \geq H(t, \bar{x}(t), u(t), p(t))$. Sendo a outra desigualdade imediata, tem-se pela separabilidade do espaço métrico U o resultado, ou seja

$$H(t, \bar{x}(t), \bar{u}(t), p(t)) = \max_{u \in U} H(t, \bar{x}(t), u(t), p(t)).$$

□

Do teorema anterior, o sistema de equações (1.14) e a equação (1.15) se denominam adjuntas e condição do máximo, respectivamente. O que foi provado foi que dada a função $p : [0, T] \rightarrow \mathbb{R}^n$ satisfazendo as equações adjuntas, o par ótimo $(\bar{x}(\cdot), \bar{u}(\cdot))$ maximiza o Hamiltoniano. Na realidade, o resultado é mais forte do que foi enunciado, ou seja, é possível garantir a otimalidade do par $(\bar{x}(\cdot), \bar{u}(\cdot))$, se este maximizar o Hamiltoniano. Formalizando,

Teorema 1.10. (Princípio do máximo determinístico, condição suficiente) Suponha que valem as hipóteses D^*1 a D^*3 e seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ um par admissível para o problema D . Seja ainda, a função $p : [0, T] \rightarrow \mathbb{R}^n$ satisfazendo as condições adjuntas. Suponha h convexa e

H côncava. Nestas condições, se a equação (1.15) é satisfeita no par $(\bar{x}(\cdot), \bar{u}(\cdot))$, então esse par é ótimo para o problema D .

Observação 1.11. Para uma demonstração da condição suficiente, veja Teorema 2.5 em [11], pág 112. No *Capítulo 2*, este resultado será demonstrado na versão estocástica.

O Princípio do máximo de Pontryagin é o ponto de partida para o que se pretende fazer neste trabalho. A função p que aparece no resultado anterior está relacionada, de certa forma, com o surgimento das BSDEs. Neste caso, p é apenas uma função determinística na variável temporal, que representa, por exemplo na ótica econômica, o preço "sombra" de um determinado produto. No *Capítulo 2*, quando se introduzir o fator estocástico nas equações (1.13), p irá ser generalizada, representando nesse caso a solução de uma BSDE.

1.3 Princípio da Programação Dinâmica

Outra forma de atacar o problema de controle é a utilização de técnicas de programação dinâmica. Este método tem como ponto de interesse as equações HJB. Será analisado o Princípio de Bellman que é o ponto fulcral de toda esta teoria, bem como os seus resultados mais salientes.

Consideremos o problema D do início do capítulo, mas agora com modificação na variável tempo no ponto inicial, isto é

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt \\ x(s) = y \in \mathbb{R}^n. \end{cases} \quad (1.19)$$

Como se observa pela equação acima, a ideia de estado inicial é generalizada, ou seja, dado agora o par $(s, y) \in [0, T] \times \mathbb{R}^n$, o controle $u \in U$ estará definido em $\mathcal{V}[s, T]$, onde $\mathcal{V}[s, T] = \{u : [s, T] \rightarrow U \text{ tal que } u \text{ é mensurável}\}$. Nestas condições, dado um controle ótimo $\bar{u}$, este terá agora de minimizar

$$J(s, y, u(\cdot)) := \int_s^T f(t, x(t), u(t))dt + h(x(T)). \quad (1.20)$$

Chame a este problema de $D_{s,y}$, ou seja, encontre $\bar{u}$ verificando a equação (1.19), tal que

$$J(s, y, \bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{V}_{ad}[s, T]} J(s, y, u(\cdot)).$$

Para este problema, considere as hipóteses:

$(D_{s,y}1)$ (U, d) é um espaço métrico completo e separável.

$(D_{s,y}2)$ As funções $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são uniformemente contínuas e existe uma constante $L > 0$ tal que para $\varphi(t, x, u) := b(t, x, u)$; $\varphi(t, x, u) = f(t, x, u)$; $\varphi(t, x, u) = h(x)$, temos que

$$\left\{ \begin{array}{l} |\varphi(t, x_1, u) - \varphi(t, x_2, u)| \leq L|x_1 - x_2| \\ \forall t \in [0, T], \quad \forall x_1, x_2 \in \mathbb{R}^n, \quad \forall u \in U \\ \\ |\varphi(t, 0, u)| \leq L \\ \forall (t, u) \in [0, T] \times U. \end{array} \right.$$

Tendo em conta a Definição (1.1), com as hipóteses $(D_{s,y}1)$, $(D_{s,y}2)$ e fixando o par (s, y) , assegura-se a existência de um único estado $x(t) := x(t, s, y, u(\cdot))$ para todo $t \in [s, T]$, que seja solução de (1.19), dado um estado inicial $x(s) = y$ e um controle u . Além disso, como $[s, T] \in \mathcal{B}[0, T]$, então $f \in L^1([s, T]; \mathbb{R})$. Assim sendo, $J(s, y, u(\cdot))$ está bem definido.

O problema $D_{s,y}$ permite generalizar o fato do tempo inicial não ser obrigatoriamente nulo, sendo um valor qualquer em $[s, T]$. A pergunta que surge inevitavelmente é, dado um $\hat{s}$ com $0 \leq s \leq \hat{s} \leq T$, é possível relacionar a equação (1.20) recursivamente em $\hat{s}$? O próximo teorema devido a Richard Bellman e que é a base da programação dinâmica determinística, dará a resposta. Para isso, defina

$$V(s, y) := \begin{cases} J(s, y, \bar{u}(\cdot)) & \text{se } (s, y) \in [0, T] \times \mathbb{R}^n \\ h(y) & \text{se } (s, y) \in T \times \mathbb{R}^n. \end{cases} \quad (1.21)$$

A função $V : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ é a função valor do problema $D_{s,y}$ e será crucial nos próximos resultados.

Teorema 1.12. (Princípio da otimalidade de Bellman determinístico) Suponha que valem as hipóteses $(D_{s,y}1)$ e $(D_{s,y}2)$. Então para qualquer $(s, y) \in [0, T] \times \mathbb{R}^n$, tem-se

$$V(s, y) = \inf_{u(\cdot) \in \mathcal{V}_{ad}[s, T]} \left\{ \int_s^{\hat{s}} f[t, x(t, s, y, u(\cdot)), u(t)] dt + V[\hat{s}, x(\hat{s}, s, y, u(\cdot))] \right\}. \quad (1.22)$$

Demonstração. Seja $0 \leq s \leq \hat{s} \leq T$ e chame de $\bar{V}(s, y)$ ao lado direito de (1.22). Por um lado, pode-se escrever

$$V(s, y) \leq J(s, y, u(\cdot)) = \int_s^T f[t, x(t, s, y, u(\cdot)), u(t)] dt + h[x(T, s, y, u(\cdot))]$$

$$\begin{aligned}
 &= \int_s^{\hat{s}} f[t, x(t, s, y, u(\cdot)), u(t)] dt + \int_{\hat{s}}^T f[t, x(t, s, y, u(\cdot)), u(t)] dt \\
 &+ h[x(T, s, y, u(\cdot))].
 \end{aligned}$$

Como o par $(s, y) \in [0, T] \times \mathbb{R}^n$ é qualquer, tome $(\hat{s}, x(\hat{s}))$ nas funções f e h na equação anterior. Assim, aplicando o ínfimo em $u(\cdot) \in \mathcal{V}_{ad}[s, T]$, fica

$$\begin{aligned}
 V(s, y) &\leq \int_s^{\hat{s}} f[t, x(t, s, y, u(\cdot)), u(t)] dt + J(\hat{s}, x(\hat{s}), u(\cdot)) \\
 &\leq \int_s^{\hat{s}} f[t, x(t, s, y, u(\cdot)), u(t)] dt + v(\hat{s}, x(\hat{s})) \\
 &\leq \bar{V}(s, y).
 \end{aligned} \tag{1.23}$$

Por outro lado, seja $\varepsilon > 0$. Por $D_{s,y}1$, existe $u_\varepsilon \in \mathcal{V}_{ad}[s, T]$, tal que

$$\begin{aligned}
 V(s, y) + \varepsilon &\geq J(s, y, u_\varepsilon(\cdot)) = \int_s^{\hat{s}} f(t, x_\varepsilon(t), u_\varepsilon(t)) dt + \int_{\hat{s}}^T f(t, x_\varepsilon(t), u_\varepsilon(t)) dt \\
 &+ h(x_\varepsilon(T)) \\
 &\geq \int_s^{\hat{s}} f(t, x_\varepsilon(t), u_\varepsilon(t)) dt + J(\hat{s}, x_\varepsilon(\hat{s}), u_\varepsilon(\cdot)) \\
 &\geq \int_s^{\hat{s}} f(t, x_\varepsilon(t), u_\varepsilon(t)) dt + V(\hat{s}, x_\varepsilon(\hat{s})) \\
 &\geq \bar{V}(s, y),
 \end{aligned} \tag{1.24}$$

em que $x_\varepsilon(t) := x_\varepsilon(t, s, y, u_\varepsilon(\cdot))$.

Logo, por (1.23) e (1.24), tem-se o pretendido. $\square$

Observação 1.13. Pelo que se acabou de provar e dado um par ótimo $(\bar{x}(\cdot), \bar{u}(\cdot))$, é coerente escrever

$$\begin{aligned}
 V(s, y) &= J(s, y, \bar{u}(\cdot)) = \int_s^{\hat{s}} f(t, \bar{x}(t), \bar{u}(t)) dt + \int_{\hat{s}}^T f(t, \bar{x}(t), \bar{u}(t)) dt + h(\bar{x}(T)) \\
 &= \int_s^{\hat{s}} f(t, \bar{x}(t), \bar{u}(t)) dt + J(\hat{s}, \bar{x}(\hat{s}), \bar{u}(\cdot))
 \end{aligned}$$

$$\geq \int_s^{\hat{s}} f(t, \bar{x}(t), \bar{u}(t)) dt + V(\hat{s}, \bar{x}(\hat{s})) \geq V(s, y),$$

com $\bar{x}(t) := \bar{x}(t, s, y, \bar{u}(\cdot))$. Assim sendo, resulta que

$$V(\hat{s}, \bar{x}(\hat{s})) = J(\hat{s}, \bar{x}(\hat{s}), \bar{u}(\cdot)) = \int_s^T f(t, \bar{x}(t), \bar{u}(t)) dt + h(\bar{x}(T)). \quad (1.25)$$

A relação (1.25) mostra que dado um par ótimo e progredindo ao longo do tempo no intervalo $[s, T]$, a otimalidade mantém-se no par, ou seja, a otimalidade global implica a otimalidade local.

O Princípio de Bellman é a base da programação dinâmica. De fato, este apresenta uma relação recursiva que é bastante útil na implementação da programação computacional. No entanto, a função valor definida anteriormente não é, de modo nenhum, de fácil manuseamento matemático. Esta questão obriga a um estudo mais aprofundado dessa função. O próximo resultado garante, sob determinadas condições, que a função valor é solução de uma equação diferencial parcial.

Teorema 1.14. Suponha que valem as hipóteses $(D_{s,y}1)$ e $(D_{s,y}2)$ e que a função valor $V \in C^1([0, T] \times \mathbb{R}^n)$. Então V é solução da seguinte equação

$$\begin{cases} -\frac{\partial v}{\partial t}(t, x) + \sup_{u \in U} H\left(t, x, u, -\frac{\partial v}{\partial x}(t, x)\right) = 0 & (t, x) \in [0, T] \times \mathbb{R}^n \\ v(T, x) = h(x) & (t, x) \in T \times \mathbb{R}^n, \end{cases} \quad (1.26)$$

onde

$$H(t, x, u, p) = \langle p, b(t, x, u) \rangle - f(t, x, u), \quad (t, x, u, p) \in [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n.$$

Demonstração. Fixe $(s, y) \in [0, T] \times \mathbb{R}^n$ e $u := u(\cdot) \in U$. Considere a trajetória $x(t, s, y, u)$ referente ao controle u fixado. Dado $\hat{s} \in [s, T]$, pelo Princípio da otimalidade de Bellman e dividindo ambos os membros por $\hat{s} - s$, tem-se para $u \in \mathcal{V}_{ad}(s, T]$

$$-\frac{V[\hat{s}, x(\hat{s}, s, y, u(\cdot))] - V(s, y)}{\hat{s} - s} \leq \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f[t, x(t, s, y, u(\cdot)), u(t)] dt. \quad (1.27)$$

Aplicando o teorema fundamental do cálculo em $v[\hat{s}, x(\hat{s}, s, y, u(\cdot))]$ e fazendo $\hat{s} \rightarrow s$ na

desigualdade (1.27), fica

$$-\left\{ \frac{\partial V}{\partial t} [\hat{s}, x(\hat{s}, s, y, u(\cdot))] + \frac{\partial V}{\partial x} [\hat{s}, x(\hat{s}, s, y, u(\cdot))] \right\} \leq f[s, x(s, s, y, u(\cdot)), u(t)],$$

donde, pela regra da cadeia, escrevemos

$$\begin{aligned} &\Leftrightarrow -\frac{\partial V}{\partial t}(s, y) - \left\langle \frac{\partial V}{\partial x}(s, y), b(s, y, u) \right\rangle - f(s, y, u) \leq 0 \\ &\Leftrightarrow -\frac{\partial V}{\partial t}(s, y) + H\left(s, y, u, -\frac{\partial V}{\partial x}(s, y)\right) \leq 0 \\ &\Rightarrow -\frac{\partial V}{\partial t}(s, y) + \sup_{u \in U} H\left(s, y, u, -\frac{\partial V}{\partial x}(s, y)\right) \leq 0. \end{aligned} \quad (1.28)$$

Além disso, dado $\varepsilon > 0$, $\hat{s} \in [s, T]$ e fazendo $\hat{s} - s$ tão pequeno quanto se queira, existe mais uma vez pelo Princípio da otimalidade de Bellman, $u := u_{\varepsilon, \hat{s}}(\cdot) \in U$ tal que

$$V(s, y) + \varepsilon(\hat{s} - s) \geq \int_s^{\hat{s}} f(t, x(t), u(t)) dt + V(\hat{s}, x(\hat{s})),$$

onde $x(t) := x_{\varepsilon, \hat{s}}(t, s, y, u_{\varepsilon, \hat{s}}(\cdot))$ e $V(\hat{s}, x(\hat{s})) := V[\hat{s}, x(\hat{s}, s, y, u(\cdot))]$.

Como $V \in C^1([0, T] \times \mathbb{R}^n)$, resulta que

$$\begin{aligned} &\Leftrightarrow \frac{V(\hat{s}, x(\hat{s})) - V(s, y)}{\hat{s} - s} \leq \varepsilon - \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f[t, x(t, s, y, u(t)), u(t)] dt \\ &\Leftrightarrow -\varepsilon \leq -\frac{V(\hat{s}, x(\hat{s})) - V(s, y)}{\hat{s} - s} - \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f[t, x(t, s, y, u(t)), u(t)] dt. \end{aligned}$$

Aplicando o teorema do valor intermediário, vem

$$\begin{aligned} -\varepsilon &\leq \frac{1}{\hat{s} - s} \int_s^{\hat{s}} \left\{ -\frac{\partial V}{\partial t}(t, x(t)) + \left\langle -\frac{\partial V}{\partial x}(t, x(t)), b(t, x(t), u(t)) \right\rangle \right. \\ &\quad \left. - f(t, x(t), u(t)) \right\} dt \\ &= \frac{1}{\hat{s} - s} \int_s^{\hat{s}} \left\{ -\frac{\partial V}{\partial t}(t, x(t)) + H[t, x(t), u(t), -\frac{\partial V}{\partial x}(t, x(t))] \right\} dt \\ &\leq \frac{1}{\hat{s} - s} \int_s^{\hat{s}} \left\{ -\frac{\partial V}{\partial t}(t, x(t)) + \sup_{u \in U} H[t, x(t), u(t), -\frac{\partial V}{\partial x}(t, x(t))] \right\} dt. \end{aligned}$$

Pelo fato de f , b e conseqüentemente H serem uniformemente contínuas, fazendo $\hat{s} \rightarrow s$, vem que

$$-\varepsilon \leq -\frac{\partial V}{\partial t}(s, y) + \sup_{u \in U} H\left(s, y, u, -\frac{\partial V}{\partial x}(s, y)\right),$$

donde tomando ε tão pequeno quanto se queira, resulta

$$0 \leq -\frac{\partial V}{\partial t}(s, y) + \sup_{u \in U} H\left(s, y, u, -\frac{\partial V}{\partial x}(s, y)\right). \quad (1.29)$$

Por fim, juntando (1.28) e (1.29), tem-se a prova. $\square$

Ao sistema (1.26) dá-se o nome de equações HJB associadas ao problema $D_{s,y}$, com $(s, y) \in [0, T] \times \mathbb{R}^n$. Além do mais, a função H é o Hamiltoniano, a mesma definida na seção anterior.

O teorema anterior fornece uma metodologia para encontrar o par ótimo de um problema de controle determinístico. Primeiramente, encontra-se a função valor $V \in C^1([0, T] \times \mathbb{R}^n)$ que seja solução das equações HJB. Depois, dado $(t, x) \in [0, T] \times \mathbb{R}^n$, elege-se $\bar{u}$ que é o controle ótimo, satisfazendo

$$H\left(t, x, \bar{u}, -\frac{\partial V}{\partial x}(t, x)\right) = \sup_{u \in U} H\left(t, x, u, -\frac{\partial V}{\partial x}(t, x)\right).$$

Por último, dado $\bar{u}$, encontra-se $\bar{x} := \bar{x}(s, y, \bar{u})$ através da equação

$$\begin{cases} d\bar{x}(t) = b(t, \bar{x}(t), \bar{u}(t))dt \\ \bar{x}(s) = y \in \mathbb{R}^n, \end{cases}$$

em que s é o tempo inicial do sistema, não necessariamente nulo.

Resumindo, para o par ótimo $(\bar{x}, \bar{u})$ tem-se

$$\frac{\partial V}{\partial t}(t, \bar{x}(t)) = \left\langle -\frac{\partial V}{\partial x}(t, \bar{x}(t)), b(t, \bar{x}(t), \bar{u}(t)) \right\rangle - f(t, \bar{x}(t), \bar{u}(t)).$$

Como, para $t \in [s, T]$

$$\frac{dV}{dt}(t, \bar{x}(t)) = \frac{\partial V}{\partial t}(t, \bar{x}(t)) + \left\langle \frac{\partial V}{\partial x}(t, \bar{x}(t)), b(t, \bar{x}(t), \bar{u}(t)) \right\rangle = -f(t, \bar{x}(t), \bar{u}(t)).$$

Aplica-se a integração de s a T , donde atendendo ao teorema fundamental do cálculo,

segue que

$$\begin{aligned}
 V(T, \bar{x}(T)) - V(s, y) &= \int_s^T \frac{dV}{dt}(t, \bar{x}(t)) dt = - \int_s^T f(t, \bar{x}(t), \bar{u}(t)) dt \\
 \Leftrightarrow h(\bar{x}(T)) - V(s, y) &= - \int_s^T f(t, \bar{x}(t), \bar{u}(t)) dt \\
 \Leftrightarrow h(\bar{x}(T)) - V(s, y) &= - \int_s^T f(t, \bar{x}(t), \bar{u}(t)) dt \\
 \Leftrightarrow V(s, y) &= \int_s^T f(t, \bar{x}(t), \bar{u}(t)) dt + h(\bar{x}(T)) = J(s, y, \bar{u}(\cdot)).
 \end{aligned}$$

Assim, como era de prever, o par ótimo realiza o ínfimo de $J(s, y, u(\cdot))$, tal como definido em (1.20).

As hipóteses do teorema anterior são muito fortes, na medida em que uma das condições para que a função valor resolva a equação HJB, é que esta tenha derivada contínua de primeira ordem, em todos os pontos do seu domínio. No entanto, os próximos exemplos apresentam funções valor que não são necessariamente de classe C^1 . Para isso, fixe a dimensão da variável estado em $n = 1$ e tome $t \in [s, T]$.

Exemplo 1.15. Considere a seguinte equação diferencial de um sistema de controle

$$\begin{cases} dx(t) = x(t)u(t)dt \\ x(s) = y \in \mathbb{R}^n, \end{cases} \quad (1.30)$$

com $(t, x) \in [s, T] \times \mathbb{R}$ e $u(t) \in [-1, 1]$. Defina para este problema, a função objetivo

$$J(s, y, u(\cdot)) = x(T).$$

Resolvendo (1.30), resulta que

$$x(t) = y \exp \left(\int_s^t u(r) dr \right), \quad \forall t \in [s, T].$$

Assim, pode-se escrever

$$V(s, y) = J(s, y, \bar{u}(\cdot)) = \inf_{|u| \leq 1} J(s, y, u(\cdot)) = \inf_{|u| \leq 1} \left\{ y \exp \left(\int_s^T u(r) dr \right) \right\}, \quad \forall t \in [s, T].$$

Como as expressões $\exp(T - t)$ e $\exp(t - T)$ são positivas $\forall t \in [s, T]$, segue então que

$$V(t, x) = \begin{cases} x \exp(T - t) & \text{se } x < 0 \\ x \exp(t - T) & \text{se } x \geq 0 \end{cases}, \quad t \in [s, T].$$

As derivadas parciais de $V(t, x)$ em relação a x à esquerda e à direita, são respectivamente, $\exp(T - t)$ e $\exp(t - T)$. Sendo assim, no ponto $x = 0$, a derivada dessa função não existe. Logo $V \notin C^1([s, T] \times \mathbb{R})$. Para este problema, a função f é nula, donde é imediato que o Hamiltoniano $H(t, x, u, p) = pxu$, com $|u| \leq 1$. Daí que

$$\sup_{|u| \leq 1} H(t, x, u, p) = |px|.$$

Para este Hamiltoniano, associe ao sistema, a equação HJB

$$\begin{cases} -\frac{\partial v}{\partial t}(t, x) + \left| \frac{\partial v}{\partial x}(t, x) x \right| = 0, & t \in [s, T) \\ v(T, x) = x. \end{cases} \quad (1.31)$$

Assim, para $x < 0$,

$$-\frac{\partial V}{\partial t}(t, x) + \left| \frac{\partial V}{\partial x}(t, x) x \right| = x \exp(T - t) - x \exp(T - t) = 0$$

e para $x > 0$,

$$-\frac{\partial V}{\partial t}(t, x) + \left| \frac{\partial V}{\partial x}(t, x) x \right| = -x \exp(t - T) + x \exp(t - T) = 0.$$

Donde, a função valor resolve a equação (1.31) em todos os pontos do domínio, exceto em $x = 0$ onde a derivada dessa função não é contínua.

Exemplo 1.16. Defina a seguinte equação de um sistema

$$\begin{cases} dx(t) = u(t)dt \\ x(s) = y \in \mathbb{R}^n, \end{cases} \quad (1.32)$$

com $(t, x) \in [s, T] \times \mathbb{R}$ e $u(t) \in \{-1, 0, 1\}$. Considere neste sistema, a função objetivo

$$J(s, y, u(\cdot)) = [x(T)]^2.$$

De (1.32), é imediato que

$$x(t) = y + \int_s^t u(r)dr, \quad \forall t \in [s, T].$$

Assim, pode-se escrever

$$V(s, y) = J(s, y, \bar{u}(\cdot)) = \inf_{u \in \{-1, 0, 1\}} J(s, y, u(\cdot)) = \inf_{u \in \{-1, 0, 1\}} \left[y + \int_s^T u(r)dr \right]^2, \quad \forall t \in [s, T].$$

Além disso,

$$J(s, y, u(\cdot)) = \begin{cases} (y + T - t)^2 & \text{se } u = 1 \\ y^2 & \text{se } u = 0 \\ (y + t - T)^2 & \text{se } u = -1 \end{cases}, \quad t \in [s, T]. \quad (1.33)$$

Comparando as expressões de cada ramo em (1.33) e observando as relações $(a + b)^2 = a^2 + b^2 + 2ab$ e $(a - b)^2 = a^2 + b^2 - 2ab$, fazendo $a := y$ e $b := T - t$, segue então que

$$V(t, x) = \begin{cases} (x + T - t)^2 & \text{se } x < -\frac{T - t}{2} \\ x^2 & \text{se } -\frac{T - t}{2} \leq x < \frac{T - t}{2} \\ (x + t - T)^2 & \text{se } x \geq \frac{T - t}{2} \end{cases}, \quad t \in [s, T].$$

Constata-se que, embora a função $V(t, x)$ seja contínua em todo o seu domínio, não é $C^1([s, T] \times \mathbb{R})$. Efetivamente,

$$\frac{\partial V}{\partial t} \left[t, \left(-\frac{T - t}{2} \right)^- \right] = -(T - t); \quad \frac{\partial V}{\partial t} \left[t, \left(-\frac{T - t}{2} \right)^+ \right] = 0$$

$$\frac{\partial V}{\partial t} \left[t, \left(\frac{T - t}{2} \right)^- \right] = 0; \quad \frac{\partial V}{\partial t} \left[t, \left(\frac{T - t}{2} \right)^+ \right] = -(T - t),$$

donde a derivada parcial de V em t não existe em $x = -\frac{T - t}{2}$ e $x = \frac{T - t}{2}$. Além disso, como no exemplo anterior, a função f é nula, donde o Hamiltoniano $H(t, x, u, p) = pu$. Portanto, neste caso com $t \in [s, T]$, fica:

Para $x < -\frac{T-t}{2}$,

$$-\frac{\partial V}{\partial t}(t, x) + \left(-\frac{\partial V}{\partial x}(t, x) \right) \times 1 = 2(x + T - t) - 2(x + T - t) = 0.$$

Para $-\frac{T-t}{2} < x < \frac{T-t}{2}$,

$$-\frac{\partial V}{\partial t}(t, x) + \left(-\frac{\partial V}{\partial x}(t, x) \right) \times 0 = 0 - 2x \times 0 = 0.$$

Para $x > \frac{T-t}{2}$,

$$-\frac{\partial V}{\partial t}(t, x) + \left(-\frac{\partial V}{\partial x}(t, x) \right) \times (-1) = -2(x + T - t) + 2(x + T - t) = 0.$$

Assim, como no exemplo anterior, a equação HJB é verificada na função valor em todos os pontos, exceto nos pontos isolados onde essa função não tem derivada. Como nota interessante, o leitor pode facilmente conferir neste exemplo que, se o espaço dos valores do controle u for alterado para $[-1, 1]$, resulta que a função valor é nula, $\forall t \in [s, T]$, com $s \in [0, T]$.

1.3.1 Unicidade da solução de viscosidade

Dos exemplos anteriores, fica claro que as funções valor de classe C^1 não são uma hipótese razoável. Desta forma, será necessário fornecer uma ideia de solução mais geral do que a solução de PDEs clássicas. Criou-se assim, o conceito de solução de viscosidade, que pode ser de dois tipos distintos, como ilustra a próxima definição.

Definição 1.17. Uma função $v \in C([0, T] \times \mathbb{R}^n)$ é uma subsolução de viscosidade de (1.26) se, e somente se,

$$v(T, x) \leq h(x) \quad \forall (t, x) \in T \times \mathbb{R}^n, \quad (1.34)$$

e para cada $\varphi \in C^1([0, T] \times \mathbb{R}^n)$, ao qual chamamos de função teste, sempre que $(t, x) \in [0, T) \times \mathbb{R}^n$ é um ponto de máximo local de $v - \varphi$, tem-se

$$-\frac{\partial \varphi}{\partial t}(t, x) + \sup_{u \in U} H\left(t, x, u, -\frac{\partial \varphi}{\partial x}(t, x)\right) \leq 0 \quad (t, x) \in [0, T) \times \mathbb{R}^n. \quad (1.35)$$

Por outro lado, uma função $v \in C([0, T] \times \mathbb{R}^n)$ é uma supersolução de viscosidade de (1.26) se, e somente se, em (1.34) e (1.35), a desigualdade " $\leq$ " for substituído por " $\geq$ " e o termo "máximo local" for substituído por "mínimo local".

O próximo teorema dá condições suficientemente fracas, para que funções valor de problemas de controle genéricos, sejam soluções únicas de viscosidade.

Teorema 1.18. Suponha que valem as hipóteses $(D_{s,y}1)$ e $(D_{s,y}2)$. Então para qualquer $(s, y) \in [0, T] \times \mathbb{R}^n$, a função valor $V(s, y)$ satisfaz

$$|V(s_1, y_1) - V(s_2, y_2)| \leq K\{|y_1 - y_2| + (1 + |y_1| \vee |y_2|)|s_1 - s_2|\}, \quad (1.36)$$

em que $(s_1, y_1) \in [0, T] \times \mathbb{R}^n$, $(s_2, y_2) \in [0, T] \times \mathbb{R}^n$, com K constante. Além disso, V é a única solução de viscosidade de (1.26).

Demonstração. Para provar (1.36), recorre-se às estimativas

$$\left\{ \begin{array}{l} |x(t, s, y, u(\cdot))| \leq K(1 + |y|) \\ \forall t \in [s, T], \quad \forall (s, y) \in [0, T] \times \mathbb{R}^n, \quad u \in \mathcal{V}[s, T] \\ \\ |x(t, s_1, y_1, u(\cdot)) - x(t, s_2, y_2, u(\cdot))| \leq K\{|y_1 - y_2| + (1 + |y_1| \vee |y_2|)|s_1 - s_2|\} \\ \forall t \in [s_1 \vee s_2, T]; \quad \forall (s_1, s_2) \in [0, T], \quad \forall (y_1, y_2) \in \mathbb{R}^n, \quad u \in \mathcal{V}[s_1 \vee s_2, T]. \end{array} \right. \quad (1.37)$$

Tendo em conta que

$$\begin{aligned} |x(t, s_1, y_1, u(\cdot)) - x(t, s_2, y_2, u(\cdot))| &= |x(t, s_1, y_1, u(\cdot)) - x(t, s_1, y_2, u(\cdot)) \\ &\quad + x(t, s_1, y_2, u(\cdot)) - x(t, s_2, y_2, u(\cdot))| \\ &\leq |x(t, s_1, y_1, u(\cdot)) - x(t, s_1, y_2, u(\cdot))| \\ &\quad + |x(t, s_1, y_2, u(\cdot)) - x(t, s_2, y_2, u(\cdot))|, \end{aligned}$$

pela hipótese $(D_{s,y}2)$ e a desigualdade de Gronwall em 1.4, chega-se às estimativas 1.37.

Assim sendo, por h ser Lipschitz, vem que

$$\begin{aligned} |J(t, s_1, y_1, u(\cdot)) - J(t, s_1, y_2, u(\cdot))| &\leq \int_{s_1}^T |f[t, x(t, s_1, y_1, u(\cdot)), u(t)] \\ &\quad - f[t, x(t, s_1, y_2, u(\cdot)), u(t)]| dt \\ &\quad + |h[x(T, s_1, y_1, u(\cdot))] - h[x(T, s_1, y_2, u(\cdot))]| \\ &\leq L \int_{s_1}^T |x(t, s_1, y_1, u(\cdot)) - x(t, s_1, y_2, u(\cdot))| dt \\ &\quad + L|x(T, s_1, y_1, u(\cdot)) - x(T, s_1, y_2, u(\cdot))|. \end{aligned}$$

Donde, por (1.37), segue que

$$|J(t, s_1, y_1, u(\cdot)) - J(t, s_1, y_2, u(\cdot))| \leq LK|y_1 - y_2|(T - s_1) + LK|y_1 - y_2|.$$

Agora, tomando $s_1 \leq s_2$, podemos escrever

$$\begin{aligned} |J(t, s_1, y_2, u(\cdot)) - J(t, s_2, y_2, u(\cdot))| &= \left| \int_{s_1}^{s_2} f[t, x(t, s_1, y_2, u(\cdot)), u(t)] dt \right. \\ &\quad + \int_{s_2}^T \left\{ f[t, x(t, s_1, y_2, u(\cdot)), u(t)] \right. \\ &\quad \left. - f[t, x(t, s_2, y_2, u(\cdot)), u(t)] \right\} dt \\ &\quad \left. + h[x(T, s_1, y_2, u(\cdot))] - h[x(T, s_2, y_2, u(\cdot))] \right|, \end{aligned}$$

donde, mais uma vez, por $(D_{s,y}2)$ e h ser Lipschitz, além de (1.37), fica

$$\begin{aligned} |J(t, s_1, y_2, u(\cdot)) - J(t, s_2, y_2, u(\cdot))| &\leq \int_{s_1}^{s_2} |f[t, x(t, s_1, y_2, u(\cdot)), u(t)]| dt \\ &\quad + \int_{s_2}^T |f[t, x(t, s_1, y_2, u(\cdot)), u(t)] \\ &\quad - f[t, x(t, s_2, y_2, u(\cdot)), u(t)]| dt \\ &\quad + |h[x(T, s_1, y_2, u(\cdot))] - h[x(T, s_2, y_2, u(\cdot))]| \\ &\leq \int_{s_1}^{s_2} L[1 + x(t, s_1, y_2, u(\cdot))] dt \\ &\quad + L \int_{s_1}^T |x(t, s_1, y_2, u(\cdot)) - x(t, s_2, y_2, u(\cdot))| dt \\ &\quad + L|x(T, s_1, y_2, u(\cdot)) - x(T, s_2, y_2, u(\cdot))|, \end{aligned}$$

ou seja

$$\begin{aligned} |J(t, s_1, y_2, u(\cdot)) - J(t, s_2, y_2, u(\cdot))| &\leq L(1 + K(1 + |y_2|))|s_1 - s_2| \\ &\quad + 2LK(1 + |y_2|)|s_1 - s_2||T - s_2|. \end{aligned}$$

Logo, como

$$\begin{aligned} |J(t, s_1, y_1, u(\cdot)) - J(t, s_2, y_2, u(\cdot))| &\leq |J(t, s_1, y_1, u(\cdot)) - J(t, s_1, y_2, u(\cdot))| \\ &\quad + |J(t, s_1, y_2, u(\cdot)) - J(t, s_2, y_2, u(\cdot))| \\ &\leq K\{|y_1 - y_2| + (1 + |y_1| \vee |y_2|)|s_1 - s_2|\}, \end{aligned}$$

vale para qualquer $u \in \mathcal{V}_{ad}[s, T]$, em particular valerá para o ínfimo, ou seja $\bar{u}$. Assim por (1.22), tem-se (1.36). Provemos agora a existência da solução V , ou seja, da subsolução e supersolução. Considere $\varphi \in C^1([0, T] \times \mathbb{R}^n)$, tal que o par $(s, y) \in [0, T] \times \mathbb{R}^n$ é ponto de máximo local de $V - \varphi$. Fixe o controle $u \in U$ e faça $x(t) := x(t, y, s, u)$ como sendo o estado referente ao controle fixado anteriormente, com $t \in [s, T]$. Tome ainda $|\hat{s} - s| \vee |y - x(t, y, s, u)|$ com $\hat{s} > s$. Usando o Princípio de Bellman, vem que

$$\begin{aligned} 0 &\leq \frac{V(s, y) - \varphi(s, y) - [V(\hat{s}, x(\hat{s})) - \varphi(\hat{s}, x(\hat{s}))]}{\hat{s} - s} \\ &\leq \frac{1}{\hat{s} - s} \left\{ \int_s^{\hat{s}} f(t, x(t), u) dt + [\varphi(\hat{s}, x(\hat{s})) - \varphi(s, y)] \right\}. \end{aligned}$$

Pela convergência de Lebesgue, fazendo $\hat{s} \rightarrow s$, resulta que

$$\begin{aligned} 0 &\leq f(s, y, u) + \frac{\partial \varphi}{\partial t}(s, y) + \left\langle \frac{\partial \varphi}{\partial x}(s, y), b(s, y, u) \right\rangle \\ &\Leftrightarrow -\frac{\partial \varphi}{\partial t}(s, y) + H\left(s, y, u, -\frac{\partial \varphi}{\partial x}(s, y)\right) \leq 0 \\ &\Leftrightarrow -\frac{\partial \varphi}{\partial t}(s, y) + \sup_{u \in U} H\left(s, y, u, -\frac{\partial \varphi}{\partial x}(s, y)\right) \leq 0. \end{aligned}$$

Nas mesmas condições anteriores, seja agora $(s, y) \in [0, T] \times \mathbb{R}^n$ um ponto de mínimo local de $V - \varphi$, tal que $\varphi \in C^1([0, T] \times \mathbb{R}^n)$. Considere $\varepsilon > 0$ e faça $\hat{s} - s$ tão pequeno quanto se queira. Assim, mais uma vez pelo Princípio de Bellman, existe $u := u_{\varepsilon, \hat{s}}(\cdot) \in U$ tal que

$$V(s, y) + \varepsilon(\hat{s} - s) \geq \int_s^{\hat{s}} f(t, x(t), u(t)) dt + V(\hat{s}, x(\hat{s})). \quad (1.38)$$

Também, segue que

$$V(s, y) - \varphi(s, y) - [V(\hat{s}, x(\hat{s})) - \varphi(\hat{s}, x(\hat{s}))] \leq 0, \quad (1.39)$$

onde $u(t) := u_{\varepsilon, \hat{s}}(t)$, $x(t) := x_{\varepsilon, \hat{s}}(t, s, y, u_{\varepsilon, \hat{s}}(\cdot))$, $V(\hat{s}, x(\hat{s})) := V[\hat{s}, x(\hat{s}, s, y, u(\cdot))]$ e $\varphi(\hat{s}, x(\hat{s})) := \varphi[\hat{s}, x(\hat{s}, s, y, u(\cdot))]$.

Assim, de (1.38) e (1.39), temos que

$$\begin{aligned} 0 &\geq -\varepsilon(\hat{s} - s) + \int_s^{\hat{s}} f(t, x(t), u(t)) dt + \varphi(\hat{s}, x(\hat{s})) - \varphi(s, y) \\ \Leftrightarrow 0 &\geq -\varepsilon + \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f(t, x(t), u(t)) dt + \frac{1}{\hat{s} - s} [\varphi(\hat{s}, x(\hat{s})) - \varphi(s, y)]. \end{aligned}$$

Aplicando o teorema fundamental do cálculo em $\varphi(t, x(t))$ com $t \in [s, \hat{s}]$, segue que

$$0 \geq -\varepsilon + \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f(t, x(t), u(t)) dt + \frac{1}{\hat{s} - s} \int_s^{\hat{s}} \frac{\partial \varphi}{\partial t}(t, x(t)) dt,$$

donde, aplicando a regra da cadeia à derivada parcial de φ em ordem a t e sabendo que $\varphi \in C^1([0, T] \times \mathbb{R}^n)$, resulta pelo teorema do valor intermediário,

$$\begin{aligned} -\varepsilon &\leq \frac{1}{\hat{s} - s} \int_s^{\hat{s}} \left\{ -\frac{\partial \varphi}{\partial t}(t, x(t)) + \left\langle -\frac{\partial \varphi}{\partial x}(t, x(t)), b(t, x(t), u(t)) \right\rangle \right. \\ &\quad \left. - f(t, x(t), u(t)) \right\} dt \\ &= \frac{1}{\hat{s} - s} \int_s^{\hat{s}} \left\{ -\frac{\partial \varphi}{\partial t}(t, x(t)) + H[t, x(t), u(t), -\frac{\partial \varphi}{\partial x}(t, x(t))] \right\} dt \\ &\leq \frac{1}{\hat{s} - s} \int_s^{\hat{s}} \left\{ -\frac{\partial \varphi}{\partial t}(t, x(t)) + \sup_{u \in U} H[t, x(t), u(t), -\frac{\partial \varphi}{\partial x}(t, x(t))] \right\} dt. \end{aligned}$$

Por H ser uniformemente contínua, por $(D_{s,y}1)$ e fazendo $\hat{s} \rightarrow s$, segue

$$-\varepsilon \leq -\frac{\partial \varphi}{\partial t}(s, y) + \sup_{u \in U} H\left(s, y, u, -\frac{\partial \varphi}{\partial x}(s, y)\right).$$

Ou seja,

$$0 \leq -\frac{\partial \varphi}{\partial t}(s, y) + \sup_{u \in U} H\left(s, y, u, -\frac{\partial \varphi}{\partial x}(s, y)\right).$$

No que concerne à unicidade da solução de viscosidade, a estimativa (1.36) tem um papel importante na prova, no entanto por esta ser extensa e técnica, não será descrita aqui. Para detalhes, veja o passo 3 do Teorema 2.5 em [11], pág 167 a 171. $\square$

1.3.2 Solução de viscosidade na diferenciabilidade

Na subseção anterior, ficou perceptível que a função teste φ substitui a função valor V na condição (1.35) na definição de subsolução de viscosidade (respectivamente supersolução de viscosidade). No entanto, como será perceptível a seguir, a função V satisfará a mesma condição anterior, dentro da noção de diferencial. Como a solução de viscosidade representa uma noção mais fraca em relação à solução clássica, torna-se necessário construir, como habitualmente se faz, ferramentas que serão úteis, na abordagem em relação ao que foi dito acima.

Definição 1.19. Para qualquer função $v : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ e $(t, x) \in [0, T] \times \mathbb{R}^n$, defina os seguintes conjuntos para $(s, y) \in [0, T] \times \mathbb{R}$

$$\begin{aligned} D_{t,x}^{1,+} v(t, x) &:= \left\{ (l, p) \in \mathbb{R} \times \mathbb{R}^n \mid \overline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \frac{v(s, y) - v(t, x) - l(s - t) - \langle p, y - x \rangle}{|s - t| + |y - x|} \leq 0 \right\} \\ D_{t,x}^{1,-} v(t, x) &:= \left\{ (l, p) \in \mathbb{R} \times \mathbb{R}^n \mid \underline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \frac{v(s, y) - v(t, x) - l(s - t) - \langle p, y - x \rangle}{|s - t| + |y - x|} \geq 0 \right\}, \end{aligned} \quad (1.40)$$

onde $D_{t,x}^{1,+} v(t, x)$ e $D_{t,x}^{1,-} v(t, x)$ são a superdiferencial e a subdiferencial de primeira ordem da função v , respectivamente. Além disso, $D_{t+,x}^{1,+} v(t, x)$ e $D_{t+,x}^{1,-} v(t, x)$, também se definem como a superdiferencial e subdiferencial de primeira ordem restrita a t à direita da mesma função.

Algumas considerações podem ser feitas a respeito da superdiferencial e subdiferencial. Primeiramente, observa-se que ambas as expressões nas desigualdades em (1.40) são coerentes, pois por exemplo, olhando para os denominadores de cada fração, constata-se que a sua expressão pode ser entendida como uma norma. Por outro lado, $D_{t,x}^{1,+} v(t, x)$, $D_{t,x}^{1,-} v(t, x)$, $D_{t+,x}^{1,+} v(t, x)$ e $D_{t+,x}^{1,-} v(t, x)$ são conjuntos fechados e convexos. De fato, considerando $\lambda \in [0, 1]$, $l_1 \in \mathbb{R}$, $l_2 \in \mathbb{R}$, $p_1 \in \mathbb{R}^n$ e $p_2 \in \mathbb{R}^n$, temos por cálculo simples,

$$\begin{aligned} & \frac{v(s, y) - v(t, x) - (\lambda l_1 + (1 - \lambda) l_2)(s - t) - \langle (\lambda p_1 + (1 - \lambda) p_2), y - x \rangle}{|s - t| + |y - x|} = \\ &= \lambda \frac{v(s, y) - v(t, x) - l_1(s - t) - \langle p_1, y - x \rangle}{|s - t| + |y - x|} \\ &+ (1 - \lambda) \frac{v(s, y) - v(t, x) - l_2(s - t) - \langle p_2, y - x \rangle}{|s - t| + |y - x|}. \end{aligned}$$

Como λ é não negativo, da igualdade anterior e da primeira desigualdade de (1.40), resulta a convexidade de $D_{t,x}^{1,+} v(t, x)$. Para os outros casos, o argumento é similar. Vejamos agora que estes conjuntos são fechados. Para isso, dado $l \in \mathbb{R}$ e $p \in \mathbb{R}^n$, tome as sequências

$l_n \in \mathbb{R}$ e $p_n \in \mathbb{R}^n$ tal que $l_n \rightarrow l$ e $p_n \rightarrow p$. Também, vem imediatamente pelas propriedades elementares dos limites, que

$$\begin{cases} \lim_{n \rightarrow +\infty} \left\{ \overline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \frac{v(s, y) - v(t, x) - l_n(s - t) - \langle p_n, y - x \rangle}{|s - t| + |y - x|} \right\} \leq 0 \\ \lim_{n \rightarrow +\infty} \left\{ \underline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \frac{v(s, y) - v(t, x) - l_n(s - t) - \langle p_n, y - x \rangle}{|s - t| + |y - x|} \right\} \geq 0. \end{cases}$$

Para $D_{t,x}^{1,+}v(t, x)$ e $D_{t,x}^{1,-}v(t, x)$, a ideia é a mesma. Além disso, é claro também por (1.40), que $D_{t,x}^{1,+}v(t, x) \subseteq D_{t,x}^{1,+}v(t, x)$ e $D_{t,x}^{1,-}v(t, x) \subseteq D_{t,x}^{1,-}v(t, x)$. Por outro lado, $D_{t,x}^{1,+}(-v)(t, x) = -D_{t,x}^{1,+}v(t, x)$ e $D_{t,x}^{1,-}(-v)(t, x) = -D_{t,x}^{1,-}v(t, x)$, pois basta lembrar que

$$\begin{aligned} & \overline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \frac{(-v)(s, y) - (-v)(t, x) - l(s - t) - \langle p, y - x \rangle}{|s - t| + |y - x|} \leq 0 \\ \Leftrightarrow & \overline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \left[- \frac{v(s, y) - v(t, x) - (-l)(s - t) - \langle (-p), y - x \rangle}{|s - t| + |y - x|} \right] \leq 0 \\ \Leftrightarrow & \underline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \left[\frac{v(s, y) - v(t, x) - (-l)(s - t) - \langle (-p), y - x \rangle}{|s - t| + |y - x|} \right] \geq 0. \end{aligned}$$

Por último, supondo que $D_{t,x}^{1,+}v(t_0, x_0) \cap D_{t,x}^{1,-}v(t_0, x_0) \neq \emptyset$ para algum $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$, então

$$D_{t,x}^{1,+}v(t_0, x_0) = D_{t,x}^{1,-}v(t_0, x_0) = \left(\frac{\partial v}{\partial t}(t_0, x_0); \frac{\partial v}{\partial x}(t_0, x_0) \right),$$

que corresponde a um ponto isolado e se deduz imediatamente, utilizando simultaneamente as duas desigualdades em (1.40) e a definição de diferenciabilidade para $v(t_0, x_0)$. Nestas condições, segue-se um importante resultado de solução de viscosidade no contexto da diferenciabilidade.

Teorema 1.20. Suponha que valem as hipóteses $(D_{s,y}1)$ e $(D_{s,y}2)$. Então a função valor $V \in C([0, T] \times \mathbb{R}^n)$ é a única que satisfaz as seguintes condições para $(t, x) \in [0, T] \times \mathbb{R}^n$, ou seja

$$\begin{cases} -l + \sup_{u \in U} H\left(t, x, u, -p\right) \leq 0 & \forall (l, p) \in D_{t,x}^{1,+}V(t, x) \\ -l + \sup_{u \in U} H\left(t, x, u, -p\right) \geq 0 & \forall (l, p) \in D_{t,x}^{1,-}V(t, x) \\ V(T, x) = h(x) & (t, x) \in T \times \mathbb{R}^n. \end{cases} \quad (1.41)$$

Demonstração. Seja $v \in C([0, T] \times \mathbb{R}^n)$.

Afirmção 1.20.1. Tome $(t_0, x_0) \in (0, T] \times \mathbb{R}^n$. Nestas condições, $(l, p) \in D_{t,x}^{1,+}v(t_0, x_0)$ se, e somente se, existe uma função $\varphi \in C^1([0, T] \times \mathbb{R}^n)$, tal que

$$\begin{cases} \left(\varphi(t_0, x_0), \frac{\partial \varphi}{\partial t}(t_0, x_0), \frac{\partial \varphi}{\partial x}(t_0, x_0) \right) = (v(t_0, x_0), l, p) \\ \varphi(t, x) > v(t, x) \quad \forall (t_0, x_0) \neq (t, x) \in [t_0, T] \times \mathbb{R}^n. \end{cases} \quad (1.42)$$

De fato, seja $(l, p) \in D_{t,x}^{1,+}v(t_0, x_0)$. Nestas condições, defina as funções

$$\phi(t, x) := \begin{cases} \frac{v(t, x) - v(t_0, x_0) - l(t - t_0) - \langle p, x - x_0 \rangle \vee 0}{t - t_0 + |x - x_0|} & \text{se } (t, x) \neq (t_0, x_0) \\ 0 & \text{se } (t, x) = (t_0, x_0), \end{cases}$$

$$\eta(r) := \begin{cases} \sup \{ \phi(t, x) \mid (t, x) \in [t_0, T] \times \mathbb{R}^n, \ t - t_0 + |x - x_0| \leq r \} & \text{se } r > 0 \\ 0 & \text{se } r \leq 0. \end{cases}$$

Olhando para (1.40), tem-se de imediato que $\eta : \mathbb{R} \rightarrow [0, +\infty]$ é contínua e não decrescente no seu domínio. Assim $\forall (t, x) \in [t_0, T] \times \mathbb{R}^n$, tem-se

$$\begin{aligned} \frac{v(t, x) - v(t_0, x_0) - l(t - t_0) - \langle p, x - x_0 \rangle}{(t - t_0 + |x - x_0|)} &\leq \eta(t - t_0 + |x - x_0|) \\ \Leftrightarrow v(t, x) - v(t_0, x_0) - l(t - t_0) - \langle p, x - x_0 \rangle &\leq (t - t_0 + |x - x_0|)\eta(t - t_0 + |x - x_0|). \end{aligned} \quad (1.43)$$

Defina agora, a seguinte função

$$\zeta(t, x) := \begin{cases} 0 & \text{se } (t, x) \in [0, t_0] \times \mathbb{R}^n \\ \int_0^{2(t-t_0+|x-x_0|)} \eta(r)dr + (t - t_0 + |x - x_0|)^2 & \text{se } (t, x) \in [t_0, T] \times \mathbb{R}^n. \end{cases}$$

Da mesma forma, $\zeta : [0, T] \times \mathbb{R}^n \rightarrow [0, +\infty]$ é contínua no seu domínio. Por outro lado,

$$\begin{aligned} \frac{\partial \zeta}{\partial t}(t, x) &= \frac{\partial}{\partial t} \left\{ \int_0^{2(t-t_0+|x-x_0|)} \eta(r)dr + (t - t_0 + |x - x_0|)^2 \right\} \\ &= \frac{\partial}{\partial t} \left\{ \int_0^{2(t-t_0+|x-x_0|)} \eta(r)dr \right\} + \frac{\partial}{\partial t} (t - t_0 + |x - x_0|)^2 \end{aligned}$$

$$= 2\eta[2(t - t_0 + |x - x_0|)] - 0\eta(0) + 2(t - t_0 + |x - x_0|)$$

(para a derivada parcial de ζ em ordem a x , procede-se de igual forma). Logo

$$\frac{\partial \zeta}{\partial t}(t_0, x_0) = \frac{\partial \zeta}{\partial x}(t_0, x_0) = 0,$$

ou seja, $\zeta \in C^1([0, T] \times \mathbb{R}^n)$. Sendo assim e tendo em conta (1.43), pode-se afirmar que, $\forall (t, x) \in [t_0, T] \times \mathbb{R}^n$ e $(t, x) \neq (t_0, x_0)$

$$\begin{aligned} \zeta(t, x) &\geq \int_{(t-t_0+|x-x_0|)}^{2(t-t_0+|x-x_0|)} \eta(r) dr + (t - t_0 + |x - x_0|)^2 \\ &\geq (t - t_0 + |x - x_0|)\eta(t - t_0 + |x - x_0|) + (t - t_0 + |x - x_0|)^2 \\ &\geq (t - t_0 + |x - x_0|)\eta(t - t_0 + |x - x_0|) \\ &\geq v(t, x) - v(t_0, x_0) - l(t - t_0) - \langle p, x - x_0 \rangle, \end{aligned}$$

donde, fazendo $\varphi(t, x) = v(t_0, x_0) + l(t - t_0) + \langle p, x - x_0 \rangle + \zeta(t, x)$ com $(t, x) \in [0, T] \times \mathbb{R}^n$, obtêm-se (1.42). Da outra implicação, sabendo que se tem $\varphi \in C^1([0, T] \times \mathbb{R}^n)$ tal que vale (1.42) e relembrando que

$$D_{t,x}^{1,+} v(t_0, x_0) = D_{t,x}^{1,-} v(t_0, x_0) = \left(\frac{\partial v}{\partial t}(t_0, x_0); \frac{\partial v}{\partial x}(t_0, x_0) \right),$$

tem-se imediatamente que $(l, p) \in D_{t,x}^{1,+} v(t_0, x_0)$. No que foi provado na Afirmação 1.20.1, pode-se, similarmente, garantir a existência de uma função $\varphi \in C^1([0, T] \times \mathbb{R}^n)$, tal que $v - \varphi$ atinge o máximo em sentido estrito no ponto (t_0, x_0) e

$$(\varphi(t_0, x_0), \frac{\partial \varphi}{\partial t}(t_0, x_0), \frac{\partial \varphi}{\partial x}(t_0, x_0)) = (v(t_0, x_0), l, p). \quad (1.44)$$

Analogamente, a Afirmação 1.20.1 continua valendo, substituindo $D_{t,x}^{1,+} v(\cdot)$ por $D_{t,x}^{1,-} v(\cdot)$, "máximo em sentido estrito" por "mínimo em sentido estrito" e mudando o sentido da desigualdade na inequação em (1.42). Observando também que, a expressão "em sentido estrito" pode ser omitida, substituindo os sinais "<" ou ">" por " $\leq$ " ou " $\geq$ ", respectivamente, em (1.42). Ainda com respeito a esta Afirmação, fica claro, que os gráficos de v e φ coincidem no ponto (t_0, x_0) , sendo que os seus gradientes são iguais a (l, p) e que na vizinhança desse ponto, φ se aproxima por baixo ou por cima a v .

Afirmação 1.20.2. A função $v \in C([0, T] \times \mathbb{R}^n)$ é solução de viscosidade de (1.26) se, e

somente se $v(T, x) = h(x) \forall (t, x) \in T \times \mathbb{R}^n$ e $\forall (t, x) \in [0, T) \times \mathbb{R}^n$

$$\begin{cases} -l + \sup_{u \in U} H(t, x, u, -p) \leq 0 & \forall (l, p) \in D_{t,x}^{1,+} v(t, x) \\ -l + \sup_{u \in U} H(t, x, u, -p) \geq 0 & \forall (l, p) \in D_{t,x}^{1,-} v(t, x). \end{cases} \quad (1.45)$$

Para provar a sua veracidade, suponha que a função $v \in C([0, T] \times \mathbb{R}^n)$ é solução de viscosidade de (1.26), então dado $(l, p) \in D_{t,x}^{1,+} v(t, x)$, pela Afirmação 1.20.1, existe uma função $\varphi \in C^1([0, T] \times \mathbb{R}^n)$, tal que $v - \varphi$ atinge o máximo local em (t, x) e

$$\left(\varphi(t, x), \frac{\partial \varphi}{\partial t}(t, x), \frac{\partial \varphi}{\partial x}(t, x) \right) = (v(t, x), l, p).$$

Pela definição de solução de viscosidade, tendo em conta (1.35), tem-se a primeira desigualdade de (1.45). Reciprocamente, se (1.45) vale, então dado $\varphi \in C^1([0, T] \times \mathbb{R}^n)$ tal que $v - \varphi$ atinge o máximo local em (t, x) , pela Afirmação 1.20.1 mais uma vez, resulta que

$$\left(\frac{\partial \varphi}{\partial t}(t, x), \frac{\partial \varphi}{\partial x}(t, x) \right) \in D_{t,x}^{1,+} v(t, x),$$

ou seja tem-se (1.35), isto é, a função $v \in C([0, T] \times \mathbb{R}^n)$ é solução de viscosidade de (1.26). Em relação à segunda desigualdade em (1.45), o raciocínio é análogo para $(l, p) \in D_{t,x}^{1,-} v(t, x)$. Olhando agora para a Afirmação 1.20.2, observa-se que não é necessariamente verdadeira, se $D_{t,x}^{1,+} v(t, x)$ e $D_{t,x}^{1,-} v(t, x)$ forem substituídos por $D_{t+,x}^{1,+} v(t, x)$ e $D_{t+,x}^{1,-} v(t, x)$ em (1.45), pois $D_{t,x}^{1,+} v(t, x) \subseteq D_{t+,x}^{1,+} v(t, x)$ e $D_{t,x}^{1,-} v(t, x) \subseteq D_{t+,x}^{1,-} v(t, x)$. No entanto, considerando as hipóteses $(D_{s,y}1)$ e $(D_{s,y}2)$ do teorema, dado $(l, p) \in D_{t+,x}^{1,+} V(t, x)$ em que V é a função valor e considerando a função teste φ da Afirmação 1.20.1, garante-se a existência da solução de viscosidade pelo Teorema 1.18, lembrando que na sua prova, utilizou-se o limite da variável tempo à direita, fazendo $\hat{s} \rightarrow s$ com $\hat{s} \geq s$. Mais ainda, pela unicidade da solução de viscosidade do mesmo Teorema 1.18 em junção com a Afirmação 1.20.2 e lembrando que $D_{t,x}^{1,+} V(t, x) \subseteq D_{t+,x}^{1,+} V(t, x)$, tem-se a primeira condição de (1.41). A ideia é a mesma para a segunda condição, fazendo $(l, p) \in D_{t+,x}^{1,-} V(t, x)$, concluindo-se assim a prova do teorema. $\square$

Capítulo 2

BSDE

Este capítulo tem como ementa, o estudo das BSDEs e a análise de sistemas não determinísticos. Nesse sentido, serão considerados também, processos de difusão na construção desses sistemas, que nos levarão inevitavelmente ao conceito de SDE. Assim sendo, serão analisados alguns resultados para este tipo de equações, onde se destaca o resultado da existência e unicidade da solução de uma BSDE. Além disso, como exemplo de conexão entre BSDEs e controle, é analisado a versão estocástica do Princípio do máximo de Pontryagin.

2.1 Existência e unicidade

Antes de estudar as BSDE, há que ter em conta algumas ideias relacionadas com processos estocásticos.

Definição 2.1. Seja $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ um espaço de medida de probabilidade com filtração em relação ao tempo. Diz-se que este espaço satisfaz as condições usuais se, e somente se $(\Omega, \mathcal{F}, P)$ é um espaço de medida completo, isto é, $\mathcal{F}_0$ contém todos os conjuntos de medida nula em $\mathcal{F}$ e $\{\mathcal{F}_t\}_{t \geq 0}$ é contínua à direita.

Definição 2.2. Nas mesmas condições da definição anterior e fixando o tempo T finito, dado o processo $X : [0, T] \times \Omega \rightarrow \mathcal{M}$, com $\omega \in \Omega$ e $(\mathcal{M}, d)$ um espaço métrico, tem-se que:

- (1) X é $\mathcal{F}$ - mensurável se, e somente, a imagem inversa da função $X(t, \omega)$ em qualquer aberto de $\mathcal{M}$, é mensurável com respeito à sigma álgebra $\mathcal{F}$.
- (2) X é $\mathcal{B}[0, T] \times \mathcal{F}$ - mensurável se, e somente, a imagem inversa da função $X(t, \omega)$ em qualquer aberto de $\mathcal{M}$, é mensurável com respeito à sigma álgebra $\mathcal{B}[0, T] \times \mathcal{F}$.
- (3) X é $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado se, e somente se, X é $\mathcal{F}_t$ - mensurável, $\forall t \in [0, T]$.
- (4) X é $\{\mathcal{F}_t\}_{t \geq 0}$ - progressivamente mensurável se, e somente se, a imagem inversa da função $X(s, \omega)$ em qualquer aberto de $\mathcal{M}$, é mensurável com respeito à sigma álgebra $\mathcal{B}[0, t] \times \mathcal{F}_t$, $\forall t \in [0, T]$.

Observação 2.3. Das definições anteriores resulta que, se um processo é $\{\mathcal{F}_t\}_{t \geq 0}$ - progressivamente mensurável então é $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado. No entanto, no contexto de espaço de probabilidade com filtração definido acima, a recíproca também vale. Ou seja, se um processo X , tomando valores em $\mathcal{M}$, é mensurável e $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado então existe uma modificação de X , pelo qual este é $\{\mathcal{F}_t\}_{t \geq 0}$ - progressivamente mensurável. Para detalhes, veja o Teorema em [2], pág 89.

Neste trabalho, os processos estocásticos estudados, tomarão valores no espaço dos números reais. Além disso, será usado o conceito de *martingale*, que é referido como um processo X que é $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado, tal que $E[X(t)]$ é finita e tem a propriedade $E[X(t)|\mathcal{F}_s] = X(s)$, P a.s., com $0 \leq s \leq t < T$. Por outro lado, chama-se a atenção para o fato de que no capítulo anterior, ao lidar com $x(t)$ determinístico, tínhamos que

$$x(t) = x(0) + \int_0^t b(s)ds \quad t \in [0, T],$$

com $b \in L^1([0, T]; \mathbb{R}^n)$. Daí que, utilizando o teorema fundamental do cálculo e dada uma função $F \in C^1([0, T] \times \mathbb{R}^n)$, esta podia ser representada da seguinte maneira:

$$F(t, x(t)) = F(0, x(0)) + \int_0^t \left\{ \frac{\partial F}{\partial s}(s, x(s)) + \left\langle \frac{\partial F}{\partial x}(s, x(s)), b(s) \right\rangle \right\} ds.$$

No entanto, no caso de um processo estocástico $X(\cdot)$, existe também o fator de difusão representado por σ . Para isso, considere um $\{\mathcal{F}_t\}_{t \geq 0}$ - movimento Browniano de dimensão m , que será definido por W . Sabendo que para $0 \leq s < t$, $W(\cdot)$ é um processo Gaussiano de valor médio nulo e variância $t-s$, além de que $W(t_1), W(t_2)-W(t_1), \dots, W(t_k)-W(t_{k-1})$ são incrementos independentes para $0 \leq t_1 < t_2 < \dots < t_k$, pode-se escrever

$$X(t) = X(0) + \int_0^t b(s)ds + \int_0^t \sigma(s)dW(s) \quad t \in [0, T], \quad (2.1)$$

em que $b \in L^1_{\mathcal{F}}([0, T]; \mathbb{R}^n)$ e $\sigma \in L^2_{\mathcal{F}}([0, T]; \mathbb{R}^{n \times m})$. Seguindo a mesma lógica, para a equação (2.1), pode-se generalizar a ideia anterior. Utilizando a fórmula de Itô que está intrinsecamente relacionada com a expansão de Taylor e sabendo que $F \in C^{1,2}([0, T] \times \mathbb{R}^n)$, tem-se

$$F(t, X(t)) = F(0, X(0)) + \int_0^t \left\{ \frac{\partial F}{\partial s}(s, X(s)) + \left\langle \frac{\partial F}{\partial x}(s, X(s)), b(s) \right\rangle \right.$$

$$\begin{aligned}
& + \frac{1}{2} \text{tr} \left[\sigma(s)^\top \frac{\partial^2 F}{\partial x^2}(s, X(s)) \sigma(s) \right] \Bigg\} ds \\
& + \int_0^t \left\langle \frac{\partial F}{\partial x}(s, X(s)), \sigma(s) dW(s) \right\rangle \quad \forall t \in [0, T] \quad P \text{ a.s.}, \quad (2.2)
\end{aligned}$$

em que as derivadas parciais de primeira ordem em t e em x , bem como a derivada de segunda ordem em x de F , são contínuas. Para detalhes da construção desta fórmula com $n = m = 1$, veja por exemplo, o Teorema 4.1.2 em [7], pág 44-48. Nestas condições, podemos afirmar que qualquer processo *martingale* pode ser identificado através da integral de Itô. Este fato é conhecido como o teorema da Representação de *martingale*.

Teorema 2.4. (Representação de martingale) Considere X um *martingale* quadrado integrável em $[0, T]$, ou seja, $E[(X(t))^2] < +\infty \quad \forall t \in [0, T]$. Nestas condições, existe um único processo $\zeta \in L^2_{\mathcal{F}}([0, T]; \mathbb{R}^m)$ progressivamente mensurável, tal que

$$X(t) = E[X(0)] + \int_0^t \zeta(s) dW(s) \quad \forall t \in [0, T] \quad P \text{ a.s.} \quad (2.3)$$

Observação 2.5. Para prova, veja Teorema 4.3.4 em [7], pág 53-54.

Uma noção também bastante importante, é a solução forte de uma SDE. Considere a seguinte equação:

$$\begin{cases} dX(t) = b(t, X)dt + \sigma(t, X)dW(t) \\ X(0) = \xi \end{cases} \quad P \text{ a.s.} \quad (2.4)$$

Definição 2.6. Dado $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ e $W(t)$ um $\{\mathcal{F}_t\}_{t \geq 0}$ - movimento Browniano de dimensão m , seja $b \in A(\mathbb{R}^n)$, $\sigma \in A(\mathbb{R}^{n \times m})$ e ξ uma variável aleatória limitada, $\mathcal{F}_0$ - mensurável. Diz-se que o processo contínuo $X(t)$, $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado, é solução forte de (2.4) se, e somente se

$$X(t) = \xi + \int_0^t b(s, X)ds + \int_0^t \sigma(s, X)dW(s) \quad \forall t \in [0, T],$$

onde

$$\int_0^t \{ |b(s, X)| + |\sigma(s, X)|^2 \} ds < +\infty \quad P \text{ a.s.}$$

Nas condições da definição anterior, assumamos a seguinte hipótese H :

$$\begin{cases} |b(t, x)|^2 + |\sigma(t, x)|^2 \leq K^2(1 + |x|^2) \\ |b(t, x_1) - b(t, x_2)| + |\sigma(t, x_1) - \sigma(t, x_2)| \leq K|x_1 - x_2|, \quad \forall t \in [0, T] \quad \forall x_1, x_2 \in \mathbb{R}^n. \\ |b(t, 0)| + |\sigma(t, 0)| \in L^2([0, T]; \mathbb{R}) \end{cases} \quad (2.5)$$

Para provar o próximo teorema e alguns resultados que se seguirão nesta dissertação, as desigualdades de Burkholder-Davis-Gundy são imprescindíveis.

Proposição 2.7. (desigualdades de Burkholder-Davis-Gundy) Dado $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ e $W(t)$ um $\{\mathcal{F}_t\}_{t \geq 0}$ - movimento Browniano de dimensão m , seja $\sigma \in L^2_{\mathcal{F}}([0, T]; \mathbb{R}^{n \times m})$. Então para qualquer $c > 0$, existe uma constante $K_c > 0$ tal que, para qualquer tempo de parada $\tau \in [0, T]$, tem-se

$$\frac{1}{K_c} E \left[\int_0^\tau |\sigma(s)|^2 ds \right]^c \leq E \left[\sup_{0 \leq t \leq \tau} \left| \int_0^t \sigma(s) dW(s) \right|^{2c} \right] \leq K_c E \left[\int_0^\tau |\sigma(s)|^2 ds \right]^c. \quad (2.6)$$

Observação 2.8. A prova deste fato pode ser encontrada no Teorema 3.28 em [5], pág 166. A ideia de "tempo de parada" em processos estocásticos que aparece no resultado, pode ser visto como um tempo aleatório τ em que esse processo apresenta um comportamento de interesse no que diz respeito aos eventos do passado, isto é, $\tau : \Omega \rightarrow [0, T]$ é tempo de parada em relação à filtração $\{\mathcal{F}_t\}_{t \geq 0}$, se $\forall t \in [0, T]$,

$$\{\tau \leq t\} := \{\omega \in \Omega \text{ tal que } \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

Teorema 2.9. Se a hipótese H vale então para qualquer variável aleatória ξ , $\mathcal{F}_0$ - mensurável tal que $E[|\xi|^2] < +\infty$, (2.4) admite uma única solução forte. Além disso, as seguintes estimativas valem

$$E \left[\sup_{0 \leq t \leq T} |X(t)|^2 \right] \leq K(1 + E[\xi^2]) \quad (2.7)$$

$$E \left[\sup_{0 \leq t \leq T} |X_1(t) - X_2(t)|^2 \right] \leq KE[\xi_1 - \xi_2]^2 \quad (2.8)$$

$$E[|X(t_1) - X(t_2)|^2] \leq K(1 + E[\xi^2])|t_1 - t_2|. \quad (2.9)$$

Demonstração. Fixe $T > 0$. Para cada $\tau \in [0, T]$, defina

$$\mathcal{L}_2[0, \tau] := \{X : [0, \tau] \times \Omega \rightarrow \mathbb{R}^n \text{ tal que } X(\cdot) \text{ é } \{\mathcal{F}_t\}_{t \geq 0} \text{ - adaptado, contínuo e } E \sup_{0 \leq t \leq \tau} |X(t)|^2 < +\infty\},$$

com norma

$$|X(\cdot)|_{\mathcal{L}_2[0,\tau]} := \left\{ E \left[\sup_{0 \leq t \leq \tau} |X(t)|^2 \right] \right\}^{\frac{1}{2}}. \quad (2.10)$$

Facilmente conclui-se que $\mathcal{L}_2[0,\tau]$ é um espaço de Hilbert munido com a norma definida em (2.10). Além disso, dado $X_1(\cdot), X_2(\cdot) \in \mathcal{L}_2[0,\tau]$, tal que

$$\begin{cases} X_1(t) = \xi + \int_0^t b(s, X_1(s)) ds + \int_0^t \sigma(s, X_1(s)) dW(s) \\ X_2(t) = \xi + \int_0^t b(s, X_2(s)) ds + \int_0^t \sigma(s, X_2(s)) dW(s) \end{cases} \quad t \in [0, \tau], \quad (2.11)$$

tem-se por (2.5), (2.11) e as desigualdades de Burkholder-Davis-Gundy em (2.6) que,

$$\begin{aligned} |X(\cdot)|_{\mathcal{L}_2[0,\tau]}^2 &= E \left[\sup_{0 \leq t \leq \tau} |X(t)|^2 \right] = E \left[\sup_{0 \leq t \leq \tau} \left| \xi + \int_0^t b(s, X(s)) ds + \int_0^t \sigma(s, X(s)) dW(s) \right|^2 \right] \\ &\leq 3^2 E \left\{ \sup_{0 \leq t \leq \tau} \left[|\xi|^2 + \left| \int_0^t b(s, X(s)) ds \right|^2 + \left| \int_0^t \sigma(s, X(s)) dW(s) \right|^2 \right] \right\} \\ &\leq 3^2 E \left[\sup_{0 \leq t \leq \tau} |\xi|^2 \right] + 3^2 \tau E \left[\int_0^\tau |b(s, X(s))|^2 ds \right] + 3^2 K_1 E \left[\int_0^\tau |\sigma(s, X(s))|^2 ds \right] \\ &\leq 9 E [|\xi|^2] + 9 \tau E \left[\int_0^\tau K^2 (1 + |X(s)|^2) ds \right] + 9 K_1 E \left[\int_0^\tau K^2 (1 + |X(s)|^2) ds \right] \\ &\leq C_1 + C_2 E \left[\int_0^\tau |X(s)|^2 ds \right] \leq C_1 + C_2 \tau |X(\cdot)|_{\mathcal{L}_2[0,\tau]}^2 < +\infty. \end{aligned} \quad (2.12)$$

Por outro lado,

$$\begin{aligned} |X_1(\cdot) - X_2(\cdot)|_{\mathcal{L}_2[0,\tau]}^2 &= E \left[\sup_{0 \leq t \leq \tau} |X_1(t) - X_2(t)|^2 \right] \\ &= E \left[\sup_{0 \leq t \leq \tau} \left| \int_0^t [b(s, X_1(s)) - b(s, X_2(s))] ds \right. \right. \\ &\quad \left. \left. + \int_0^t [\sigma(s, X_1(s)) - \sigma(s, X_2(s))] dW(s) \right|^2 \right] \\ &\leq 2^2 E \left\{ \sup_{0 \leq t \leq \tau} \left[\left| \int_0^t [b(s, X_1(s)) - b(s, X_2(s))] ds \right|^2 \right. \right. \end{aligned}$$

$$\begin{aligned}
& + \left| \int_0^t [\sigma(s, X_1(s)) - \sigma(s, X_2(s))] dW(s) \right|^2 \Bigg\} \\
& \leq 4\tau E \left\{ \int_0^\tau |b(s, X_1(s)) - b(s, X_2(s))|^2 ds \right\} \\
& + 4K_1 E \left[\int_0^\tau |\sigma(s, X_1(s)) - \sigma(s, X_2(s))|^2 ds \right] \\
& \leq 4K^2\tau E \left[\int_0^\tau |X_1(s) - X_2(s)|^2 ds \right] \\
& + 4K_1 K^2 E \left[\int_0^\tau |X_1(s) - X_2(s)|^2 ds \right] \\
& = (4K^2\tau^2 + 4K_1 K^2\tau) \|X_1(\cdot) - X_2(\cdot)\|_{\mathcal{L}_2[0, \tau]}^2. \tag{2.13}
\end{aligned}$$

Como $\mathcal{L}_2[0, \tau]$ é espaço de Banach e tomando $4K^2\tau^2 + 4K_1 K^2\tau < 1$, tem-se uma contração em (2.13). Logo, existe um único ponto fixo, ou seja uma única solução forte de (2.4) no intervalo $[0, \tau]$. Aplicando o mesmo raciocínio para $[\tau, 2\tau]$ e assim sucessivamente, obtemos uma única solução forte em $[0, T]$. Em relação às estimativas, (2.7) é imediato devido a (2.12) e à desigualdade de Grownall. Para verificar (2.8), notamos que

$$\begin{aligned}
E \left[\sup_{0 \leq t \leq T} |X_1(t) - X_2(t)|^2 \right] &= E \left[\sup_{0 \leq t \leq T} \left| (\xi_1 - \xi_2) + \int_0^t [b(s, X_1(s)) - b(s, X_2(s))] ds \right. \right. \\
& \quad \left. \left. + \int_0^t [\sigma(s, X_1(s)) - \sigma(s, X_2(s))] dW(s) \right|^2 \right] \\
&\leq 3^2 E \left\{ \sup_{0 \leq t \leq T} \left[|\xi_1 - \xi_2|^2 + \left| \int_0^t [b(s, X_1(s)) - b(s, X_2(s))] ds \right|^2 \right. \right. \\
& \quad \left. \left. + \left| \int_0^t [\sigma(s, X_1(s)) - \sigma(s, X_2(s))] dW(s) \right|^2 \right] \right\} \\
&\leq 9E|\xi_1 - \xi_2|^2 + 9TE \left[\int_0^T |b(s, X_1(s)) - b(s, X_2(s))|^2 ds \right] \\
& + 9KE \left[\int_0^T |\sigma(s, X_1(s)) - \sigma(s, X_2(s))|^2 ds \right] \\
&\leq 9E|\xi_1 - \xi_2|^2 + 9K_1 E \left[\int_0^T |X_1(s) - X_2(s)|^2 ds \right]
\end{aligned}$$

$$\begin{aligned}
& + 9K_2 E \left[\int_0^T |(X_1(s) - X_2(s))|^2 ds \right] \\
& \leq 9E|\xi_1 - \xi_2|^2 + 9K_3 \int_0^T E \left[\sup_{0 \leq s \leq T} |X_1(s) - X_2(s)|^2 \right] ds. \quad (2.14)
\end{aligned}$$

Utilizando a desigualdade de Grownall, observa-se que (2.8) vale. Por último, tem-se

$$\begin{aligned}
E[|X(t_1) - X(t_2)|^2] &= E \left[\left| \int_{t_1}^{t_2} b(s, X(s)) ds + \int_{t_1}^{t_2} \sigma(s, X(s)) dW(s) \right|^2 \right] \\
&\leq 2^2 E \left[\left| \int_{t_1}^{t_2} b(s, X(s)) ds \right|^2 + \left| \int_{t_1}^{t_2} \sigma(s, X(s)) dW(s) \right|^2 \right] \\
&\leq 4|t_1 - t_2| E \left[\int_{t_1}^{t_2} |b(s, X(s))|^2 ds \right] + KE \left[\int_{t_1}^{t_2} |\sigma(s, X(s))|^2 ds \right] \\
&\leq 4|t_1 - t_2| E \left[\int_{t_1}^{t_2} |b(s, X(s)) - b(s, 0)|^2 ds \right] \\
&\quad + KE \left[\int_{t_1}^{t_2} |\sigma(s, X(s)) - \sigma(s, 0)|^2 ds \right] \\
&\quad + 4|t_1 - t_2| E \left[\int_{t_1}^{t_2} |b(s, 0)|^2 ds \right] + KE \left[\int_{t_1}^{t_2} |\sigma(s, 0)|^2 ds \right],
\end{aligned}$$

isto é

$$\begin{aligned}
E[|X(t_1) - X(t_2)|^2] &\leq 4K_1|t_1 - t_2| E \left[\int_0^T |X(s)|^2 ds \right] + K_2 E \left[\int_{t_1}^{t_2} |X(s)|^2 ds \right] \\
&\quad + 4|t_1 - t_2| E \left[\int_0^T |(b(s, 0))|^2 ds \right] + KE \left[\int_{t_1}^{t_2} |(\sigma(s, 0))|^2 ds \right]. \quad (2.15)
\end{aligned}$$

Assim, por (2.7) e o fato de $|b(t, 0)| + |\sigma(t, 0)| \in L^2([0, T]; \mathbb{R})$, chega-se à desigualdade (2.9). $\square$

Observação 2.10. Do teorema anterior, algumas observações são importantes. Em primeiro lugar, na prova da unicidade da solução forte, não é necessário que τ seja tempo de parada. Além disso, evidencia-se o fato do anterior resultado ser válido, também, para espaços de Banach $\mathcal{L}_p[0, T]$, com $p \geq 1$. Para o que se seguirá neste trabalho, será suficiente

considerar $p = 2$. Ainda como consequência de tal teorema, temos que se $X_1(t)$ e $X_2(t)$ são soluções fortes de (2.4), então

$$P(X_1(t) = X_2(t), \quad \forall t \in [0, T]) = 1.$$

Imagine agora um investidor que daqui a 25 anos deseja ter um lucro de p reais. Iniciando hoje o seu investimento que estratégias, ele deve adotar para atingir sua meta? Por exemplo, fazendo $T = 25$, fica claro para este problema, que ao contrário das SDEs, tem que existir uma "certeza" no tempo futuro, ou seja, sendo $Y(t)$ o processo estocástico do investimento em causa, com $t \in [0, 25]$, ter-se-ia $Y(25) = p$. Para formalizar a ideia, considere o seguinte exemplo simplista, que é dada pela equação diferencial

$$\begin{cases} dY(t) = 0 \\ Y(T) = \xi \end{cases} \quad P \text{ a.s.} \quad (2.16)$$

Uma solução natural para esta equação, seria considerar $Y(t) = Y(0) = \xi$. Assim, pelo Teorema 2.9 e assumindo que ξ é $\mathcal{F}_0$ - mensurável, existiria um único processo Y em (2.16). No entanto neste caso, como $Y(T) = \xi$ é $\mathcal{F}_T$ - mensurável, para que (2.16) faça sentido, há que adaptar ou condicionar $Y(t)$ ao tempo passado, fazendo $Y(t) = E[Y(T)|\mathcal{F}_t]$, com $0 \leq t \leq T$. Considerando o movimento Browniano habitual $W(\cdot)$ e $Y(\cdot)$ um *martingale* quadrado integrável, pelo teorema da representação de *martingale*, existe um único processo $Z(\cdot) \in L^2_{\mathcal{F}}([0, T]; \mathbb{R}^{k \times m})$, $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado tal que

$$Y(t) = E[\xi] + \int_0^t Z(s) dW(s) \quad \forall t \in [0, T] \quad P \text{ a.s.}$$

Pondo $Y(t) = E[Y(T)|\mathcal{F}_t]$, podemos escrever

$$Y(t) = E[\xi] + \int_0^t Z(s) dW(s) = Y(0) + \int_0^t Z(s) dW(s) \quad \forall t \in [0, T] \quad P \text{ a.s.}$$

Como $\int_0^t Z(s) dW(s) = \int_0^T Z(s) dW(s) - \int_t^T Z(s) dW(s)$, vem que

$$Y(t) = \xi - \int_t^T Z(s) dW(s) \quad \text{com} \quad dY(t) = Z(t) dW(t) \quad \text{e} \quad Y(T) = \xi.$$

Assim sendo, $\forall t \in [0, T]$ conclui-se que o processo $Z(t)$ surge na equação, para que $Y(t)$ seja $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado.

Proposição 2.11. Dado ξ , $\mathcal{F}_T$ - mensurável tal que $E[|\xi|^2] < +\infty$, o par de processos adaptados (Y, Z) são únicos em

$$\begin{cases} dY(t) = Z(t)dW(t) \\ Y(T) = \xi \quad P \text{ a.s.} \end{cases} \quad (2.17)$$

A equação diferencial estocástica (2.17) é chamada de BSDE, pois está fixado o tempo terminal T .

Demonstração. Sejam os pares (Y_1, Z_1) e (Y_2, Z_2) duas soluções de (2.17), tais que

$$Y_1(t) = \xi - \int_t^T Z_1(s)dW(s) \quad \text{e} \quad Y_2(t) = \xi - \int_t^T Z_2(s)dW(s).$$

Resulta então que

$$\begin{aligned} Y_1(t) - Y_2(t) &= \xi - \int_t^T Z_1(s)dW(s) - \xi + \int_t^T Z_2(s)dW(s) = \int_t^T (Z_2(s) - Z_1(s))dW(s) \\ &= \int_0^T (Z_2(s) - Z_1(s))dW(s) - \int_0^t (Z_2(s) - Z_1(s))dW(s). \end{aligned}$$

Fazendo $Y(t) := Y_1(t) - Y_2(t)$, é imediato que $Y(T) = 0$, em que

$$\int_0^T (Z_2(s) - Z_1(s))dW(s) \quad (\mathcal{F}_0 \text{ - mensurável}).$$

Assim, tendo em conta que $Y(t) = E[Y(T)|\mathcal{F}_t] = E[0|\mathcal{F}_t] = 0$, resulta $Y_1(t) = Y_2(t)$ e $Z_1(t) = Z_2(t)$ com $t \in [0, T]$. $\square$

Pode-se generalizar a equação (2.17) da seguinte forma. Considere uma função $f : [0, T] \times \Omega \times \mathbb{R}^k \times \mathbb{R}^{k \times m} \rightarrow \mathbb{R}^k$, tal que

$$\begin{cases} dY(t) = -f(t, Y(t), Z(t))dt + Z(t)dW(t) \\ Y(T) = \xi \end{cases}, \quad t \in [0, T] \quad P \text{ a.s.} \quad (2.18)$$

A função f é estocástica pois depende de pelo menos um processo. Chamemos f de função geradora da BSDE (2.18). Olhando para esta BSDE, verifica-se que a sua solução será um processo de Itô, ou seja

$$Y(t) = \xi + \int_t^T f(s, Y(s), Z(s)) ds - \int_t^T Z(s) dW(s). \quad (2.19)$$

Denote $S^2(\mathbb{R}^k)$, como o conjunto de todos os processos progressivamente mensuráveis tomando valores em $\mathbb{R}^k$ tal que $\forall Y(\cdot) \in S^2(\mathbb{R}^k)$, tem-se

$$|Y(\cdot)|_{S^2(\mathbb{R}^k)} := \left\{ E \left[\sup_{0 \leq t \leq T} |Y(t)|^2 \right] \right\}^{\frac{1}{2}} < +\infty.$$

Defina também $M^2(\mathbb{R}^{k \times m})$, como o conjunto de todos os processos progressivamente mensuráveis tomando valores em $\mathbb{R}^{k \times m}$ tal que $\forall Z(\cdot) \in M^2(\mathbb{R}^{k \times m})$, tem-se

$$|Z(\cdot)|_{M^2(\mathbb{R}^{k \times m})} := \left\{ E \left[\int_0^T |Z(s)|^2 ds \right] \right\}^{\frac{1}{2}} < +\infty.$$

Observação 2.12. É importante ressaltar, que para cada $s \in [0, T]$, a norma de um elemento $z \in \mathbb{R}^{k \times m}$ é dada pela raiz quadrada do traço do produto dela com a sua transposta, ou seja,

$$|z| = \sqrt{\text{tr}(z^\top z)}.$$

Para $\alpha \geq 0$, defina $B^2 := S^2(\mathbb{R}^k) \times M^2(\mathbb{R}^{k \times m})$, em que

$$|Y(\cdot), Z(\cdot)|_{B^2} := \left\{ E \left[\sup_{0 \leq t \leq T} |Y(t)|^2 e^{\alpha t} \right] + E \left[\int_0^T |Z(s)|^2 e^{\alpha s} ds \right] \right\}^{\frac{1}{2}} < +\infty.$$

Pode-se sem maior dificuldade, concluir que B^2 é um espaço de Banach. Este fato será importante mais adiante.

Definição 2.13. Uma solução de (2.18), é um par de processos $(Y(\cdot), Z(\cdot))$, satisfazendo

- (1) $Y(\cdot)$ e $Z(\cdot)$ são $\{\mathcal{F}_t\}_{t \geq 0}$ - progressivamente mensuráveis,
- (2) $\int_0^T [|f(s, Y(s), Z(s))| + |Z(s)|^2] ds < +\infty \quad P \text{ a.s.},$
- (3) $Y(t) = \xi + \int_t^T f(s, Y(s), Z(s)) ds - \int_t^T Z(s) dW(s) \quad P \text{ a.s.} \quad 0 \leq t \leq T.$

Lema 2.14. Suponha que exista um processo estocástico $g \in M^2(\mathbb{R}^{k \times m})$ e uma constante $\lambda > 0$ tal que $|f(t, y, z)| \leq g(t) + \lambda(|y| + |z|)$, $\forall (t, y, z) \in [0, T] \times \mathbb{R}^k \times \mathbb{R}^{k \times m}$. Se $(Y(\cdot), Z(\cdot))$ é solução de (2.18) tal que $Z(\cdot) \in M^2(\mathbb{R}^{k \times m})$ então $Y(\cdot) \in S^2(\mathbb{R}^k)$.

Demonstração. De fato,

$$\begin{aligned} |Y(t)| &= \left| Y(0) - \int_0^t f(s, Y(s), Z(s)) ds + \int_0^t Z(s) dW(s) \right| \\ &\leq |Y(0)| + \int_0^t |f(s, Y(s), Z(s))| ds + \left| \sup_{0 \leq t \leq T} \int_0^t Z(s) dW(s) \right| \\ &\leq |Y(0)| + \int_0^t [|g(s)| + \lambda|Y(s)| + \lambda|Z(s)|] ds + \left| \sup_{0 \leq t \leq T} \int_0^t Z(s) dW(s) \right|. \end{aligned} \quad (2.20)$$

Faça

$$\alpha := |Y(0)| + \int_0^t [|g(t)| + \lambda|Z(t)|] ds + \left| \sup_{0 \leq t \leq T} \int_0^t Z(s) dW(s) \right|.$$

Por hipótese, g e $(Z(\cdot))$ estão em $M^2(\mathbb{R}^{k \times m})$, o mesmo acontecendo com $Y(0)$ por ser determinístico e portanto quadrado integrável. Por outro lado, pela desigualdade maximal de Doob (iv em [5], pág 14), segue que a última parcela de (2.20) é quadrado integrável. Assim, temos que $E|\alpha|^2 < +\infty$ e

$$|Y(t)| \leq \alpha + \int_0^t \lambda|Y(s)| ds.$$

Donde, pela desigualdade de Gronwall em (1.4), resulta que

$$\begin{aligned} |Y(t)| &\leq \alpha e^{\lambda t} \\ \Rightarrow \sup_{0 \leq t \leq T} |Y(t)|^2 &\leq \sup_{0 \leq t \leq T} \alpha^2 e^{2\lambda t} \\ \Rightarrow |Y(\cdot)|_{S^2(\mathbb{R}^k)}^2 &= E \left[\sup_{0 \leq t \leq T} |Y(t)|^2 \right] < +\infty. \end{aligned}$$

Logo $Y(\cdot) \in S^2(\mathbb{R}^k)$. □

Com o objetivo de analisar o principal teorema desta seção, considere a variável aleatória ξ nas mesmas condições da Proposição 2.11. Introduzimos as seguintes hipóteses:

(L1) Existe uma constante $\lambda > 0$ tal que

$$|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq \lambda(|y_1 - y_2| + |z_1 - z_2|)$$

$$\forall (t, y_1, z_1) \in [0, T] \times \mathbb{R}^k \times \mathbb{R}^{k \times m} \quad \forall (t, y_2, z_2) \in [0, T] \times \mathbb{R}^k \times \mathbb{R}^{k \times m} \quad P \text{ a.s.}$$

$$(L2) \quad E \left[|\xi|^2 + \int_0^T |f(s, 0, 0)|^2 ds \right] < +\infty \quad P \text{ a.s.}$$

No momento, suponha a existência de uma função F que não dependa de y e z . Considere então, $F \in M^2(\mathbb{R}^k)$, com $t \in [0, T]$. Nesta situação tem-se o seguinte lema:

Lema 2.15. Se ξ e F satisfazem L1 e L2 então existe uma solução única $(Y(\cdot), Z(\cdot)) \in B^2$, satisfazendo

$$Y(t) = \xi + \int_t^T F(s)ds - \int_t^T Z(s)dW(s), \quad 0 \leq t \leq T. \quad (2.21)$$

Demonstração. O objetivo é mostrar que existe um $Z(\cdot) \in M^2(\mathbb{R}^{k \times m})$, tal que $(Y(\cdot), Z(\cdot))$ é solução de (2.21)

Definindo $Y(t) := E \left[\xi + \int_t^T F(s)ds \middle| \mathcal{F}_t \right]$, tem-se

$$\begin{aligned} Y(t) &= E \left[\xi + \int_t^T F(s)ds \middle| \mathcal{F}_t \right] = E \left[\xi + \left(\int_0^T F(s)ds - \int_0^t F(s)ds \right) \middle| \mathcal{F}_t \right] \\ &= E \left[\xi + \int_0^T F(s)ds \middle| \mathcal{F}_t \right] - \int_0^t F(s)ds \middle| \mathcal{F}_t \\ &= E \left[\xi + \int_0^T F(s)ds \middle| \mathcal{F}_t \right] - \int_0^t F(s)ds = m(t) - \int_0^t F(s)ds, \end{aligned}$$

onde m é um *martingale* quadrado integrável e F é $\{\mathcal{F}_t\}_{t \geq 0}$ - progressivamente mensurável. Assim, pelo teorema da Representação de *martingale*, segue que existe $Z(\cdot) \in M^2(\mathbb{R}^{k \times m})$, tal que

$$m(t) = m(0) + \int_0^t Z(s)dW(s), \quad 0 \leq t \leq T.$$

Basta verificar agora que $(Y(\cdot), Z(\cdot))$ é a solução de (2.21). Ora, sabendo que $\xi = Y(T)$,

pode se escrever

$$\begin{aligned}
Y(t) &= m(t) - \int_0^t F(s)ds = m(0) + \int_0^t Z(s)dW(s) - \int_0^t F(s)ds \\
&= m(0) + \int_0^T Z(s)dW(s) - \int_t^T Z(s)dW(s) - \int_0^t F(s)ds + \int_t^T F(s)ds \\
&= m(0) + \int_0^T Z(s)dW(s) - \int_0^t F(s)ds - \int_t^T Z(s)dW(s) + \int_t^T F(s)ds.
\end{aligned}$$

Logo, fazendo $\xi = m(T) - \int_0^T F(s)ds$, garante-se a existência da solução.

A unicidade segue imediatamente da Proposição 2.11. $\square$

A seguir, apresentaremos um resultado fundamental da existência e unicidade da solução de uma BSDE.

Teorema 2.16. (Pardoux-Peng) Se as hipóteses $L1$ e $L2$ valem, então a BSDE (2.18) admite uma única solução $(Y(\cdot), Z(\cdot))$ em B^2 .

Demonstração. A ideia da prova é a utilização do argumento do ponto fixo para espaços de Banach. Com efeito, dado $(U, V) \in B^2$, defina $(Y, Z) = \phi(U, V)$, tal que o par (Y, Z) resolva a BSDE

$$Y(t) = \xi + \int_t^T f(s, U(s), V(s))ds - \int_t^T Z(s)dW(s), \quad 0 \leq t \leq T. \quad (2.22)$$

Observando que, fazendo $F(t) = f(t, U(t), V(t))$ e supondo que existe $f(\cdot, 0, 0) \in M^2(\mathbb{R}^{k \times m})$ tal que $|F(t)| = |f(t, U(t), V(t))| \leq f(t, 0, 0) + \lambda(|U(t)| + |V(t)|)$ para $\lambda > 0$, garante-se, juntando o Lema 2.14 e o Lema 2.15, que $(Y, Z) \in B^2$ é a única solução em (2.22). Além disso, como os processos em (2.22) são quadrados integráveis, resulta que $(Y, Z) = \phi(U, V)$ está bem definida. Assim, considere $(U_1, V_1) \in B^2$ e $(U_2, V_2) \in B^2$, tal que $(Y_1, Z_1) = \phi(U_1, V_1)$ e $(Y_2, Z_2) = \phi(U_2, V_2)$, onde $Y := Y_1 - Y_2$, $Z := Z_1 - Z_2$, $U := U_1 - U_2$ e $V := V_1 - V_2$. Daqui resulta que $Y(T) = 0$ e portanto, temos

$$dY(t) = -[f(t, U_1(t), V_1(t)) - f(t, U_2(t), V_2(t))]dt + Z(t)dW(t).$$

Por outro lado, aplicando a fórmula de Itô na função $g(t, Y(t)) := e^{\alpha t}|Y(t)|^2$, com $\alpha \geq 0$ e

$t \in [0, T]$, resulta que

$$\begin{aligned} d(e^{\alpha t}|Y(t)|^2) &= \alpha e^{\alpha t}|Y(t)|^2 dt + 2e^{\alpha t}Y(t)dY(t) + e^{\alpha t}(dY(t))^2 \\ &= \alpha e^{\alpha t}|Y(t)|^2 dt - 2e^{\alpha t}\langle Y(t), [f(t, U_1(t), V_1(t)) - f(t, U_2(t), V_2(t))] \rangle dt \\ &\quad + 2e^{\alpha t}\langle Y(t), Z(t)dW(t) \rangle + e^{\alpha t}|Z(t)|^2 dt, \end{aligned}$$

ou seja,

$$\begin{aligned} e^{\alpha t}|Y(t)|^2 &= \int_0^t \{ \alpha e^{\alpha s}|Y(s)|^2 - 2e^{\alpha s}\langle Y(s), [f(s, U_1(s), V_1(s)) - f(s, U_2(s), V_2(s))] \rangle \} ds \\ &\quad + \int_0^t 2e^{\alpha s}\langle Y(s), Z(s)dW(s) \rangle + \int_0^t e^{\alpha s}|Z(s)|^2 ds \\ &= \int_t^T e^{\alpha s} \{ -\alpha|Y(s)|^2 + 2\langle Y(s), [f(s, U_1(s), V_1(s)) - f(s, U_2(s), V_2(s))] \rangle \} ds \\ &\quad - \int_t^T 2e^{\alpha s}\langle Y(s), Z(s)dW(s) \rangle - \int_t^T e^{\alpha s}|Z(s)|^2 ds. \end{aligned}$$

Como f é Lipschitz, então $L1$ vale e portanto tem-se

$$\begin{aligned} e^{\alpha t}|Y(t)|^2 &\leq \int_t^T e^{\alpha s} [-\alpha|Y(s)|^2 + 2\lambda Y(s)(|U(s)| + |V(s)|)] ds \\ &\quad - \int_t^T 2e^{\alpha s}\langle Y(s), Z(s)dW(s) \rangle - \int_t^T e^{\alpha s}|Z(s)|^2 ds. \end{aligned}$$

Dado $\varepsilon > 0$ e tendo em conta a relação $2ab \leq \frac{a^2}{\varepsilon} + \varepsilon b^2$, para a e b reais, pode-se escrever

$$\begin{aligned} e^{\alpha t}|Y(t)|^2 &\leq \int_t^T e^{\alpha s} [-\alpha|Y(s)|^2 + 2\lambda Y(s)(|U(s)| + |V(s)|)] ds \\ &\quad - \int_t^T 2e^{\alpha s}\langle Y(s), Z(s)dW(s) \rangle - \int_t^T e^{\alpha s}|Z(s)|^2 ds \\ &\leq \int_t^T e^{\alpha s} \left[-\alpha|Y(s)|^2 + \frac{\lambda^2}{\varepsilon}|Y(s)|^2 + \varepsilon|U(s)|^2 + \frac{\lambda^2}{\varepsilon}|Y(s)|^2 + \varepsilon|V(s)|^2 \right] ds \end{aligned}$$

$$\begin{aligned}
& - \int_t^T 2e^{\alpha s} \langle Y(s), Z(s) dW(s) \rangle - \int_t^T e^{\alpha s} |Z(s)|^2 ds \\
& = \int_t^T e^{\alpha s} \left[\left(-\alpha + \frac{2\lambda^2}{\varepsilon} \right) |Y(s)|^2 + \varepsilon (|U(s)|^2 + |V(s)|^2) \right] ds \\
& - \int_t^T 2e^{\alpha s} \langle Y(s), Z(s) dW(s) \rangle - \int_t^T e^{\alpha s} |Z(s)|^2 ds.
\end{aligned}$$

Pondo $\alpha = \frac{2\lambda^2}{\varepsilon}$, resulta

$$\begin{aligned}
e^{\alpha t} |Y(t)|^2 + \int_t^T e^{\alpha s} |Z(s)|^2 ds & \leq \varepsilon \int_0^T e^{\alpha s} (|U(s)|^2 + |V(s)|^2) ds \\
& - 2 \int_t^T e^{\alpha s} \langle Y(s), Z(s) dW(s) \rangle. \tag{2.23}
\end{aligned}$$

Afirmação 2.16.1. Com $Y(\cdot) \in S^2(\mathbb{R}^k)$ e $Z(\cdot) \in M^2(\mathbb{R}^{k \times m})$, tem-se que

$$\int_t^T e^{\alpha s} \langle Y(s), Z(s) dW(s) \rangle,$$

é *martingale* quadrado integrável.

De fato, aplicando as desigualdades de Burkholder-Davis-Gundy com $c = \frac{1}{2}$ e $\tau = T$ em (2.6), temos que para $K > 0$,

$$\begin{aligned}
E \left[\sup_{0 \leq t \leq T} \left| \int_0^t e^{\alpha s} \langle Y(s), Z(s) dW(s) \rangle \right| \right] & \leq K E \left[\int_0^T e^{\alpha s} |Y(s)|^2 e^{\alpha s} |Z(s)|^2 ds \right]^{\frac{1}{2}} \\
& \leq K E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 \left(\int_0^T e^{\alpha s} |Z(s)|^2 ds \right)^{\frac{1}{2}} \right].
\end{aligned}$$

Sabendo que $|ab| \leq \frac{a^2}{2} + \frac{b^2}{2}$ para a e b reais, pode-se escrever

$$E \left[\sup_{0 \leq t \leq T} \left| \int_0^t e^{\alpha s} \langle Y(s), Z(s) dW(s) \rangle \right| \right] \leq \frac{1}{2} E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 \right] + \frac{K^2}{2} E \left[\int_0^T e^{\alpha s} |Z(s)|^2 ds \right]. \tag{2.24}$$

Como o membro da direita na desigualdade (2.24) é finito por hipótese, a Afirmção 2.16.1

vale. Denote agora

$$R(\varepsilon) := \varepsilon \int_0^T e^{\alpha s} (|U(s)|^2 + |V(s)|^2) ds.$$

Observe que para $t = 0$ na desigualdade (2.23), temos

$$E \left[\int_0^T e^{\alpha s} |Z(s)|^2 ds \right] \leq E[R(\varepsilon)]. \quad (2.25)$$

Assim por (2.24), pode-se afirmar que para $K > 0$,

$$\begin{aligned} E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 \right] &\leq E[R(\varepsilon)] + KE \left[\int_0^T e^{2\alpha s} |Y(s)|^2 |Z(s)|^2 ds \right]^{\frac{1}{2}} \\ &\leq E[R(\varepsilon)] + KE \left[\sup_{0 \leq t \leq T} e^{\frac{\alpha t}{2}} |Y(t)| \left(\int_0^T e^{\alpha s} |Z(s)|^2 ds \right)^{\frac{1}{2}} \right] \\ &\leq E[R(\varepsilon)] + \frac{1}{2} E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 \right] + \frac{K^2}{2} E \left[\int_0^T e^{\alpha s} |Z(s)|^2 ds \right]. \end{aligned}$$

Portanto, da desigualdade anterior e notando (2.25), conclui-se que

$$\begin{aligned} \frac{1}{2} E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 \right] &\leq E[R(\varepsilon)] + \frac{K^2 + 1}{2} E[R(\varepsilon)] - \frac{1}{2} E \left[\int_0^T e^{\alpha s} |Z(s)|^2 ds \right] \\ \Leftrightarrow \frac{1}{2} \left\{ E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 \right] + E \left[\int_0^T e^{\alpha s} |Z(s)|^2 ds \right] \right\} &\leq \frac{K^2 + 3}{2} E[R(\varepsilon)] \\ \Leftrightarrow E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 \right] + E \left[\int_0^T e^{\alpha s} |Z(s)|^2 ds \right] &\leq (K^2 + 3) E[R(\varepsilon)], \end{aligned}$$

isto é

$$\begin{aligned} E \left[\sup_{0 \leq t \leq T} e^{\alpha t} |Y(t)|^2 + \int_0^T e^{\alpha s} |Z(s)|^2 ds \right] &\leq (K^2 + 3) \varepsilon E \left[\int_0^T e^{\alpha s} (|U(s)|^2 + |V(s)|^2) ds \right] \\ &\leq LE \left[\sup_{0 \leq t \leq T} e^{\alpha t} |U(t)|^2 + \int_0^T e^{\alpha s} |V(s)|^2 ds \right], \quad (2.26) \end{aligned}$$

onde $L = (K^2 + 3)\varepsilon(1 \vee T)$. Assim para um $\varepsilon > 0$ e fazendo $L = \frac{1}{2}$, tem-se uma contração

em (2.26). Ou seja,

$$|Y, Z|_{B^2} = |\phi(U, V)|_{B^2} \leq \frac{\sqrt{2}}{2} |U, V|_{B^2}.$$

Pelo argumento do ponto fixo e sendo B^2 um espaço de Banach, (2.18) tem uma única solução $(Y(\cdot), Z(\cdot))$ em B^2 , provando assim o teorema. $\square$

No desenrolar deste trabalho, será importante falar também das BSDEs lineares, ou seja, quando a função geradora f em (2.18) tem forma linear. Neste caso e com as mesmas hipóteses do teorema anterior, o processo $Y(\cdot)$ pode ser explicitado. Apresenta-se então a versão linear com $k = 1$ para aliviar notação. A versão para $k \geq 1$ com a respectiva prova, pode ser encontrada no Teorema 2.2 em [11], pág 349 a 353.

Teorema 2.17. Sejam $A(\cdot)$ e $B(\cdot)$ processos em $\mathbb{R}$ e $\mathbb{R}^m$ respectivamente, progressivamente mensuráveis e limitados. Considere ainda $C(\cdot) \in M^2(\mathbb{R})$ e ξ uma variável $\mathcal{F}_T$ - mensurável, tal que $E[|\xi|^2] < +\infty$. Nestas condições, a BSDE

$$\begin{cases} dY(t) = -[A(t)Y(t) + \langle B(t), Z(t) \rangle + C(t)]dt + Z(t)dW(t) \\ Y(T) = \xi \quad P \text{ a.s.} \end{cases} \quad (2.27)$$

admite uma única solução $(Y(\cdot), Z(\cdot))$, o qual é dada por

$$Y(t) = L^{-1}(t)E\left[\xi L(T) + \int_t^T C(s)L(s)ds \middle| \mathcal{F}_t\right] \quad 0 \leq t \leq T, \quad (2.28)$$

$$\text{com } L(t) = \exp\left[\int_0^t B(s)dW(s) + \frac{1}{2}\int_0^t |B(s)|^2 ds + \int_0^t A(s)ds\right], \quad 0 \leq t \leq T.$$

Demonstração. Pela fórmula de Itô, pode-se deduzir que $L(t)$ verifica

$$\begin{cases} dL(t) = L(t)[A(t)dt + B(t)dW(t)] \\ L(0) = 1. \end{cases} \quad (2.29)$$

Além disso, pela desigualdade Maximal de Doob, temos que

$$|L(\cdot)|_{S^2}^2 = E\left[\sup_{0 \leq t \leq T} |L(t)|^2\right] \leq E\left[\sup_{0 \leq t \leq T} |L(t)|\right]^2 \leq KE\left[|L(t)|^2\right].$$

Como $A(\cdot)$ e $B(\cdot)$ são limitados por hipótese e notando a definição de $L(t)$ no enunciado do teorema, resulta que $|L(\cdot)|_{S^2}^2 < +\infty$ e portanto $L(\cdot) \in S^2$. Por outro lado, considerando

a função geradora $f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$, tal que $f(t, Y(t), Z(t)) = A(t)Y(t) + \langle B(t), Z(t) \rangle + C(t)$, pelo teorema anterior assegura-se a existência de uma única solução $(Y(\cdot), Z(\cdot)) \in B^2$ da BSDE (2.27). Resta então mostrar que (2.28) vale. Integrando por partes, resulta que

$$\begin{aligned} L(t)Y(t) - L(0)Y(0) &= \int_0^t (L(s)dY(s))ds + \int_0^t (dL(s)Y(s))ds \\ &\quad + \int_0^t (L(s)\langle B(s), Z(s) \rangle)ds. \end{aligned}$$

Por (2.27) e (2.29), temos

$$\begin{aligned} L(t)Y(t) &= Y(0) + \int_0^t (L(s)\langle B(s), Z(s) \rangle)ds - \int_0^t [L(s)f(t, Y(t), Z(t))]ds \\ &\quad + \int_0^t (L(s)Z(s))dW(s) + \int_0^t (Y(s)L(s)A(s))ds + \int_0^t (Y(s)L(s)B(s))dW(s). \end{aligned}$$

Como $f(t, Y(t), Z(t)) = A(t)Y(t) + \langle B(t), Z(t) \rangle + C(t)$, conclui-se que

$$L(t)Y(t) = Y(0) - \int_0^t (L(s)C(s))ds + \int_0^t (L(s)Z(s))dW(s) + \int_0^t (Y(s)L(s)B(s))dW(s),$$

isto é,

$$L(t)Y(t) - Y(0) + \int_0^t (L(s)C(s))ds = \int_0^t (L(s)Z(s))dW(s) + \int_0^t (Y(s)L(s)B(s))dW(s). \quad (2.30)$$

Pelas desigualdades de Burkholder-Davis-Gundy e aplicando a linearidade da esperança no membro da direita da igualdade anterior, observa-se de (2.24), que o *martingale* (2.30) é uniformemente integrável. Isto é,

$$E \left[\int_0^T |L(s)Z(s)|^2 ds \right]^{\frac{1}{2}} + E \left[\int_0^T |Y(s)L(s)B(s)|^2 ds \right]^{\frac{1}{2}} < +\infty.$$

Assim, aplicando outra vez a esperança, pode-se escrever

$$L(t)Y(t) - Y(0) + \int_0^t (L(s)C(s))ds = E \left[L(T)Y(T) - Y(0) + \int_0^T (L(s)C(s))ds \middle| \mathcal{F}_t \right],$$

ou seja,

$$L(t)Y(t) = E \left[L(T)Y(T) + \int_t^T (L(s)C(s))ds \middle| \mathcal{F}_t \right].$$

Isto conclui (2.28). $\square$

O último teorema analisado permite afirmar que se $Y(T) = \xi \geq 0$ e $C(\cdot) \geq 0$ a.s, então $Y(\cdot) \geq 0$ a.s. Isto motiva o próximo resultado, conhecido pelo Princípio da comparação.

Teorema 2.18. (Princípio da comparação) Considere $k = 1$ e seja (ξ_1, f_1) e (ξ_2, f_2) os pares associados às soluções (Y_1, Z_1) e (Y_2, Z_2) respectivas, em cada BSDE. Suponha que ambas estão nas condições do Teorema 2.16. Assuma ainda que $\xi_1 \leq \xi_2$ em $\mathbb{P}$ a.s e $f_1(t, Y_1, Z_1) \leq f_2(t, Y_1, Z_1)$ em $m_L \times \mathbb{P}$ a.s (m_L medida de Lebesgue). Nestas condições, tem-se $Y_1(t) \leq Y_2(t) \quad \forall t \in [0, T]$ $\mathbb{P}$ a.s.

Demonstração. Faça $Y(t) := Y_2(t) - Y_1(t)$, $Z(t) := Z_2(t) - Z_1(t)$ e $\xi := \xi_2 - \xi_1$. Considere a equação

$$Y(t) = \xi + \int_t^T [f_2(s, Y_2(s), Z_2(s)) - f_1(s, Y_1(s), Z_1(s))]ds - \int_t^T Z(s)dW(s), \quad (2.31)$$

e escreva

$$\begin{aligned} f_2(t, Y_2(t), Z_2(t)) - f_1(t, Y_1(t), Z_1(t)) &= f_2(t, Y_2(t), Z_2(t)) - f_2(t, Y_1(t), Z_2(t)) \\ &\quad + f_2(t, Y_1(t), Z_2(t)) - f_2(t, Y_1(t), Z_1(t)) \\ &\quad + f_2(s, Y_1(t), Z_1(t)) - f_1(t, Y_1(t), Z_1(t)). \end{aligned} \quad (2.32)$$

Seja A um processo tomando valores em $\mathbb{R}$ e B um processo com valores em $\mathbb{R}^m$, tais que $\forall t \in [0, T]$,

$$A(t) = \begin{cases} \frac{f_2(t, Y_2(t), Z_2(t)) - f_2(t, Y_1(t), Z_2(t))}{Y_2(t) - Y_1(t)} & \text{se } Y_1(t) \neq Y_2(t) \\ 0 & \text{se } Y_1(t) = Y_2(t) \end{cases}$$

e para cada $1 \leq i \leq m$, com $B(t) = (B^{(1)}(t), \dots, B^{(m)}(t))$

$$B^{(i)}(t) = \begin{cases} \frac{f_2(t, Y_1(t), Z_2^{(i)}(t)) - f_2(t, Y_1(t), Z_1^{(i)}(t))}{Z_2^{(i)}(t) - Z_1^{(i)}(t)} & \text{se } Z_1^{(i)}(t) \neq Z_2^{(i)}(t) \\ 0 & \text{se } Z_1^{(i)}(t) = Z_2^{(i)}(t). \end{cases}$$

Como por hipótese, f_2 é Lipschitz, os processos A e B são progressivamente mensuráveis. Além disso, os processos A e B são limitados, pois $\forall t \in [0, T]$, existe um $\lambda > 0$ tal que

$$|A(t)| = \frac{f_2(t, Y_2(t), Z_2(t)) - f_2(t, Y_1(t), Z_2(t))}{Y_2(t) - Y_1(t)} \leq \frac{\lambda(Y_2(t) - Y_1(t))}{Y_2(t) - Y_1(t)} = \lambda,$$

e

$$|B^{(i)}(t)| = \frac{f_2(t, Y_1(t), Z_2^{(i)}(t)) - f_2(t, Y_1(t), Z_1^{(i)}(t))}{Z_2^{(i)}(t) - Z_1^{(i)}(t)} \leq \frac{\lambda(Z_2^{(i)}(t) - Z_1^{(i)}(t))}{Z_2^{(i)}(t) - Z_1^{(i)}(t)} = \lambda$$

$$\Rightarrow |B(t)| \leq \lambda.$$

Assim, por (2.32) e com as definições de $A(t)$ e $B(t)$, pode afirmar-se que (2.31) tem forma linear, isto é

$$Y(t) = \xi + \int_t^T [A(s)Y(s) + \langle B(s), Z(s) \rangle + C(s)] ds - \int_t^T Z(s) dW(s). \quad (2.33)$$

Como f_1 é também Lipschitz e $f_1(t, Y_1, Z_1) \leq f_2(t, Y_1, Z_1)$, tem-se $C(t) = f_2(t, Y_1(t), Z_1(t)) - f_1(t, Y_1(t), Z_1(t)) \geq 0$, bem como, $\xi = \xi_2 - \xi_1 \geq 0$. Logo, aplicando o teorema anterior a (2.33), tem-se por (2.28) que $Y(t) \geq 0$, ou seja, $Y_1(t) \leq Y_2(t) \quad \forall t \in [0, T]$. $\square$

Observação 2.19. O Princípio da comparação permite facilmente concluir que se $\xi \geq 0$ e a função geradora da BSDE verifica $f(\cdot, 0, 0) \geq 0$ a.s, então $Y(\cdot) \geq 0$ a.s. Basta tomar $\xi_2 = \xi$, $\xi_1 = 0$, $f_2 = f$ e $f_1 = 0$ na prova anterior.

2.2 Princípio de Pontryagin Estocástico

Nesta seção exemplificamos a conexão que as BSDEs têm com problemas de controle, concretamente as BSDEs lineares, através do Princípio de Pontryagin estocástico e que desempenha um papel muito importante no controle de sistemas. Aqui, ao contrário do que foi abordado na seção 2 do *Capítulo 1*, existe um fator de difusão estocástico que está associado ao movimento Browniano definido na seção anterior.

2.2.1 Controle Estocástico

Antes de analisar o Princípio de Pontryagin, abordamos o problema de controle estocástico, o qual será analisado com pormenor, a existência e unicidade da solução forte do caso linear. Nesse sentido, considere o espaço de probabilidade $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ com as condições usuais e $W(\cdot)$ um movimento Browniano associado. Defina o sistema de controle através da seguinte equação diferencial linear

$$\begin{cases} dx(t) = [Ax(t) + Bu(t)]dt + [Cx(t) + Du(t)]dW(t), & t \in [0, T] \\ x(0) = x_0, \end{cases} \quad (2.34)$$

onde A, B, C e D são matrizes associadas a esse sistema. Além disso, o estado do sistema $x(\cdot)$ toma valores em $\mathbb{R}^n$, enquanto o controle $u(\cdot)$ está em $\mathcal{U}^L[0, T]$, tal que

$$\mathcal{U}^L[0, T] := \left\{ u : [0, T] \times \Omega \rightarrow U \text{ tal que } u(\cdot) \in L^2_{\mathcal{F}}([0, T]; \mathbb{R}^l) \right. \\ \left. \text{e } u(\cdot) \text{ é } \{\mathcal{F}_t\}_{t \geq 0} \text{ - adaptado, com } t \in [0, T] \text{ } P \text{ a.s.} \right\},$$

e em que $U \subseteq \mathbb{R}^l$. Neste contexto, existirá um controle ótimo $\bar{u}$ entre os possíveis pares admissíveis estocásticos $(x(\cdot), u(\cdot))$ (veja Definição 4.1 em [11] pág 63), que terá de minimizar,

$$J(u(\cdot)) := E \left[\int_0^T f(x(t), u(t))dt + h(x(T)) \right], \quad (2.35)$$

com $f : \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$. Chame a este problema de SL , ou seja, encontre $\bar{u}$ verificando a equação (2.34) tal que,

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}^L[0, T]} J(u(\cdot)).$$

Para este problema, assuma que

(SL1) $U \subseteq \mathbb{R}^l$ é um espaço convexo e compacto, além de que, as funções f e h são convexas.

(SL2) $U \subseteq \mathbb{R}^l$ é um espaço convexo e fechado, as funções f e h são convexas, além de que, existe $K, \delta > 0$, tal que

$$\begin{cases} f(x, u) \geq \delta|u|^2 - K \\ h(x) \geq -K \end{cases}, \quad \forall (x, u) \in \mathbb{R}^n \times U.$$

Nestas condições tem-se o seguinte resultado:

Teorema 2.20. Nas hipóteses $(SL1)$ ou $(SL2)$, se o problema SL é finito então admite um controle ótimo $\bar{u} \in \mathcal{U}_{ad}^L[0, T]$, tal que $J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}^L[0, T]} J(u(\cdot))$.

Demonstração. Por hipótese, o problema é finito, ou seja existe $c > 0$, tal que $c \leq J(u(\cdot))$, $\forall u(\cdot) \in \mathcal{U}_{ad}^L[0, T]$. Portanto, tome um par de sequencias admissíveis (x_N, u_N) com $N \in \mathbb{N}$ em que

$$\lim_{N \rightarrow +\infty} J(u_N(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}^L[0, T]} J(u(\cdot)),$$

e

$$\begin{cases} dx_N(t) = [Ax_N(t) + Bu_N(t)]dt + [Cx_N(t) + Du_N(t)]dW(t), & t \in [0, T] \\ x(0) = x_0. \end{cases}$$

Por $(SL2)$, tem-se com a norma euclideana de $\mathbb{R}^l$, que

$$|u_N(t)|^2 \leq \frac{K + f(x_N(t), u_N(t))}{\delta},$$

donde

$$E \left[\int_0^T |u_N(t)|^2 dt \right] \leq \frac{KT + E \left[\int_0^T f(x_N(t), u_N(t)) dt \right]}{\delta}.$$

Como $J(u_N(\cdot))$ é limitado, resulta que $E \left[\int_0^T |u_N(t)|^2 dt \right] < +\infty$, $\forall N \geq 1$, $t \in [0, T]$.

Assim, considerando uma subsequência $\tilde{u}_N(\cdot)$ de $u_N(\cdot)$, tal que $u_N(\cdot)$ converge fracamente para $\bar{u}$ em $L_{\mathcal{F}}^2([0, T]; \mathbb{R}^l)$, pelo resultado de Mazur (veja o Corolário 3.8 em [1], pág. 61), segue que $\tilde{u}_N(\cdot)$ converge fortemente para $\bar{u}$ em $L_{\mathcal{F}}^2([0, T]; \mathbb{R}^l)$ e pode ser escrita como combinação linear convexa de elementos de $u_N(\cdot)$, isto é

$$\tilde{u}_N(\cdot) = \sum_{i=1}^l a_{iN} u_{i+N}(\cdot), \quad a_{iN} \geq 0, \quad \sum_{i=1}^l a_{iN} = 1. \quad (2.36)$$

Assim, pode-se supor que $\tilde{u}_N(\cdot)$ converge para $\bar{u}$ quase certamente em medida de Lebesgue e P-a.s. Como U é convexo e fechado por hipótese, vem que $\bar{u}(t) \in U$, $\forall t \in [0, T]$, donde efetivamente, $\bar{u} \in \mathcal{U}_{ad}^L[0, T]$. Considere então a sequência $\tilde{u}_N(\cdot) \in \mathcal{U}_{ad}^L[0, T]$, definida em

(2.36). Associe a ela, a sequência dos estados $\tilde{x}_N(\cdot)$ respectiva, tal que

$$\begin{cases} d\tilde{x}_N(t) = [A\tilde{x}_N(t) + B\tilde{u}_N(t)]dt + [C\tilde{x}_N(t) + D\tilde{u}_N(t)]dW(t), & t \in [0, T] \\ \tilde{x}(0) = x_0. \end{cases}$$

Basta mostrar que $\bar{x}$ é $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptado e o par $(\bar{x}(\cdot), \bar{u}(\cdot))$ é solução da equação (2.34). Pela linearidade da equação e sabendo que $\tilde{u}_N \rightarrow \bar{u}$ em $L^2_{\mathcal{F}}([0, T]; \mathbb{R}^l)$, pode-se concluir que $\tilde{x}_N \rightarrow \bar{x}$, ou seja,

$$\lim_{N \rightarrow +\infty} |\tilde{x}_N - \bar{x}|^2 = \lim_{N \rightarrow +\infty} E \left[\sup_{0 \leq t \leq T} |\tilde{x}_N(t) - \bar{x}(t)|^2 \right] = 0. \quad (2.37)$$

Note que, por um lado

$$\begin{aligned} \lim_{N \rightarrow +\infty} \left| \int_0^t (B\tilde{u}_N(s) - B\bar{u}(s))ds \right|^2 &\leq \lim_{N \rightarrow +\infty} \int_0^t |B\tilde{u}_N(s) - B\bar{u}(s)|^2 ds \\ &\leq |B|^2 \lim_{N \rightarrow +\infty} \int_0^T |\tilde{u}_N(s) - \bar{u}(s)|^2 ds = 0, \end{aligned}$$

e por outro lado pelas desigualdades de Burkholder-Davis-Gundy, para $K \geq 0$

$$\begin{aligned} E \left[\sup_{0 \leq t \leq T} \left| \int_0^t (D\tilde{u}_N(s) - D\bar{u}(s))dW(s) \right|^2 \right] &\leq KE \left[\int_0^T |D\tilde{u}_N(s) - D\bar{u}(s)|^2 ds \right] \\ &\leq KE \left[|D|^2 \int_0^T |\tilde{u}_N(s) - \bar{u}(s)|^2 ds \right], \end{aligned}$$

donde

$$\lim_{N \rightarrow +\infty} E \left[\sup_{0 \leq t \leq T} \left| \int_0^t (D\tilde{u}_N(s) - D\bar{u}(s))dW(s) \right|^2 \right] = 0.$$

Logo, pela desigualdade triangular das normas, tem-se (2.37). Finalmente, pela convexidade de f e h , o fato de $(\bar{x}(\cdot), \bar{u}(\cdot))$ ser um par admissível para (2.34) e notando (2.36), pode-se escrever

$$\begin{aligned} J(\bar{u}(\cdot)) &= \lim_{N \rightarrow +\infty} J(\tilde{u}_N(\cdot)) = \lim_{N \rightarrow +\infty} J \left(\sum_{i=1}^l a_{iN} u_{i+N}(\cdot) \right) \\ &\leq \lim_{N \rightarrow +\infty} \sum_{i=1}^l a_{iN} J(u_{i+N}(\cdot)) = \sum_{i=1}^l a_{iN} \lim_{N \rightarrow +\infty} J(u_{i+N}(\cdot)) \end{aligned}$$

$$= \inf_{u(\cdot) \in \mathcal{U}_{ad}^L[0,T]} J(u(\cdot)),$$

isto é $J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}^L[0,T]} J(u(\cdot))$.

No caso de (SL1), a prova é idêntica, pois sendo U compacto, este é limitado e portanto para qualquer sequência $u_N(t)$ de U tem-se $|u_N(t)| \leq K$, com $t \in [0, T]$. $\square$

Observação 2.21. Ficou claro na prova do teorema anterior que o fato de U ser convexo e fechado, garante a convergência forte de $\tilde{u}_N$ e consequentemente pela linearidade de (2.34), a convergência forte de $\tilde{x}_N$. No entanto, essa convergência pode não existir, caso a equação dos estados (2.34) seja não linear. Neste sentido, a estratégia será a utilização da formulação fraca para o problema de controle estocástico.

Para um problema não necessariamente linear, a equação dos estados pode ser descrita seguinte maneira:

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), & t \in [0, T] \\ x(0) = x_0, \end{cases} \quad (2.38)$$

em que $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$ e $\sigma : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^{n \times m}$ são funções mensuráveis, $\forall t \in [0, T]$. Além disso, sendo $x(\cdot) \in \mathbb{R}^n$ e $u(\cdot) \in \mathcal{U}[0, T]$, $\mathcal{U}[0, T]$ tem a mesma construção do caso linear. Chamemos este problema de S . Para definir o espaço dos controles admissíveis, não basta fixar apenas o espaço de probabilidade com filtração associado a um movimento Browniano à priori, pelas razões já mencionadas anteriormente. Para a formulação fraca, $u(\cdot)$ fará variar $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, donde para qualquer controle admissível $\pi \in \mathcal{U}_{ad}[0, T]$, teremos a 6-tupla $\pi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P, W(\cdot), u(\cdot))$ (veja Definição 4.2 em [11] pág 63). Assim, considerando hipóteses análogas às hipóteses D no controle determinístico do *Capítulo 1* (também para σ), existirá um controle ótimo $\bar{\pi} = (\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}_{t \geq 0}, \bar{P}, \bar{W}(\cdot), \bar{u}(\cdot))$, tal que

$$J(\bar{\pi}) = \inf_{\pi \in \mathcal{U}_{ad}[0,T]} J(\pi).$$

Para resultado e prova da existência de um controle ótimo $\bar{\pi}$ para o problema estocástico geral, veja o Teorema 5.3 em [11] pág. 71-75.

2.2.2 BSDE e Controle

Antes de falar do Princípio de Pontryagin estocástico, apresentam-se dois resultados que mostram como as BSDEs podem ser úteis em problemas de controle estocástico.

Lema 2.22. Sejam (ξ, f) e (ξ^α, f^α) com $\alpha \in I$. Para cada (ξ^α, f^α) associe a respectiva solução (Y^α, Z^α) de uma BSDE. Suponha que existe $\hat{\alpha} \in I$ tal que

$$f(t, Y(t), Z(t)) = \operatorname{ess\,inf}_{\alpha \in I} f^\alpha(t, Y(t), Z(t)) = f^{\hat{\alpha}}(t, Y(t), Z(t)) \quad , \quad m_L \times P \text{ a.s.},$$

e

$$\xi = \operatorname{ess\,inf}_{\alpha \in I} \xi^\alpha = \xi^{\hat{\alpha}}.$$

Então

$$Y(t) = \operatorname{ess\,inf}_{\alpha \in I} Y^\alpha(t) = Y^{\hat{\alpha}}(t) \quad , \quad P \text{ a.s.}$$

Demonstração. Para provar o lema, será importante o conceito de variáveis $\mathcal{F}$ - mensuráveis não negativas (veja Teorema 1.1.1 em [4], pág 3-4). Seja $\{X^\alpha, \alpha \in I\}$ uma família de variáveis aleatórias, sendo I o conjunto de índices, no máximo não numerável. Nestas condições, existe uma única variável aleatória Y , $\mathcal{F}$ - mensurável, tal que

- (1) $Y \leq X^\alpha \quad , \quad \forall \alpha \in I \quad P \text{ a.s.},$
- (2) Se X é $\mathcal{F}$ - mensurável e $X \leq X^\alpha \quad \forall \alpha \in I$ então $Y \geq X \quad P \text{ a.s.}$

A variável Y é denominada por $\operatorname{ess\,inf}_{\alpha \in I} X^\alpha$. Assim neste caso, como $\xi \leq \xi^\alpha$ e $f(t, Y(t), Z(t)) \leq f^\alpha(t, Y(t), Z(t))$, tem-se, pelo Princípio da comparação, que $Y(t) \leq Y^\alpha(t)$, isto é, $\operatorname{ess\,inf}_{\alpha \in I} Y^\alpha \geq Y$. Por, outro lado, por hipótese, existe $\hat{\alpha} \in I$ tal que $f(t, Y(t), Z(t)) = f^{\hat{\alpha}}(t, Y(t), Z(t))$ e $\xi = \xi^{\hat{\alpha}}$, então pode-se afirmar que (Y, Z) e $(Y^{\hat{\alpha}}, Z^{\hat{\alpha}})$ são duas soluções da mesma BSDE associada a $(\xi^{\hat{\alpha}}, f^{\hat{\alpha}})$. Devido à unicidade da solução da BSDE, temos que $Y = Y^{\hat{\alpha}}$ e $Z = Z^{\hat{\alpha}}$. Resulta então que $\operatorname{ess\,inf}_{\alpha \in I} Y^\alpha \leq Y$, ou seja

$$Y(t) = \operatorname{ess\,inf}_{\alpha \in I} Y^\alpha(t) = Y^{\hat{\alpha}}(t) \quad , \quad P \text{ a.s.} \quad \forall t \in [0, T].$$

□

Assumindo a concavidade da função geradora f de uma BSDE, podemos provar que a solução desse tipo de equações diferenciais estocásticas representa na sua essência, a função valor em problemas de controle. Este fato será perceptível no *Capítulo 3*, quando se abordar a relação entre PDEs e BSDEs. Nesse sentido, considere $k = 1$ e tome a função geradora $f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$, côncava em $\mathbb{R} \times \mathbb{R}^m$. Seja (Y, Z) , a solução da BSDE associada ao par (ξ, f) . Considere ainda a função convexa $F : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ como sendo a transformada de Fenchel-Legendre, definida por

$$F(t, b, c) := \sup_{(y, z) \in \mathbb{R} \times \mathbb{R}^m} [f(t, y, z) - yb - \langle z, c \rangle].$$

Como f é uma função côncava em $\mathbb{R} \times \mathbb{R}^m$, pode-se fazer a seguinte transformação, mantendo toda a informação de F , isto é

$$f(t, y, z) = \inf_{(b, c) \in \mathbb{R} \times \mathbb{R}^m} [F(t, b, c) + yb + \langle z, c \rangle].$$

Seja $\mathcal{A}$ o conjunto de todos os processos estocásticos limitados e $\{\mathcal{F}_t\}_{t \geq 0}$ - progressivamente mensuráveis, com $(\beta, \gamma) : [0, T] \times \Omega \rightarrow \mathbb{R} \times \mathbb{R}^m$ dois processos em $\mathcal{A}$, tal que

$$E \left[\int_0^T |F(t, \beta(t), \gamma(t))|^2 dt \right] < +\infty.$$

Seja $f^{(\beta, \gamma)} : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$, uma família de geradores lineares definida por

$$f^{(\beta, \gamma)}(t, y, z) := F(t, \beta(t), \gamma(t)) + y\beta(t) + \langle z, \gamma(t) \rangle, \quad (\beta, \gamma) \in \mathcal{A}, \quad \forall t \in [0, T],$$

tal que $(Y^{(\beta, \gamma)}, Z^{(\beta, \gamma)})$ é solução da BSDE associada a $(\xi, f^{(\beta, \gamma)})$. O lema anterior fornece o seguinte resultado:

Teorema 2.23. Seja $\xi, \mathcal{F}_T$ - mensurável tal que $E[|\xi|^2] < +\infty$ e f um gerador côncavo satisfazendo as condições do Teorema 2.16. Seja ainda (Y, Z) a solução da BSDE linear associada ao par (ξ, f) . Nestas condições, tem-se

$$f(t, Y(t), Z(t)) = \operatorname{ess\,inf}_{(\beta, \gamma) \in \mathcal{A}} f^{(\beta, \gamma)}(t, Y(t), Z(t)), \quad (2.39)$$

onde

$$Y^{(\beta, \gamma)}(t) = E_{Q^\gamma} \left[\xi \exp \left(\int_t^T \beta(s) ds \right) + \int_t^T \exp \left(\int_t^s \beta(u) du \right) F(s, \beta(s), \gamma(s)) ds \middle| \mathcal{F}_t \right], \quad (2.40)$$

tal que $Q^\gamma \ll P$ e o processo $L(t) = E \left[\frac{dQ^\gamma}{dP} \middle| \mathcal{F}_t \right]$ satisfaz a SDE

$$\begin{cases} dL(t) = L(t)\gamma(t)dW(t) \\ L(0) = 1 \end{cases}, \quad \forall t \in [0, T]. \quad (2.41)$$

Demonstração. Pela definição da geradora f na transformada de Fenchel-Legendre, segue que

$$f(t, Y(t), Z(t)) = \inf_{(b, c) \in \mathbb{R} \times \mathbb{R}^m} [F(t, b, c) + Y(t)b + \langle Z(t), c \rangle]$$

$$\leq F(t, b, c) + Y(t)b + \langle Z(t), c \rangle,$$

ou seja

$$\begin{aligned} f(t, Y(t), Z(t)) &\leq F(t, \beta(t), \gamma(t)) + Y(t)\beta(t) + \langle Z(t), \gamma(t) \rangle \\ &= f^{(\beta, \gamma)}(t, Y(t), Z(t)) \text{ , } P \text{ a.s. , } 0 \leq t \leq T, \end{aligned}$$

donde $f(t, Y(t), Z(t)) \leq \text{ess inf}_{(\beta, \gamma) \in A} f^{(\beta, \gamma)}(t, Y(t), Z(t))$, $\forall t \in [0, T]$. Por outro lado, como F é convexa, tem crescimento linear e portanto para cada $(t, \omega, y, z) \in [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^m$, resulta que o ínfimo de $F(t, b, c) + yb + \langle z, c \rangle$ é atingido em

$$(\tilde{b}(t, y, z), \tilde{c}(t, y, z)) \in \mathbb{R} \times \mathbb{R}^m.$$

Considerando a função $-f$ convexa, resulta que para cada $t \in [0, T]$, $(\tilde{b}(y, z), \tilde{c}(y, z))$, pertence ao subdiferencial de $-f$. Sendo

$$D_{b,c}^{1,-}(-f)(b, c) := \left\{ (p, q) \in \mathbb{R} \times \mathbb{R}^m \mid \lim_{\substack{u \rightarrow b \\ v \rightarrow c}} \frac{(-f)(u, v) - (-f)(b, c) - p(u - b) - \langle q, v - c \rangle}{|u - b| + |v - c|} \geq 0 \right\},$$

temos que, como f é Lipschitz, verifica-se a hipótese $L1$ da página 44. Portanto para $(p, q) = (\tilde{b}(y, z), \tilde{c}(y, z))$, existe $\lambda > 0$ tal que

$$|(\tilde{b}(t, y, z), \tilde{c}(t, y, z))| \leq |\nabla(-f(\cdot))| \leq |\nabla f(\cdot)| \leq \lambda. \quad (2.42)$$

Isto é, $(\tilde{b}(y, z), \tilde{c}(y, z))$ é limitado. Logo, tendo em conta (2.42), escrevemos

$$\begin{aligned} f(t, y, z) &= F(t, \tilde{b}(t, y, z), \tilde{c}(t, y, z)) + y\tilde{b}(t, y, z) + \langle z, \tilde{c}(t, y, z) \rangle \\ \Rightarrow f(t, Y(t), Z(t)) &= F[t, \tilde{b}(t, Y(t), Z(t)), \tilde{c}(t, Y(t), Z(t))] + Y(t)\tilde{b}(t, Y(t), Z(t)) \\ &\quad + \langle Z(t), \tilde{c}(t, Y(t), Z(t)) \rangle \text{ , } 0 \leq t \leq T. \end{aligned}$$

Através do resultado da seleção mensurável de Filipov (veja Teorema 2.3 em [10] pág. 36-37), existe um par de funções implícitas mensuráveis $(\hat{\beta}, \hat{\gamma}) \in A$, tal que

$$f(t, Y(t), Z(t)) = f^{(\hat{\beta}, \hat{\gamma})}(t, Y(t), Z(t)) = F(t, \hat{\beta}(t), \hat{\gamma}(t)) + Y(t)\hat{\beta}(t) + \langle Z(t), \hat{\gamma}(t) \rangle,$$

ou seja, $f(t, Y(t), Z(t)) \geq \text{ess inf}_{(\beta, \gamma) \in A} f^{(\beta, \gamma)}(t, Y(t), Z(t))$, $\forall t \in [0, T]$. Logo está provado (2.39). Para demonstrar (2.40), tendo em conta que $\xi = \xi^{(\beta, \gamma)}$ e (2.39), segue pelo teorema

anterior que

$$Y(t) = \operatorname{ess\,inf}_{(\beta, \gamma) \in A} Y^{(\beta, \gamma)}(t), \quad P \text{ a.s. } \forall t \in [0, T].$$

Assim, $Y^{(\beta, \gamma)}$ é solução de uma BSDE linear associada a $(\xi, f^{(\beta, \gamma)})$, onde

$$f^{(\beta, \gamma)}(t, y, z) = F(t, \beta(t), \gamma(t)) + y\beta(t) + \langle z, \gamma(t) \rangle, \quad (\beta, \gamma) \in A, \quad \forall t \in [0, T].$$

Daí, pelo Teorema 2.17, vem que

$$\bar{L}(t)Y^{(\beta, \gamma)}(t) = E \left[\xi \bar{L}(T) + \int_t^T \bar{L}(s)F(s, \beta(s), \gamma(s))ds \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T.$$

Assim tendo em conta que $d\bar{L}(t) = \bar{L}(t)[\beta(t)dt + \gamma(t)dW(t)]$, (2.41) e a fórmula de Itô, tem-se

$$\bar{L}(t) = \exp \left(\int_0^t \beta(s)ds \right) L(t),$$

donde usando a extensão do resultado de Bayes (veja 34 em [9] pág. 231), chega-se à relação (2.40). $\square$

Com o objetivo de analisar o Princípio de Pontryagin conhecido por Princípio do máximo, introduza no problema S , as seguintes hipóteses:

(S*1) (U, d) é um espaço métrico separável e convexo em $\mathbb{R}^l$.

(S*2) Existe uma constante $L > 0$ e existe um módulo de continuidade $\bar{w} : [0, +\infty) \rightarrow [0, +\infty)$ tal que para $\varphi(t, x, u) := b(t, x, u)$; $\varphi(t, x, u) := \sigma(t, x, u)$; $\varphi(t, x, u) := f(t, x, u)$; $\varphi(t, x, u) := h(x)$ se tem,

$$\left\{ \begin{array}{l} |\varphi(t, x_1, u_1) - \varphi(t, x_2, u_2)| \leq L|x_1 - x_2| + \bar{w}(d(u_1, u_2)) \\ \forall t \in [0, T]; \quad x_1, x_2 \in \mathbb{R}^n; \quad u_1, u_2 \in U \\ \\ |\varphi(t, 0, u)| \leq L \\ \forall (t, u) \in [0, T] \times U. \end{array} \right.$$

(S*3) As funções $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^{n \times m}$, $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são de classe C^2 em x e existe um módulo de continuidade $\bar{w} : [0, +\infty) \rightarrow [0, +\infty)$ tal que para $\varphi(t, x, u) := b(t, x, u)$; $\varphi(t, x, u) :=$

$\sigma(t, x, u); \varphi(t, x, u) := f(t, x, u); \varphi(t, x, u) := h(x)$ se tem,

$$\left| \frac{\partial \varphi}{\partial x}(t, x_1, u_1) - \frac{\partial \varphi}{\partial x}(t, x_2, u_2) \right| \leq \bar{w}(|x_1 - x_2| + d(u_1, u_2))$$

$$\left| \frac{\partial^2 \varphi}{\partial x^2}(t, x_1, u_1) - \frac{\partial^2 \varphi}{\partial x^2}(t, x_2, u_2) \right| \leq \bar{w}(|x_1 - x_2| + d(u_1, u_2))$$

$$\forall t \in [0, T]; \quad x_1, x_2 \in \mathbb{R}^n; \quad u_1, u_2 \in U.$$

(S*4) As funções f , b e σ são localmente Lipschitz e as suas derivadas parciais em x são contínuas $\forall (x, u) \in \mathbb{R}^n \times U$.

Para um par de processos $(Y(\cdot), Z(\cdot))$ $\{\mathcal{F}_t\}_{t \geq 0}$ - adaptados, considere as seguintes equações adjuntas associadas ao Hamiltoniano H na vertente estocástica, que nada mais são, que uma BSDE linear.

$$\begin{cases} dY(t) = -\frac{\partial H}{\partial x}(t, x(t), u(t), Y(t), Z(t))dt + Z(t)dW(t), & t \in [0, T] \\ Y(T) = -\frac{dh}{dx}(x(T)), \end{cases} \quad (2.43)$$

onde

$$\begin{aligned} H(t, x, u, p, q) &:= \langle b(t, x, u), p \rangle + \text{tr}[\sigma^\top(t, x, u) q] - f(t, x, u) \\ 0 \leq t \leq T; \quad x &\in \mathbb{R}^n; \quad u \in U; \quad p \in \mathbb{R}^n; \quad q \in \mathbb{R}^n \times \mathbb{R}^m. \end{aligned} \quad (2.44)$$

Nestas condições, garante-se então que o seguinte resultado vale.

Teorema 2.24. (Princípio do máximo estocástico, condição necessária) Seja o espaço de probabilidade $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ com $W(t)$ o Movimento Browniano associado. Suponha que valem as hipóteses S*1 a S*4 e seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ o par ótimo do problema S . Considere os processos $(\bar{Y}, \bar{Z})$ como sendo a solução de (2.43) associado a esse par ótimo, tal que

$$\begin{cases} d\bar{Y}(t) = -\frac{\partial H}{\partial x}(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t))dt + \bar{Z}(t)dW(t), & t \in [0, T] \\ \bar{Y}(T) = -\frac{dh}{dx}(dx(T)). \end{cases} \quad (2.45)$$

Então

$$\mathcal{H}(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) = \max_{u \in U} \mathcal{H}(t, \bar{x}(t), u(t), \bar{Y}(t), \bar{Z}(t)),$$

onde

$$\begin{aligned} \mathcal{H}(t, \bar{x}(t), u(t), \bar{Y}(t), \bar{Z}(t)) &:= G(t, \bar{x}(t), u(t), \bar{Y}(t), \bar{p}(t)) \\ &+ tr\{\sigma^\top(t, \bar{x}, u) [\bar{Z}(t) - \bar{p}(t)\sigma(t, \bar{x}, \bar{u})]\}, \end{aligned} \quad (2.46)$$

sendo

$$\begin{aligned} G(t, x, u, p, \tilde{p}) &:= \langle b(t, x, u), p \rangle + \frac{1}{2} tr[\sigma^\top(t, x, u) \tilde{p} \sigma(t, x, u)] - f(t, x, u) \\ 0 \leq t \leq T; \quad x &\in \mathbb{R}^n; \quad u \in U; \quad p \in \mathbb{R}^n; \quad \tilde{p} \in S^n. \end{aligned} \quad (2.47)$$

Observação 2.25. Uma abordagem a este resultado com as respectivas considerações e prova, pode ser encontrada em [11] pág 118-137. O leitor mais atento observa que no caso estocástico, o par ótimo $(\bar{x}(\cdot), \bar{u}(\cdot))$ realiza o máximo na função $\mathcal{H}$ definida em (2.46). Supondo que σ em (2.38) dependa do controle $u(\cdot)$ e não assumindo a concavidade de H em (2.44), pode suceder que o controle ótimo $\bar{u}$ do problema S não maximize o Hamiltoniano H . O seguinte exemplo esclarece esse fato.

Exemplo 2.26. Faça $n = m = 1$ e considere o sistema

$$\begin{cases} dx(t) = 3u(t)dW(t), & t \in [0, 2] \\ x(0) = 0, \end{cases} \quad (2.48)$$

tal que $u \in [-1, 1]$. Para este sistema, considere a função objetivo

$$J(u(\cdot)) = E \left[\int_0^2 \left(x^2(t) - \frac{1}{4}u^2(t) \right) dt + x^2(2) \right]. \quad (2.49)$$

Para determinar o par ótimo que minimiza a função de custo anterior, de (2.48), tem-se que

$$J(u(\cdot)) = E \left\{ \int_0^2 \left[\left(\int_0^t 3u(s)dW(s) \right)^2 - \frac{1}{4}u^2(t) \right] dt + \left(\int_0^2 3u(t)dW(t) \right)^2 \right\},$$

ou seja, aplicando a isometria de Itô e o teorema de Fubini, resulta que

$$J(u(\cdot)) = E \left[\int_0^2 \left(\int_0^t 9u^2(s)ds - \frac{1}{4}u^2(t) \right) dt + \int_0^2 9u^2(t)dt \right]$$

$$\begin{aligned}
&= E \left[\int_0^2 \left(\int_0^t 9u^2(s) ds \right) dt + \int_0^2 \frac{35}{4} u^2(t) dt \right] \\
&= E \left[\int_0^2 9u^2(s) \left(\int_s^2 dt \right) ds + \int_0^2 \frac{35}{4} u^2(s) ds \right] \\
&= E \left[\int_0^2 (18 - 9s) u^2(s) ds + \int_0^2 \frac{35}{4} u^2(s) ds \right] \\
&= E \left[\int_0^2 \left(\frac{107}{4} - 9s \right) u^2(s) ds \right].
\end{aligned}$$

Assim, o par que minimiza $J(u(\cdot))$ é $(\bar{x}, \bar{u}) = (0, 0)$. Por outro lado, por (2.44), tendo em conta (2.48) e (2.49), vem imediatamente que

$$H(t, \bar{x}, u, p, q) = \frac{1}{4}u^2 + 3qu.$$

Portanto H é convexa e não poderá ser maximizada em $\bar{u} = 0$, $\forall t \in [0, 2]$, P a.s.

Teorema 2.27. (Princípio do máximo estocástico, condição suficiente) Seja o espaço de probabilidade $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ com $W(t)$ o Movimento Browniano associado. Suponha que valem as hipóteses S^*1 a S^*4 e seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ um par admissível do problema S . Suponha ainda h convexa e H côncava. Nestas condições, se

$$H(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) = \max_{u \in U} H(t, \bar{x}(t), u(t), \bar{Y}(t), \bar{Z}(t)),$$

no par $(\bar{x}(\cdot), \bar{u}(\cdot))$, tal que $(\bar{Y}, \bar{Z})$ é solução de (2.43), então esse par é ótimo para o problema S , ou seja

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}[0, T]} J(u(\cdot)).$$

Demonstração. Assuma que $(\bar{x}(\cdot), \bar{u}(\cdot))$ é um par, tal que $(\bar{Y}, \bar{Z})$ é solução de

$$\begin{cases} d\bar{Y}(t) = -\frac{\partial H}{\partial x}(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t))dt + \bar{Z}(t)dW(t), & t \in [0, T] \\ \bar{Y}(T) = -\frac{dh}{dx}(\bar{x}(T)). \end{cases} \quad (2.50)$$

Faça $n = m = 1$ para simplicidade de notação. Ora,

$$J(u) - J(\bar{u}) = E \left[\int_0^T f(t, x(t), u(t))dt + h(x(T)) \right]$$

$$\begin{aligned}
& - E \left[\int_0^T f(t, \bar{x}(t), \bar{u}(t)) dt + h(\bar{x}(T)) \right] \\
& = E \left[\int_0^T [f(t, x(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] dt \right] \\
& + E \left[h(x(T)) - h(\bar{x}(T)) \right].
\end{aligned}$$

Analise-se as esperanças $E_1(u; \bar{u})$ e $E_2(u; \bar{u})$, onde

$$E_1(u; \bar{u}) := E \left[\int_0^T [f(t, x(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] dt \right], \quad (2.51)$$

e

$$E_2(u; \bar{u}) := E \left[h(x(T)) - h(\bar{x}(T)) \right]. \quad (2.52)$$

Pela definição de H em (2.44) e (2.51), temos que

$$\begin{aligned}
E_1(u; \bar{u}) &= E \left[\int_0^T [H(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) - H(t, x(t), u(t), \bar{Y}(t), \bar{Z}(t))] dt \right] \\
&- E \left[\int_0^T [b(t, \bar{x}(t), \bar{u}(t)) - b(t, x(t), u(t))] \bar{Y}(t) dt \right] \\
&- E \left[\int_0^T [\sigma(t, \bar{x}(t), \bar{u}(t)) - \sigma(t, x(t), u(t))] \bar{Z}(t) dt \right].
\end{aligned} \quad (2.53)$$

Por outro lado, de (2.52), resulta pela convexidade de h que

$$\begin{aligned}
h(\bar{x}(T)) - h(x(T)) &\leq (\bar{x}(T) - x(T)) \frac{dh}{dx}(\bar{x}(T)) \\
\Leftrightarrow h(x(T)) - h(\bar{x}(T)) &\geq (\bar{x}(T) - x(T)) \bar{Y}(T),
\end{aligned}$$

donde

$$E_2(u; \bar{u}) \geq E \left[(\bar{x}(T) - x(T)) \bar{Y}(T) \right].$$

Além disso, por integração por partes, sabendo que $(\bar{x}(0) - x(0)) \bar{Y}(0) = 0$, tendo em conta a equação dos estados (2.38) e a BSDE (2.50), segue que

$$(\bar{x}(T) - x(T)) \bar{Y}(T) = \int_0^T (\bar{x}(t) - x(t)) d\bar{Y}(t) + \int_0^T \bar{Y}(t) d(\bar{x}(t) - x(t))$$

$$+ \int_0^T [\sigma(t, \bar{x}(t), \bar{u}(t)) - \sigma(t, x(t), u(t))] \bar{Z}(t) dt,$$

onde

$$\begin{aligned} \int_0^T (\bar{x}(t) - x(t)) d\bar{Y}(t) &= - \int_0^T (\bar{x}(t) - x(t)) \left(\frac{\partial H}{\partial x}(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) \right) dt \\ &+ \int_0^T (\bar{x}(t) - x(t)) \bar{Z}(t) dW(t), \end{aligned}$$

e

$$\begin{aligned} \int_0^T \bar{Y}(t) d(\bar{x}(t) - x(t)) &= \int_0^T \bar{Y}(t) [b(t, \bar{x}(t), \bar{u}(t)) - b(t, x(t), u(t))] dt \\ &+ \int_0^T \bar{Y}(t) [\sigma(t, \bar{x}(t), \bar{u}(t)) - \sigma(t, x(t), u(t))] dW(t). \end{aligned}$$

Passando à esperança, temos imediatamente

$$E \left[\int_0^T (\bar{x}(t) - x(t)) d\bar{Y}(t) \right] = -E \left[\int_0^T (\bar{x}(t) - x(t)) \left(\frac{\partial H}{\partial x}(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) \right) dt \right],$$

e

$$E \left[\int_0^T \bar{Y}(t) d(\bar{x}(t) - x(t)) \right] = E \left[\int_0^T \bar{Y}(t) [b(t, \bar{x}(t), \bar{u}(t)) - b(t, x(t), u(t))] dt \right].$$

Sendo assim, tem-se

$$\begin{aligned} E \left[(\bar{x}(T) - x(T)) \bar{Y}(T) \right] &= -E \left[\int_0^T (\bar{x}(t) - x(t)) \left(\frac{\partial H}{\partial x}(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) \right) dt \right] \\ &+ E \left[\int_0^T \bar{Y}(t) [b(t, \bar{x}(t), \bar{u}(t)) - b(t, x(t), u(t))] dt \right] \\ &+ E \left[\int_0^T [\sigma(t, \bar{x}(t), \bar{u}(t)) - \sigma(t, x(t), u(t))] \bar{Z}(t) dt \right]. \end{aligned} \quad (2.54)$$

Logo, notando (2.51) e (2.52), resulta que

$$\begin{aligned} J(u) - J(\bar{u}) &= E_1(u; \bar{u}) + E_2(u; \bar{u}) \geq E \left[\int_0^T [f(t, x(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] dt \right] \\ &\quad + E \left[(\bar{x}(T) - x(T)) \bar{Y}(T) \right]. \end{aligned}$$

Ou seja, juntando (2.53) e (2.54), temos que

$$\begin{aligned} J(u) - J(\bar{u}) &\geq E \left[\int_0^T [H(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) - H(t, x(t), u(t), \bar{Y}(t), \bar{Z}(t))] dt \right] \\ &\quad - E \left[\int_0^T (\bar{x}(t) - x(t)) \left(\frac{\partial H}{\partial x}(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) \right) dt \right]. \end{aligned} \quad (2.55)$$

Como $H(t, \bar{x}(t), \bar{u}(t), \bar{Y}(t), \bar{Z}(t)) = \max_{u \in U} H(t, \bar{x}(t), u(t), \bar{Y}(t), \bar{Z}(t))$ e por hipótese H é côncava, conclui-se que o membro da direita de (2.55) é positivo. Logo, tem-se $J(u) \geq J(\bar{u})$, donde

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}[0, T]} J(u(\cdot)).$$

□

Capítulo 3

BSDE e PDE

Este capítulo representa o objetivo principal de estudo da dissertação. Será analisado a importância que as BSDEs têm nas soluções de certas PDEs, através da fórmula de Feynman-Kac. Será perceptível o papel que a programação dinâmica tem nesse sentido, aliado ao problema de controle.

3.1 Programação Dinâmica estocástica

Igualmente, como no *Capítulo 1* para o caso determinístico, a programação dinâmica estocástica é uma ferramenta poderosa em toda esta teoria, sendo que neste caso, ficará claro a sua aplicabilidade nas BSDEs e PDEs, através de resultados que serão analisados posteriormente. O caso estocástico é de certa forma similar ao caso determinístico, com as devidas adaptações referentes ao fator de difusão introduzido. Iremos assumir que o leitor já está familiarizado com os conceitos e resultados apresentados, referentes ao caso determinístico e portanto, abreviaremos um pouco este assunto. No entanto, não deixaremos de enfatizar as principais diferenças entre os dois casos. Pegando a deixa da equação (1.19), tem-se nesta situação

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t) , & t \in [s, T] \\ x(s) = y \in \mathbb{R}^n, \end{cases} \quad (3.1)$$

tal que $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$ e $\sigma : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^{n \times m}$ são funções mensuráveis, $\forall t \in [s, T]$. Defina $\mathcal{U}_{ad}[s, T]$ como o conjunto das 5-uplas $(\Omega, \mathcal{F}, P, W(\cdot), u(\cdot))$, tal que $(\Omega, \mathcal{F}, \{\mathcal{F}_t^{(s)}\}_{t \geq s}, P)$ é espaço de probabilidade com as condições usuais, definida no começo do *Capítulo 2*, com o movimento Browniano $W(\cdot)$ iniciado em s , ou seja, $W(s) = 0$ P a.s. e gerando a filtração $\{\mathcal{F}_t^{(s)}\}_{t \geq s}$ com $t \in [s, T]$. Além disso, $u : [s, T] \times \Omega \rightarrow U$ é $\mathcal{F}_t^{(s)}$ -

adaptado no espaço de Probabilidade considerado. Definindo a função objetivo

$$J(s, y, u(\cdot)) := E \left[\int_s^T f(t, x(t), u(t)) dt + h(x(T)) \right], \quad (3.2)$$

e assumindo que $x(\cdot)$ é solução forte de (3.1), notando abusivamente que a 5-upla acima pode ser identificada simplesmente por $u(\cdot)$. Define-se para o problema $S_{s,y}$, fixando $(s, y) \in [0, T] \times \mathbb{R}^n$, a seguinte finalidade

$$J(s, y, \bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}[s, T]} J(s, y, u(\cdot)) \quad \bar{u} \in \mathcal{U}_{ad}[s, T],$$

onde $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são funções mensuráveis, tal que $E[|h|] < +\infty$. A função valor $\mathbb{V}$, tal como no caso determinístico, será da forma

$$\mathbb{V}(s, y) := \begin{cases} J(s, y, \bar{u}(\cdot)) & \text{se } (s, y) \in [0, T) \times \mathbb{R}^n \\ h(y) & \text{se } (s, y) \in T \times \mathbb{R}^n. \end{cases} \quad (3.3)$$

Nestas condições, assuma que

($S_{s,y}1$) (U, d) é um espaço métrico completo e separável.

($S_{s,y}2$) As funções $b : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$; $\sigma : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^{n \times m}$; $f : [0, T] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são uniformemente contínuas e existe uma constante $L > 0$ tal que para $\varphi(t, x, u) := b(t, x, u)$; $\varphi(t, x, u) := \sigma(t, x, u)$; $\varphi(t, x, u) := f(t, x, u)$; $\varphi(t, x, u) := h(x)$, temos

$$\begin{cases} |\varphi(t, x_1, u) - \varphi(t, x_2, u)| \leq L|x_1 - x_2| \\ \forall t \in [0, T]; \quad \forall x_1, x_2 \in \mathbb{R}^n; \quad \forall u \in U \\ \\ |\varphi(t, 0, u)| \leq L \\ \forall (t, u) \in [0, T] \times U. \end{cases}$$

No caso estocástico, quando se pretende usar uma forma recorrente de definir a função valor para $\hat{s}$ com $\hat{s} \geq s$, não basta aplicar somente o Princípio da otimalidade de Bellman como no caso determinístico. Ao invés, temos que tomar cuidado, pois dentro do contexto de variáveis aleatórias, o controle aplicado a partir de $\hat{s}$ está naturalmente condicionado à filtração $\mathcal{F}_{\hat{s}}^{(s)}$ com $\hat{s} \in [s, T]$ gerada pela lei do movimento Browniano em $\hat{s}$, ou seja, por cada iterada na variável tempo, as 5-uplas definidas acima, terão de ser readaptadas. No entanto, o próximo lema, cuja demonstração pode ser encontrada em [11], pág 179-180, resolve essa situação, permitindo que o resultado de Bellman se mantenha válido neste caso.

Lema 3.1. Seja $(s, y) \in [0, T) \times \mathbb{R}^n$ e $(\Omega, \mathcal{F}, P, W(\cdot), u(\cdot)) \in \mathcal{U}_{ad}[s, T]$. Se $\hat{s} \in [s, T]$ e a variável aleatória $\xi(\omega) := x(\hat{s}, s, y, u(\cdot))$ é $\mathcal{F}_{\hat{s}}^{(s)}$ -mensurável, com $\omega \in \Omega$, então

$$J(\hat{s}, \xi(\omega), u(\cdot)) = E \left[\int_{\hat{s}}^T f[t, x(t, \hat{s}, \xi(\omega), u(\cdot)), u(t)] dt + h[x(T, \hat{s}, \xi(\omega), u(\cdot))] \middle| \mathcal{F}_{\hat{s}}^{(s)} \right] (\omega).$$

Teorema 3.2. (Princípio da otimalidade de Bellman estocástico) Suponha que valem as hipóteses $(S_{s,y}1)$ e $(S_{s,y}2)$. Então para qualquer $(s, y) \in [0, T] \times \mathbb{R}^n$, com $s \leq \hat{s} \leq T$, tem-se

$$\mathbb{V}(s, y) = \inf_{u \in \mathcal{U}_{ad}[s, T]} E \left[\int_s^{\hat{s}} f[t, x(t, s, y, u(\cdot)), u(t)] dt + \mathbb{V}[\hat{s}, x(\hat{s}, s, y, u(\cdot))] \right]. \quad (3.4)$$

Observação 3.3. O teorema anterior apresenta algumas variantes quando se considera o caso não determinístico. A versão aqui enunciada é Markoviana, ou seja, a informação que podemos extrair da variável X em um tempo futuro $\tilde{s}$ condicionada ao tempo presente e passado, depende apenas do tempo presente $\hat{s}$. Formalmente, tem-se para $\tilde{s} \geq \hat{s} \geq s$, $E[X(\tilde{s}) | \mathcal{F}_{\tilde{s}}^{(s)}] = E[X(\tilde{s}) | X(\hat{s})]$. A prova deste resultado segue a mesma ideia do *Capítulo 1*, com as devidas modificações necessárias no âmbito estocástico (veja Teorema 3.3 em [11], pág 180-181).

Como consequência do Princípio da programação dinâmica anterior, tem-se o seguinte resultado que será útil mais adiante.

Teorema 3.4. Suponha que as hipóteses $(S_{s,y}1)$ e $(S_{s,y}2)$ são verdadeiras. Se $(\bar{x}, \bar{u})$ é o par ótimo do problema $S_{s,y}$ então $\forall t \in [s, T]$, tem-se

$$\mathbb{V}(t, \bar{x}(t)) = E \left[\int_t^T f(r, \bar{x}(r), \bar{u}(r)) dr + h(\bar{x}(T)) \middle| \mathcal{F}_t^{(s)} \right] \quad P \text{ a.s.} \quad (3.5)$$

para $(s, y) \in [0, T) \times \mathbb{R}^n$ fixado.

Demonstração. Fixando, $(s, y) \in [0, T) \times \mathbb{R}^n$ e dado o par $(\bar{x}, \bar{u})$, sabemos que este é um par que realiza o ínfimo na função J , isto é

$$J(s, y, \bar{u}) = \inf_{u \in \mathcal{U}_{ad}[s, T]} J(s, y, u) = \mathbb{V}(s, y) \quad , \quad \bar{u} \in \mathcal{U}_{ad}[s, T].$$

Dai que

$$\mathbb{V}(s, y) = J(s, y, \bar{u}) = E \left[\int_s^T f(t, \bar{x}(t), \bar{u}(t)) dt + h(\bar{x}(T)) \right]$$

$$\begin{aligned}
 &= E \left[\int_s^t f(t, \bar{x}(t), \bar{u}(t)) dt \right] + E \left[\int_t^T f(t, \bar{x}(t), \bar{u}(t)) dt + h(\bar{x}(T)) \middle| \mathcal{F}_t^{(s)} \right] \\
 &= E \left[\int_s^t f(t, \bar{x}(t), \bar{u}(t)) dt \right] + E[J(t, \bar{x}(t), \bar{u})] \\
 &\geq E \left[\int_s^t f(t, \bar{x}(t), \bar{u}(t)) dt \right] + E[\mathbb{V}(t, \bar{x}(t))] \geq \mathbb{V}(s, y),
 \end{aligned}$$

sendo que a última desigualdade resulta de (3.4). Logo é imediato que

$$E[J(t, \bar{x}(t), \bar{u})] = E[\mathbb{V}(t, \bar{x}(t))].$$

Como $J(t, \bar{x}(t), \bar{u}) \geq \mathbb{V}(t, \bar{x}(t))$ P a.s., vem que

$$\mathbb{V}(t, \bar{x}(t)) = J(t, \bar{x}(t), \bar{u}) = E \left[\int_t^T f(r, \bar{x}(r), \bar{u}(r)) dr + h(\bar{x}(T)) \middle| \mathcal{F}_t^{(s)} \right].$$

□

Similarmente ao analisado no *Capítulo 1*, apresentam-se em seguida numa forma mais resumida, os resultados de existência e unicidade em relação ao caso estocástico.

Teorema 3.5. Suponha que valem as hipóteses $(S_{s,y}1)$ e $(S_{s,y}2)$ e a função valor $\mathbb{V}$ definida em (3.3) é de classe $C^{1,2}([0, T] \times \mathbb{R}^n)$. Então $\mathbb{V}$ é solução da equação diferencial

$$\begin{cases} -\frac{\partial v}{\partial t}(t, x) + \sup_{u \in U} G\left(t, x, u, -\frac{\partial v}{\partial x}(t, x), -\frac{\partial^2 v}{\partial x^2}(t, x)\right) = 0 & (t, x) \in [0, T] \times \mathbb{R}^n \\ v(T, x) = h(x) & (t, x) \in T \times \mathbb{R}^n, \end{cases} \quad (3.6)$$

onde G é o Hamiltoniano generalizado definido em (2.47) no *Capítulo 2*.

Demonstração. A prova é sensivelmente similar ao caso determinístico, com as devidas adaptações na esperança. Efetivamente, fixe $(s, y) \in [0, T] \times \mathbb{R}^n$ e $u := u(\cdot) \in U$. Considere o estado $x(t, s, y, u)$ referente ao controle u fixado. Dado $\hat{s} \in [s, T]$, pelo Princípio da otimalidade de Bellman estocástico e dividindo ambos os membros por $\hat{s} - s$, tem-se para $u \in \mathcal{U}_{ad}[s, T]$,

$$-\frac{E\{\mathbb{V}[\hat{s}, x(\hat{s}, s, y, u)] - \mathbb{V}(s, y)\}}{\hat{s} - s} \leq \frac{1}{\hat{s} - s} E \left\{ \int_s^{\hat{s}} f[t, x(t, s, y, u), u(t)] dt \right\}. \quad (3.7)$$

Aplicando a fórmula de Itô em $\mathbb{V}[\hat{s}, x(\hat{s}, s, y, u)] - \mathbb{V}(s, y)$ e aliviando notação, fica

$$\begin{aligned} \mathbb{V}(\hat{s}, x(\hat{s})) - \mathbb{V}(s, x(s)) &= \int_s^{\hat{s}} \left\{ \frac{\partial \mathbb{V}}{\partial t}(t, x(t)) + \left\langle \frac{\partial \mathbb{V}}{\partial x}(t, x(t)), b(t, x(t), u) \right\rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr} \left[\sigma^\top(t, x(t), u) \frac{\partial^2 \mathbb{V}}{\partial x^2}(t, x(t)) \sigma(t, x(t), u) \right] \right\} dt \\ &\quad + \int_s^{\hat{s}} \left\langle \frac{\partial \mathbb{V}}{\partial x}(t, x(t)), \sigma(t, x(t), u) dW(t) \right\rangle, \end{aligned} \quad (3.8)$$

donde, aplicando a esperança em (3.8), tem-se por (3.7)

$$\begin{aligned} & - \frac{1}{\hat{s} - s} E \left\{ \int_s^{\hat{s}} \left\{ \frac{\partial \mathbb{V}}{\partial t}(t, x(t)) + \left\langle \frac{\partial \mathbb{V}}{\partial x}(t, x(t)), b(t, x(t), u) \right\rangle \right. \right. \\ & \quad \left. \left. + \frac{1}{2} \text{tr} \left[\sigma^\top(t, x(t), u) \frac{\partial^2 \mathbb{V}}{\partial x^2}(t, x(t)) \sigma(t, x(t), u) \right] \right\} dt \right. \\ & \quad \left. + \int_s^{\hat{s}} \left\langle \frac{\partial \mathbb{V}}{\partial x}(t, x(t)), \sigma(t, x(t), u) dW(t) \right\rangle \right\} \\ & - \frac{1}{\hat{s} - s} E \left\{ \int_s^{\hat{s}} f[t, x(t, s, y, u), u(t)] dt \right\} \\ & \leq 0. \end{aligned}$$

Ou seja, como a esperança anula a integral estocástica, resulta que

$$\frac{1}{\hat{s} - s} E \int_s^{\hat{s}} \left\{ - \frac{\partial \mathbb{V}}{\partial t}(t, x(t)) + G \left[t, x(t), u, - \frac{\partial \mathbb{V}}{\partial x}(t, x(t)), - \frac{\partial^2 \mathbb{V}}{\partial x^2}(t, x(t)) \right] \right\} dt \leq 0. \quad (3.9)$$

Como por hipótese $\mathbb{V}$ é de classe $C^{1,2}([0, T] \times \mathbb{R}^n)$, tem-se, notando (3.9) com $\hat{s} \rightarrow s$,

$$- \frac{\partial \mathbb{V}}{\partial t}(s, y) + G \left(s, y, u, - \frac{\partial \mathbb{V}}{\partial x}(s, y), - \frac{\partial^2 \mathbb{V}}{\partial x^2}(s, y) \right) \leq 0,$$

ou seja

$$- \frac{\partial \mathbb{V}}{\partial t}(s, y) + \sup_{u \in U} G \left(s, y, u, - \frac{\partial \mathbb{V}}{\partial x}(s, y), - \frac{\partial^2 \mathbb{V}}{\partial x^2}(s, y) \right) \leq 0. \quad (3.10)$$

Por outro lado, dado $\varepsilon > 0$, $\hat{s} \in [s, T]$ e fazendo $\hat{s} - s$ tão pequeno quanto se queira, existe mais uma vez pelo Princípio da otimalidade de Bellman estocástico, $u := u_{\varepsilon, \hat{s}}(\cdot) \in U$ tal

que

$$\mathbb{V}(s, y) + \varepsilon(\hat{s} - s) \geq \int_s^{\hat{s}} f(t, x(t), u(t)) dt + \mathbb{V}(\hat{s}, x(\hat{s})),$$

onde $x(t) := x_{\varepsilon, \hat{s}}(t, s, y, u_{\varepsilon, \hat{s}}(\cdot))$ e $V(\hat{s}, x(\hat{s})) := V[\hat{s}, x(\hat{s}, s, y, u(\cdot))]$.

Assim,

$$-\frac{E\{\mathbb{V}(\hat{s}, x(\hat{s})) - \mathbb{V}(s, y)\}}{\hat{s} - s} \geq -\varepsilon + \frac{1}{\hat{s} - s} E\left\{ \int_s^{\hat{s}} f[t, x(t, s, y, u), u(t)] dt \right\}.$$

Da mesma forma anterior, pode-se escrever

$$\frac{1}{\hat{s} - s} E \int_s^{\hat{s}} \left\{ -\frac{\partial \mathbb{V}}{\partial t}(t, x(t)) + G\left[t, x(t), u, -\frac{\partial \mathbb{V}}{\partial x}(t, x(t)), -\frac{\partial^2 \mathbb{V}}{\partial x^2}(t, x(t))\right] \right\} dt \geq -\varepsilon. \quad (3.11)$$

Pela hipótese $(S_{s,y}2)$, garante-se a continuidade da função G , pois esta depende de funções uniformemente contínuas $(b, \sigma \text{ e } f)$. Portanto, pelo teorema da diferenciação de Lebesgue, $\hat{s}$ converge para s em todo o ponto e assim (3.11) dá origem a

$$-\frac{\partial \mathbb{V}}{\partial t}(s, y) + \sup_{u \in U} G\left(s, y, u, -\frac{\partial \mathbb{V}}{\partial x}(s, y), -\frac{\partial^2 \mathbb{V}}{\partial x^2}(s, y)\right) \geq -\varepsilon. \quad (3.12)$$

Tomando ε tão pequeno quanto se queira, juntando (3.10) e (3.12), o resultado segue. $\square$

Observação 3.6. Embora não seja uma ideia central no desenrolar deste trabalho, é importante referir uma relação interessante entre o Princípio de Pontryagin estocástico e a programação dinâmica estocástica (veja Teorema 6.4.7 em [8], pág. 151-152). Isto é, quando o problema de otimização S é substituído pelo problema $S_{s,y}$ e assumindo a BSDE (2.43), existe uma identidade entre a solução $(\bar{Y}, \bar{Z})$ dessa BSDE e as derivadas de primeira e segunda ordem em x da função valor $\mathbb{V}$. Neste caso, essa relação é dada por

$$(\bar{Y}, \bar{Z}) = \left(-\frac{\partial \mathbb{V}}{\partial x}(t, \bar{x}), -\frac{\partial^2 \mathbb{V}}{\partial x^2}(t, \bar{x}) \sigma(t, \bar{x}, \bar{u}) \right). \quad (3.13)$$

Facilmente observa-se que, nestas condições, $\mathcal{H}$ em (2.46) é exatamente o Hamiltoniano generalizado G , que é maximizado no par ótimo do problema $S_{s,y}$. Este fato, de certa forma, justifica a utilização da função Hamiltoniana G nas equações HJB (3.6).

Igualmente como no caso determinístico, podemos introduzir o conceito de solução de viscosidade para o caso estocástico.

Definição 3.7. Uma função $v \in C([0, T] \times \mathbb{R}^n)$ é uma subsolução de viscosidade de (3.6)

se, e somente se,

$$v(T, x) \leq h(x) \quad \forall (t, x) \in T \times \mathbb{R}^n, \quad (3.14)$$

e para cada $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$, ao qual chamamos de função teste, sempre que $(t, x) \in [0, T) \times \mathbb{R}^n$ é um ponto de máximo local de $v - \varphi$, tem-se

$$-\frac{\partial \varphi}{\partial t}(t, x) + \sup_{u \in U} G\left(t, x, u, -\frac{\partial \varphi}{\partial x}(t, x), -\frac{\partial^2 \varphi}{\partial x^2}(t, x)\right) \leq 0 \quad (t, x) \in [0, T) \times \mathbb{R}^n. \quad (3.15)$$

Por outro lado, uma função $v \in C([0, T] \times \mathbb{R}^n)$ é uma supersolução de viscosidade de (3.6) se, e somente se, em (3.14) e (3.15), a desigualdade " $\leq$ " for substituído por " $\geq$ " e o termo "máximo local" for substituído por "mínimo local". Além disso, quando $v \in C([0, T] \times \mathbb{R}^n)$ é subsolução e supersolução de viscosidade, então é chamada solução de viscosidade de (3.6).

Assim, em sintonia com a programação dinâmica determinística, temos o seguinte resultado.

Teorema 3.8. Suponha que valem as hipóteses $(S_{s,y}1)$ e $(S_{s,y}2)$ e a função valor $\mathbb{V}$ definida em (3.3) é de classe $C([0, T] \times \mathbb{R}^n)$. Então para qualquer $(s, y) \in [0, T] \times \mathbb{R}^n$, a função valor $\mathbb{V}(s, y)$ satisfaz

$$|\mathbb{V}(s_1, y_1) - \mathbb{V}(s_2, y_2)| \leq K\{|y_1 - y_2| + (1 + |y_1| \vee |y_2|)|s_1 - s_2|^{\frac{1}{2}}\}, \quad (3.16)$$

em que $(s_1, y_1) \in [0, T] \times \mathbb{R}^n$ e $(s_2, y_2) \in [0, T] \times \mathbb{R}^n$, com K constante positiva. Além disso, $\mathbb{V}$ é a única solução de viscosidade de (3.6).

Esboço da demonstração. Suponha que valem as estimativas

$$\left\{ \begin{array}{l} |x(t, s, y, u(\cdot))| \leq K(1 + |y|) \\ \forall t \in [s, T]; \quad \forall (s, y) \in [0, T] \times \mathbb{R}^n; \quad u \in \mathcal{V}[s, T] \\ \\ |x(t, s_1, y_1, u(\cdot)) - x(t, s_2, y_2, u(\cdot))| \leq K\{|y_1 - y_2| + (1 + |y_1| \vee |y_2|)|s_1 - s_2|^{\frac{1}{2}}\} \\ \forall t \in [s_1 \vee s_2, T]; \quad \forall (s_1, s_2) \in [0, T]; \quad \forall (y_1, y_2) \in \mathbb{R}^n; \quad u \in \mathcal{V}[s_1 \vee s_2, T]. \end{array} \right.$$

Nesse sentido, notando (2.7), (2.8) e (2.9) no *Capítulo 2*, resulta a desigualdade (3.16). Por outro lado, a existência da solução de viscosidade $\mathbb{V}$ deduz-se da prova da existência da solução de viscosidade V no Teorema 1.18. De fato, sendo $(s, y) \in [0, T) \times \mathbb{R}^n$ um ponto de máximo local de $\mathbb{V} - \varphi$, tal que $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$, tem-se

$$-\frac{\partial \varphi}{\partial t}(s, y) + \sup_{u \in U} G\left(s, y, u, -\frac{\partial \varphi}{\partial x}(s, y), -\frac{\partial^2 \varphi}{\partial x^2}(s, y)\right) \leq 0,$$

e sendo $(s, y) \in [0, T] \times \mathbb{R}^n$ um ponto de mínimo local de $\mathbb{V} - \varphi$, tal que $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$ com $\varepsilon \geq 0$, segue que

$$-\frac{\partial \varphi}{\partial t}(s, y) + \sup_{u \in U} G\left(s, y, u, -\frac{\partial \varphi}{\partial x}(s, y), -\frac{\partial^2 \varphi}{\partial x^2}(s, y)\right) \geq -\varepsilon.$$

□

Observação 3.9. A prova da unicidade da solução de viscosidade anterior é bem delicada (veja Teorema 6.1 em [11], pág. 198-212). Iremos assumir, de agora em diante, a unicidade desse tipo de soluções.

Dentro do espírito da programação dinâmica analisada no *Capítulo 1*, daremos a noção de subdiferencial e superdiferencial de segunda ordem. Nesta dissertação e em visto ao que se seguirá, considera-se apenas o caso parabólico.

Definição 3.10. Para qualquer função $v : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ e $(t, x) \in [0, T] \times \mathbb{R}^n$, defina

$$\mathcal{DP}_{t,x}^{1,2,-} v(t, x) := \left\{ (l, p, \tilde{p}) \in \mathbb{R} \times \mathbb{R}^n \times S^n \mid \lim_{\substack{s \rightarrow t \\ y \rightarrow x}} \frac{1}{|s - t| + |y - x|^2} \left[v(s, y) - v(t, x) - l(s - t) - \langle p, y - x \rangle - \frac{1}{2}(y - x)^\top \tilde{p}(y - x) \right] \geq 0 \right\}, \quad (3.17)$$

a subdiferencial parabólica de segunda ordem e

$$\mathcal{DP}_{t,x}^{1,2,+} v(t, x) := \left\{ (l, p, \tilde{p}) \in \mathbb{R} \times \mathbb{R}^n \times S^n \mid \overline{\lim}_{\substack{s \rightarrow t \\ y \rightarrow x}} \frac{1}{|s - t| + |y - x|^2} \left[v(s, y) - v(t, x) - l(s - t) - \langle p, y - x \rangle - \frac{1}{2}(y - x)^\top \tilde{p}(y - x) \right] \leq 0 \right\}, \quad (3.18)$$

a superdiferencial parabólica de segunda ordem. Também, como no *Capítulo 1*, se define $\mathcal{DP}_{t+,x}^{1,2,-} v(t, x)$ e $\mathcal{DP}_{t+,x}^{1,2,+} v(t, x)$ a subdiferencial e a superdiferencial à direita na variável t .

Em relação às definições em (1.40), embora as propriedades sejam praticamente idênticas a (3.17) e (3.18), existem duas diferenças essenciais. A primeira, é que estes conjuntos podem não ser fechados (veja Exemplo 5.3 em [11], pág. 192-193). A segunda diferença interessante, é fato dos conjuntos definidos anteriormente, não serem necessariamente pontos isolados. Para verificar isso, suponha para $(t_0, x_0) \in [0, T] \times \mathbb{R}^n$, a existência das derivadas parciais de primeira ordem em t e de primeira e segunda ordem em x da função

$v \in C([0, T] \times \mathbb{R}^n)$. Nestas condições, temos $\mathcal{DP}_{t,x}^{1,2,-}v(t_0, x_0) \cap \mathcal{DP}_{t,x}^{1,2,+}v(t_0, x_0) \neq \emptyset$, onde

$$\begin{cases} \mathcal{DP}_{t,x}^{1,2,-}v(t_0, x_0) = \left\{ \frac{\partial v}{\partial t}(t_0, x_0); \frac{\partial v}{\partial x}(t_0, x_0); \left(-\infty, \frac{\partial^2 v}{\partial x^2}(t_0, x_0) \right] \right\} \\ \mathcal{DP}_{t,x}^{1,2,+}v(t_0, x_0) = \left\{ \frac{\partial v}{\partial t}(t_0, x_0); \frac{\partial v}{\partial x}(t_0, x_0); \left[\frac{\partial^2 v}{\partial x^2}(t_0, x_0), +\infty \right) \right\}. \end{cases} \quad (3.19)$$

Sabendo que a derivada parcial de segunda ordem em x da função v é um elemento de S^n e definindo $(-\infty, \hat{p}] = \{\tilde{p} \in S^n \text{ tal que } \tilde{p} \leq \hat{p}\}$ e $[\hat{p}, +\infty) = \{\tilde{p} \in S^n \text{ tal que } \tilde{p} \geq \hat{p}\}$, resulta que os conjuntos em (3.19) não são pontos isolados.

Teorema 3.11. Suponha que valem as hipóteses $(S_{s,y}1)$ e $(S_{s,y}2)$. Então a função valor $\mathbb{V} \in C([0, T] \times \mathbb{R}^n)$ é a única que satisfaz as seguintes condições para $(t, x) \in [0, T] \times \mathbb{R}^n$, ou seja

$$\begin{cases} -l + \sup_{u \in U} G\left(t, x, u, -p, -\tilde{p}\right) \leq 0 & \forall (l, p, \tilde{p}) \in \mathcal{DP}_{t,x}^{1,2,+}\mathbb{V}(t, x) \\ -l + \sup_{u \in U} G\left(t, x, u, -p, -\tilde{p}\right) \geq 0 & \forall (l, p, \tilde{p}) \in \mathcal{DP}_{t,x}^{1,2,-}\mathbb{V}(t, x) \\ \mathbb{V}(T, x) = h(x), & (t, x) \in T \times \mathbb{R}^n. \end{cases} \quad (3.20)$$

Demonstração. Em analogia com o Teorema 1.20 do *Capítulo 1*, a essência da prova é a construção da função teste φ em que, sendo $(t_0, x_0) \in (0, T] \times \mathbb{R}^n$, resulta que $(l, p, \tilde{p}) \in \mathcal{DP}_{t,x}^{1,2,+}v(t_0, x_0)$ se, e somente se existe uma função $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$, tal que $v - \varphi$ atinge o máximo em sentido estrito em (t_0, x_0) e

$$\left(\varphi(t_0, x_0), \frac{\partial \varphi}{\partial t}(t_0, x_0), \frac{\partial \varphi}{\partial x}(t_0, x_0), \frac{\partial^2 \varphi}{\partial x^2}(t_0, x_0) \right) = (v(t_0, x_0), l, p, \tilde{p}). \quad (3.21)$$

Similarmente, $(l, p, \tilde{p}) \in \mathcal{DP}_{t,x}^{1,2,-}v(t_0, x_0)$ se, e somente se existe uma função $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$, tal que $v - \varphi$ atinge o mínimo em sentido estrito em (t_0, x_0) e tem-se (3.21). Para detalhes dessa construção, veja o Lema 4.1 em [3], pág. 211-214. O resto da prova, segue a mesma diretriz do Teorema 1.20, substituindo H pelo Hamiltoniano generalizado G em (1.45), donde assumindo a unicidade da solução de viscosidade na equação de HJB (3.6), resulta (3.20). $\square$

Observação 3.12. Constata-se sem grande dificuldade que o hamiltoniano G , mantém um comportamento monótono não decrescente em relação a $\tilde{p} \in S^n$, ou seja, dados $\tilde{p}_1, \tilde{p}_2 \in S^n$,

tem-se por (2.47),

$$\tilde{p}_1 \leq \tilde{p}_2 \Rightarrow G(t, x, u, p, \tilde{p}_1) \leq G(t, x, u, p, \tilde{p}_2).$$

Por outro lado, pelo teorema anterior e pela definição de solução de viscosidade, $v \in C([0, T] \times \mathbb{R}^n)$ é subsolução de viscosidade se, e somente se

$$\begin{cases} -l + \sup_{u \in U} G(t, x, u, -p, -\tilde{p}) \leq 0 & \forall (l, p, \tilde{p}) \in \mathcal{DP}_{t,x}^{1,2,+} v(t, x) \\ v(T, x) \leq h(x) & (t, x) \in T \times \mathbb{R}^n, \end{cases} \quad (3.22)$$

ou é supersolução de viscosidade se, e somente se

$$\begin{cases} -l + \sup_{u \in U} G(t, x, u, -p, -\tilde{p}) \geq 0 & \forall (l, p, \tilde{p}) \in \mathcal{DP}_{t,x}^{1,2,-} v(t, x) \\ v(T, x) \geq h(x) & (t, x) \in T \times \mathbb{R}^n. \end{cases} \quad (3.23)$$

Desta forma e supondo que v é de classe $C^{1,2}([0, T] \times \mathbb{R}^n)$, é imediato que a definição de solução clássica e de viscosidade são equivalentes.

3.2 Fórmulas de Feynman-Kac

Nesta seção iremos apresentar resultados, que permitem relacionar SDEs e BSDEs com as PDEs de segunda ordem. Algumas classes dessas PDEs podem ter soluções de natureza probabilística, que são determinadas pelas fórmulas de Feynman-Kac, em homenagem a Richard Feynman e Mark Kac. Neste trabalho, apenas consideramos o caso parabólico. De fato, será provado que a solução de uma PDE parabólica pode ser determinada pela esperança e para isso, as equações estocásticas têm um papel relevante nessa abordagem. Essa relação faz todo o sentido, pois a maioria dos fenômenos que se apresentam na natureza, são por si só, sistemas de difusão e portanto de certa forma, com componente aleatória ao longo do tempo.

3.2.1 Solução de uma PDE parabólica linear

Como motivação, considere a seguinte equação diferencial parcial terminal:

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + A \frac{\partial^2 v}{\partial x^2}(t, x) + B \frac{\partial v}{\partial x}(t, x) = 0 & (t, x) \in [0, T) \times \mathbb{R}^n \\ v(T, x) = h(x) & (t, x) \in T \times \mathbb{R}^n, \end{cases} \quad (3.24)$$

em que definimos,

$$A \frac{\partial^2 v}{\partial x^2}(t, x) := \sum_{i,j=1}^n a_{ij}(t, x) \frac{\partial^2 v}{\partial x_i \partial x_j}(t, x) \quad \text{e} \quad B \frac{\partial v}{\partial x}(t, x) := \sum_{i=1}^n b_i(t, x) \frac{\partial v}{\partial x_i}(t, x),$$

sendo $a_{ij} : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$; $b_i : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$. Será que a solução desta PDE, pode ser expressa em termos estocásticos? Para responder à questão, considere uma versão mais geral da equação (3.24), ou seja

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + A \frac{\partial^2 v}{\partial x^2}(t, x) + B \frac{\partial v}{\partial x}(t, x) + g(t, x) = 0 & (t, x) \in [0, T) \times \mathbb{R}^n \\ v(T, x) = h(x) & (t, x) \in T \times \mathbb{R}^n, \end{cases} \quad (3.25)$$

com $g : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$. Considere as funções $b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ e $\sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ da Definição 2.6, tal que $\forall (t, x) \in [0, T] \times \mathbb{R}^n$, tenhamos as matrizes $b = [b_{i1}]$ e $\sigma \sigma^\top = [2a_{ij}]$, com $i, j \in \{1, \dots, n\}$. Assim sendo, a PDE (3.25) poderá ser rescrita na forma

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + \frac{1}{2} \text{tr} \left[\sigma(t, x) \sigma^\top(t, x) \frac{\partial^2 v}{\partial x^2}(t, x) \right] + \left\langle b(t, x), \frac{\partial v}{\partial x}(t, x) \right\rangle + g(t, x) = 0 \\ v(T, x) = h(x) & (t, x) \in T \times \mathbb{R}^n. \end{cases} \quad (3.26)$$

Nestas condições, assuma que,

(HF) As funções $b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$; $\sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$; $g : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são uniformemente contínuas e existe uma constante $L > 0$ tal que para

$\varphi(t, x) := b(t, x)$; $\varphi(t, x) := \sigma(t, x)$; $\varphi(t, x) := g(t, x)$; $\varphi(t, x) := h(x)$, resulta

$$\left\{ \begin{array}{l} |\varphi(t, x_1) - \varphi(t, x_2)| \leq L|x_1 - x_2| \\ \forall t \in [0, T]; \quad \forall x_1, x_2 \in \mathbb{R}^n \\ \\ |\varphi(t, 0)| \leq L \\ \forall t \in [0, T]. \end{array} \right.$$

Teorema 3.13. Se vale a hipótese (HF) então v é a única solução de viscosidade de (3.25), com a representação

$$v(t, x) = E \left[\int_t^T g(s, X(s, t, x)) ds + h(X(T)) \right] \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.27)$$

onde $X(\cdot, t, x)$ é a única solução forte da SDE,

$$\left\{ \begin{array}{l} dX(s) = b(s, X(s))ds + \sigma(s, X(s))dW(s) \\ X(t) = x \end{array} \right. \quad s \in [t, T], \quad (3.28)$$

sendo $(t, x) \in [0, T] \times \mathbb{R}^n$ e $W(s)$, o movimento Browniano usual de dimensão m com $s \in [t, T]$ e $W(t) = 0$. Além disso, se (3.25) admite uma única solução clássica então esta é representada por (3.27).

Demonstração. Para simplificar notação, faça $X(s, t, x) := X(s)$ com $s \in [t, T]$. Defina $U = \{u_0\}$ como sendo um ponto isolado. Assim, a equação dos estados (3.28) pode ser associada ao espaço do controle U representado apenas por u_0 , donde, é válido abordar este problema, como um problema de controle estocástico. Observar que assim, (HF) será exatamente $(S_{s,y}2)$, com as devidas adaptações na notação. Assim sendo, notando (3.5) com $f(s, y, u_0) = g(s, y)$, $\forall (s, y) \in [t, T] \times \mathbb{R}^n$, resulta que

$$\begin{aligned} v(t, X(t)) &= \inf_{\{u_0\}} E \left[\int_t^T f(s, X(s), u_0) ds + h(X(T)) \right] \\ &= E \left[\int_t^T f(s, X(s), u_0) ds + h(X(T)) \right], \quad (t, x) \in [0, T] \times \mathbb{R}^n. \end{aligned} \quad (3.29)$$

Assim, considerando $(X(t), u_0)$ o para ótimo do problema de controle associado, tem-se (3.27). Note que a fórmula de Itô também permite obter a representação (3.27). Com efeito, assumindo a suavidade da função v , isto é, de classe $C^{1,2}([0, T] \times \mathbb{R}^n)$ e sabendo

que o processo estocástico $X(\cdot, t, x)$ é $\mathcal{F}_s$ - progressivamente mensurável e solução única de (3.28), tem-se

$$\begin{aligned} v(T, X(T)) - v(t, X(t)) &= \int_t^T \left\{ \frac{\partial v}{\partial s}(s, X(s)) + \left\langle \frac{\partial v}{\partial x}(s, X(s)), b(s, X(s)) \right\rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr} \left[\sigma(s, X(s)) \sigma^\top(s, X(s)) \frac{\partial^2 v}{\partial x^2}(s, X(s)) \right] \right\} ds \\ &\quad + \int_t^T \left\langle \frac{\partial v}{\partial x}(s, X(s)), \sigma(s, X(s)) dW(s) \right\rangle. \end{aligned}$$

Por (3.26) e (3.28), segue que

$$\begin{aligned} h(X(T)) - v(t, x) &= - \int_t^T g(s, X(s)) ds \\ &\quad + \int_t^T \left\langle \frac{\partial v}{\partial x}(s, X(s)), \sigma(s, X(s)) dW(s) \right\rangle, \end{aligned} \quad (3.30)$$

e aplicando a esperança nos dois membros de (3.30), fica para cada $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$\begin{aligned} E[h(X(T))] - v(t, x) &= -E \left[\int_t^T g(s, X(s)) ds \right] \\ &\quad + E \left[\int_t^T \left\langle \frac{\partial v}{\partial x}(s, X(s)), \sigma(s, X(s)) dW(s) \right\rangle \right]. \end{aligned}$$

Isto mostra (3.27). Como (3.26) corresponde precisamente às equações HJB (3.6) para $U = \{u_0\}$ e pelo que foi analisado anteriormente neste capítulo, existe uma única solução clássica e portanto de viscosidade para (3.26). Logo, (3.25) admite uma única solução de viscosidade ou clássica, que é dada por (3.27). $\square$

Observação 3.14. O teorema anterior permite responder à questão do início desta seção, bastando para isso, fazer $g = 0$ em (3.27). Assim, (3.24) tem uma solução representada estocasticamente por

$$v(t, x) = E[h(X(T))] , \quad (t, x) \in [0, T] \times \mathbb{R}^n. \quad (3.31)$$

Seguindo o mesmo raciocínio em relação (3.25), considere agora a seguinte PDE pa-

rabólica:

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + A \frac{\partial^2 v}{\partial x^2}(t, x) + B \frac{\partial v}{\partial x}(t, x) + c(t, x)v(t, x) + g(t, x) = 0 \\ v(T, x) = h(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.32)$$

com $c : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$, tal que $c(t, x)$ é uniformemente contínua e limitada no seu domínio.

Teorema 3.15. Se vale a hipótese (HF) e c é uma função com as condições acima, então v é a única solução de viscosidade de (3.32). Além disso, essa solução pode ser representada, para cada $(t, x) \in [0, T] \times \mathbb{R}^n$, por

$$v(t, x) = E \left\{ \int_t^T g(s, X(s)) \exp \left[\int_t^s c(r, X(r)) dr \right] ds + h(X(T)) \exp \left[\int_t^T c(r, X(r)) dr \right] \right\}, \quad (3.33)$$

onde $X(\cdot) := X(\cdot, t, x)$ é a única solução forte (3.28).

Demonstração. Podemos escrever (3.32) como

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + \frac{1}{2} \text{tr} \left[\sigma(t, x) \sigma^\top(t, x) \frac{\partial^2 v}{\partial x^2}(t, x) \right] + \left\langle b(t, x), \frac{\partial v}{\partial x}(t, x) \right\rangle + F(t, x) = 0 \\ v(T, x) = h(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.34)$$

em que $F(t, x) := c(t, x)v(t, x) + g(t, x)$ tem forma linear. Lembrando que o Hamiltoniano G é dado por

$$\begin{aligned} G(t, x, u_0, p, \tilde{p}) &= \left\langle b(t, x, u_0), p \right\rangle + \frac{1}{2} \text{tr} \left[\sigma(t, x, u_0) \sigma^\top(t, x, u_0) \tilde{p} \right] - f(t, x, u_0) \\ 0 \leq t \leq T; \quad x &\in \mathbb{R}^n; \quad U = \{u_0\}; \quad p \in \mathbb{R}^n; \quad \tilde{p} \in S^n, \end{aligned}$$

com $f(t, x, u_0) = g(t, x) \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n$, resulta que as equações HJB associadas ao controle $\{u_0\}$ para $v \in C^{1,2}([0, T] \times \mathbb{R}^n)$, têm a forma

$$\begin{cases} -\frac{\partial v}{\partial t}(t, x) + \sup_{\{u_0\}} G \left(t, x, u_0, -\frac{\partial v}{\partial x}(t, x), -\frac{\partial^2 v}{\partial x^2}(t, x) \right) - c(t, x)v(t, x) = 0 \\ v(T, x) = h(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n,$$

que nada mais são que as equações (3.34). Assim, será de esperar que para $c(t, x)$ não necessariamente nula, a lógica se mantenha idêntica ao que foi analisado na programação dinâmica estocástica, anteriormente neste capítulo, ou seja, que existe uma única solução de

viscosidade para (3.32). No entanto, essa ideia só será devidamente justificada na próxima seção. Por agora, considere a função $R : [t, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$, que é dada por

$$R(s, y) := v(s, y) \exp \left(\int_t^s c(r, z) dr \right), \quad (3.35)$$

com $y := x(s)$ e $z := x(r)$. Observe que para cada $t \in [0, T]$, $R(s, y)$ é de classe $C^{1,2}([t, T] \times \mathbb{R}^n)$ pois c é limitada e uniformemente contínua. Considerando o processo $X(\cdot)$ como a solução única de (3.28), pode-se escrever

$$R(s, X(s)) = v(s, X(s)) \exp \left[\int_t^s c(r, X(r)) dr \right].$$

Agora, aplicando a fórmula de Itô em $R(s, X(s))$, fica

$$\begin{aligned} d[R(s, X(s))] &= \frac{\partial R}{\partial s}(s, X(s))d(s) + \frac{\partial R}{\partial x}(s, X(s))d(X(s)) + \frac{1}{2} \frac{\partial^2 R}{\partial x^2}(s, X(s)) [d(X(s))]^2 \\ &= \left\{ v(s, X(s))c(s, X(s)) \exp \left[\int_t^s c(r, X(r)) dr \right] \right. \\ &\quad \left. + \frac{\partial v}{\partial s}(s, X(s)) \exp \left[\int_t^s c(r, X(r)) dr \right] \right\} d(s) \\ &\quad + \frac{\partial v}{\partial x}(s, X(s)) \exp \left[\int_t^s c(r, X(r)) dr \right] d(X(s)) \\ &\quad + \frac{1}{2} \frac{\partial^2 v}{\partial x^2}(s, X(s)) \exp \left[\int_t^s c(r, X(r)) dr \right] [d(X(s))]^2. \end{aligned}$$

Portanto, tendo em conta a definição da função R em (3.35), tem-se que

$$\begin{aligned} R(T, X(T)) - R(t, X(t)) &= v(T, X(T)) \exp \left[\int_t^T c(r, X(r)) dr \right] - v(t, X(t)) \\ &= \int_t^T \exp \left[\int_t^s c(r, X(r)) dr \right] \left\{ \frac{\partial v}{\partial s}(s, X(s)) \right. \\ &\quad \left. + v(s, X(s))c(s, X(s)) + \left\langle \frac{\partial v}{\partial x}(s, X(s)), b(s, X(s)) \right\rangle \right. \\ &\quad \left. + \frac{1}{2} tr \left[\sigma(s, X(s)) \sigma^\top(s, X(s)) \frac{\partial^2 v}{\partial x^2}(s, X(s)) \right] \right\} ds \end{aligned}$$

$$+ \int_t^T \exp \left[\int_t^s c(r, X(r)) dr \right] \left\langle \frac{\partial v}{\partial x}(s, X(s)), \sigma(s, X(s)) dW(s) \right\rangle,$$

donde por (3.28) e (3.34), vem que

$$\begin{aligned} h(X(T)) \exp \left[\int_t^T c(r, X(r)) dr \right] - v(t, x) &= - \int_t^T g(s, X(s)) \exp \left[\int_t^s c(r, X(r)) dr \right] ds \\ &+ \int_t^T \exp \left[\int_t^s c(r, X(r)) dr \right] \left\langle \frac{\partial v}{\partial x}(s, X(s)), \sigma(s, X(s)) dW(s) \right\rangle. \end{aligned} \quad (3.36)$$

Assim, aplicando a esperança nos dois membros da igualdade (3.36), resulta para cada $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$\begin{aligned} E \left\{ h(X(T)) \exp \left[\int_t^T c(r, X(r)) dr \right] \right\} - v(t, x) \\ = -E \left\{ \int_t^T \exp \left[\int_t^s c(r, X(r)) dr \right] g(s, X(s)) ds \right\}. \end{aligned}$$

Isto prova (3.33). □

A relação (3.33) tem o nome de fórmula de Feynman-Kac.

3.2.2 Solução de uma PDE parabólica não linear

Na seção anterior, ficou claro que soluções de PDEs parabólicas lineares, têm uma representação explicitada em termos de esperança matemática. Nesta última seção será provado que PDEs não necessariamente lineares, embora não tenham soluções explicitadas, têm uma única solução de viscosidade. Nesse sentido, ficará também justificada a unicidade deste tipo de soluções para a equação (3.32). Assim, para terminar a dissertação, será considerado um caso mais geral deste tipo equações diferenciais, ou seja

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + A \frac{\partial^2 v}{\partial x^2}(t, x) + B \frac{\partial v}{\partial x}(t, x) + F\left(t, x, v(t, x), r(t, x) \frac{\partial v}{\partial x}(t, x)\right) = 0 \\ v(T, x) = h(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.37)$$

em que $F : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ é uma função não necessariamente linear. Da mesma maneira da seção anterior, a equação (3.37) pode ser rescrita, isto é

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + \frac{1}{2} \text{tr} \left[\sigma(t, x) \sigma^\top(t, x) \frac{\partial^2 v}{\partial x^2}(t, x) \right] + \left\langle b(t, x), \frac{\partial v}{\partial x}(t, x) \right\rangle + F(t, x, \cdot, \cdot) = 0 \\ v(T, x) = h(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.38)$$

com $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $r(t, x) = \sigma^\top(t, x)$ e

$$F(t, x, \cdot, \cdot) := F\left(t, x, v(t, x), \sigma^\top(t, x) \frac{\partial v}{\partial x}(t, x)\right).$$

No entanto, antes de enunciar e provar o último teorema, há que dar alguns conceitos e resultados preliminares. Nesse contexto, considere uma família de problemas de controle $S_{s,y}$ parametrizados por $\varepsilon \in [0, 1]$. Assim por cada problema $S_{s,y}^\varepsilon$, associa-se a equação dos estados (3.1) em que b e σ são substituídos por b^ε e σ^ε , respectivamente. Além disso tem-se a função objetivo $J^\varepsilon(s, y, u(\cdot))$ que coincide com $J(s, y, u(\cdot))$ para o problema $S_{s,y}$, mudando f e h para f^ε e h^ε . Nestas condições e considerando que valem as hipóteses $(S_{s,y}1)$ e $(S_{s,y}2)$ para b^ε , σ^ε , f^ε e h^ε , pode-se tomar a função valor $\mathbb{V}^\varepsilon$ definida da mesma forma que $\mathbb{V}$, tal que a toda a teoria da programação dinâmica, analisada no início deste capítulo, seja verdadeira para esta família de problemas de controle.

Lema 3.16. Suponha que as hipóteses $(S_{s,y}1)$ e $(S_{s,y}2)$ são verdadeiras com $\varphi^\varepsilon(t, x, u) := b^\varepsilon(t, x, u)$; $\varphi^\varepsilon(t, x, u) := \sigma^\varepsilon(t, x, u)$; $\varphi^\varepsilon(t, x, u) := f^\varepsilon(t, x, u)$; $\varphi^\varepsilon(t, x, u) := h^\varepsilon(x)$ para $\varepsilon \in [0, 1]$, tal que

$$\lim_{\varepsilon \rightarrow 0} \varphi^\varepsilon(t, x, u) = \varphi(t, x, u), \quad (3.39)$$

uniformemente em $(t, x, u) \in [0, T] \times \mathbb{K} \times U$. Nestas condições, tem-se que

$$\lim_{\varepsilon \rightarrow 0} \mathbb{V}^\varepsilon(s, y) = \mathbb{V}(s, y), \quad (3.40)$$

uniformemente em $(s, y) \in [0, T] \times \mathbb{K}$, sendo $\mathbb{K}$ qualquer compacto de $\mathbb{R}^n$.

Demonstração. Por (3.39) e tendo em conta a convergência uniforme em compactos, é possível encontrar uma função contínua $\eta : [0, 1] \times \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ não decrescente em seu domínio, com $\eta(0, R) = 0 \ \forall R \geq 0$, tal que

$$|\varphi^\varepsilon(t, x, u) - \varphi(t, x, u)| \leq \eta(\varepsilon, |x|), \quad \forall (t, x, u) \in [0, T] \times \mathbb{R}^n \times U. \quad (3.41)$$

Fixando $(s, y) \in [0, T] \times \mathbb{R}^n$, tome $u(\cdot) \in \mathcal{U}_{ad}[s, T]$, tal que para $\varepsilon \in [0, 1]$, $x^\varepsilon(\cdot) :=$

$x^\varepsilon(\cdot, s, y, u(\cdot))$ é solução de

$$\begin{cases} dx(t) = b^\varepsilon(t, x(t), u(t))dt + \sigma^\varepsilon(t, x(t), u(t))dW(t), & t \in [s, T] \\ x(s) = y \in \mathbb{R}^n, \end{cases} \quad (3.42)$$

em que $x^\varepsilon(s) = x(s) = y$. Assim para $r \in [s, T]$, tem-se

$$\begin{aligned} |x^\varepsilon(t) - x(t)|^2 &= \left| \int_s^t [b^\varepsilon(r, x^\varepsilon(r), u(r)) - b(r, x(r), u(r))] dr \right. \\ &\quad \left. + \int_s^t [\sigma^\varepsilon(r, x^\varepsilon(r), u(r)) - \sigma(r, x(r), u(r))] dW(r) \right|^2 \\ &\leq 4^2 \left| \int_s^t [b^\varepsilon(r, x^\varepsilon(r), u(r)) - b^\varepsilon(r, x(r), u(r))] dr \right|^2 \\ &\quad + 4^2 \left| \int_s^t [\sigma^\varepsilon(r, x^\varepsilon(r), u(r)) - \sigma^\varepsilon(r, x(r), u(r))] dW(r) \right|^2 \\ &\quad + 4^2 \left| \int_s^t [b^\varepsilon(r, x(r), u(r)) - b(r, x(r), u(r))] dr \right|^2 \\ &\quad + 4^2 \left| \int_s^t [\sigma^\varepsilon(r, x(r), u(r)) - \sigma(r, x(r), u(r))] dW(r) \right|^2. \end{aligned} \quad (3.43)$$

Logo, tendo em conta $(S_{s,y}2)$, as desigualdades de Burkholder-Davis-Gundy e (3.43) (compare também com (2.14)), vem que

$$\begin{aligned} E \left[\sup_{s \leq t \leq T} |x^\varepsilon(t) - x(t)|^2 \right] &\leq 16K_1 \int_s^T E \left[\sup_{s \leq r \leq T} |x^\varepsilon(r) - x(r)|^2 \right] dr \\ &\quad + 16K_2(T-s) E \left[\sup_{s \leq t \leq T} |b^\varepsilon(t, x(t), u(t)) - b(t, x(t), u(t))|^2 \right] \\ &\quad + 16K_3(T-s) E \left[\sup_{s \leq t \leq T} |\sigma^\varepsilon(t, x(t), u(t)) - \sigma(t, x(t), u(t))|^2 \right], \end{aligned}$$

donde pela desigualdade de Gronwall e (3.41), resulta

$$E \left[\sup_{s \leq t \leq T} |x^\varepsilon(t) - x(t)|^2 \right] \leq KE \left[\int_s^T |b^\varepsilon(r, x(r), u(r)) - b(r, x(r), u(r))|^2 dr \right]$$

$$\begin{aligned}
 & + KE \left[\int_s^T |\sigma^\varepsilon(r, x(r), u(r)) - \sigma(r, x(r), u(r))|^2 dr \right] \\
 & \leq K \int_s^T E\eta(\varepsilon, |x(r)|)^2 dr.
 \end{aligned} \tag{3.44}$$

Ora, notando outra vez $(S_{s,y}2)$, sabe-se que $\eta(\varepsilon, |x(t)|)^2 \leq 4L^2(1 + |x(t)|^2)$, $\forall t \in [s, T]$. Assim sendo, pela definição de esperança e pelo fato de η ser não decrescente, podemos escrever para cada $R > 0$,

$$\begin{aligned}
 E\eta(\varepsilon, |x(t)|)^2 &= \int_{|x(t)| \leq R} \eta(\varepsilon, |x(t)|)^2 dP + \int_{|x(t)| > R} \eta(\varepsilon, |x(t)|)^2 dP \\
 &\leq \int_{|x(t)| \leq R} \eta(\varepsilon, R)^2 dP + \int_{|x(t)| > R} 4L^2(1 + |x(t)|^2) dP \\
 &\leq \eta(\varepsilon, R)^2 + \int_{\sup_\varepsilon |x(t)| > R} 4L^2(1 + \sup_\varepsilon |x(t)|^2) dP.
 \end{aligned} \tag{3.45}$$

Faça

$$\begin{cases} \mathcal{A}_1(\varepsilon) := \eta(\varepsilon, R)^2 \\ \mathcal{A}_2(\varepsilon) := \int_{\sup_\varepsilon |x(t)| > R} 4L^2(1 + \sup_\varepsilon |x(t)|^2) dP \end{cases}$$

Logo, dado $\delta > 0$ e fazendo $R \rightarrow +\infty$, segue que $\mathcal{A}_1(\varepsilon) + \mathcal{A}_2(\varepsilon) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$, pois

$$\lim_{R \rightarrow +\infty} P\left\{ \sup_\varepsilon |x(t)| > R \right\} = 0, \tag{3.46}$$

lembrando que, neste caso, se tem $\sup_\varepsilon |x(t)| < +\infty$ P a.s, sendo que $E[\sup_\varepsilon |x(t)|] < +\infty$, $\forall t \in [s, T]$. Daí que, tomando δ tão pequeno quanto se queira e notando (3.45), fica

$$\lim_{\varepsilon \rightarrow 0} E\eta(\varepsilon, |x(t)|)^2 = 0,$$

ou seja por (3.44),

$$\lim_{\varepsilon \rightarrow 0} E\left[\sup_{s \leq t \leq T} |x^\varepsilon(t) - x(t)|^2 \right] = 0. \tag{3.47}$$

Finalmente, para concluir (3.40), basta da mesma forma como se fez em (3.43), ter em

atenção que

$$\begin{aligned}
 |J^\varepsilon(s, y, u(\cdot)) - J(s, y, u(\cdot))| &\leq K_1 E \left[\sup_{s \leq t \leq T} |x^\varepsilon(t) - x(t)| \right] + K_2 E [|h^\varepsilon(x(T)) - h(x(T))|] \\
 &\quad + K_3 E \left[\int_s^T |f^\varepsilon(t, x(t), u(t)) - f(t, x(t), u(t))| dt \right] \\
 &\leq K_1 E \left[\sup_{s \leq t \leq T} |x^\varepsilon(t) - x(t)| \right] + K_2 E \eta(\varepsilon, |x(T)|) \\
 &\quad + K_3 E \left[\int_s^T E \eta(\varepsilon, |x(t)|) dt \right]. \tag{3.48}
 \end{aligned}$$

Portanto, por (3.47) e (3.48), tem-se

$$\lim_{\varepsilon \rightarrow 0} |J^\varepsilon(s, y, u(\cdot)) - J(s, y, u(\cdot))| = 0.$$

Como o limite anterior vale para qualquer $u(\cdot) \in \mathcal{U}_{ad}[s, T]$, em especial valerá para o $\bar{u}$ que realiza o ínfimo no problema de controle, isto é, (3.40) também vale. $\square$

Sabe-se que se a função valor $\mathbb{V}$ é de classe $C^{1,2}([0, T] \times \mathbb{R}^n)$, esta é a única solução clássica da equação (3.6). No entanto, pode acontecer que essa equação seja degenerada, ou seja, que o traço da matriz $\sigma\sigma^\top$ no Hamiltoniano G seja nulo. Nesse caso, não se garante a solução clássica em (3.6). Sendo assim, fazendo $\sigma^\varepsilon(\sigma^\varepsilon)^\top = \sigma\sigma^\top + \varepsilon^2 I_n$ (veja [6], pág. 200-203), essa equação se transforma em não degenerada e tem a forma

$$\begin{cases} -\frac{\partial v^\varepsilon}{\partial t}(t, x) + \sup_{u \in U} G\left(t, x, u, -\frac{\partial v^\varepsilon}{\partial x}(t, x), -\frac{\partial^2 v^\varepsilon}{\partial x^2}(t, x)\right) + \frac{1}{2}\varepsilon^2 \Delta v^\varepsilon(t, x) = 0 \\ v^\varepsilon(T, x) = h(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \tag{3.49}$$

em que $\mathbb{V}^\varepsilon$ é a única solução clássica de (3.49), com $\varepsilon \in [0, 1]$. A unicidade da solução clássica pode ser encontrada em [6], pág. 228-242. Pelo lema anterior, resulta que tomando ε tão pequeno quanto se queira, a solução da PDE degenerada será limite de soluções de PDEs não degeneradas. Para o próximo resultado, faça $U = \{u_0\}$.

Lema 3.17. Seja a função valor $\mathbb{V}^\varepsilon$ não necessariamente de classe $C^{1,2}([0, T] \times \mathbb{R}^n)$. Isto é, assuma que $\mathbb{V}^\varepsilon \in C([0, T] \times \mathbb{R}^n)$ e é solução de viscosidade de

$$\begin{cases} -\frac{\partial v^\varepsilon}{\partial t}(t, x) + G^\varepsilon\left(t, x, -\frac{\partial v^\varepsilon}{\partial x}(t, x), -\frac{\partial^2 v^\varepsilon}{\partial x^2}(t, x)\right) = 0 \\ v^\varepsilon(T, x) = h^\varepsilon(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \tag{3.50}$$

onde $G^\varepsilon : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times S^n \rightarrow \mathbb{R}$ e $h^\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}$ são funções contínuas. Admita também que

$$\begin{cases} \lim_{\varepsilon \rightarrow 0} G^\varepsilon(t, x, p, \tilde{p}) = G(t, x, p, \tilde{p}) \\ \lim_{\varepsilon \rightarrow 0} h^\varepsilon(x) = h(x) \\ \lim_{\varepsilon \rightarrow 0} \mathbb{V}^\varepsilon(t, x) = \mathbb{V}(t, x), \end{cases} \quad (3.51)$$

uniformemente para qualquer compacto, nos seus domínios. Nestas condições, a função valor $\mathbb{V}$ é solução de viscosidade de

$$\begin{cases} -\frac{\partial v}{\partial t}(t, x) + G\left(t, x, -\frac{\partial v}{\partial x}(t, x), -\frac{\partial^2 v}{\partial x^2}(t, x)\right) = 0 \\ v(T, x) = h(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n. \quad (3.52)$$

Demonstração. Considere o caso da subsolução de viscosidade (para a supersolução de viscosidade, a prova é análoga). Seja então $\mathbb{V} \in C([0, T] \times \mathbb{R}^n)$ e tome $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$, tal que $\mathbb{V}(t_0, x_0) - \varphi(t_0, x_0)$ é ponto de máximo local estrito em $(t_0, x_0) \in [0, T] \times \mathbb{R}^n$. Fazendo ε tão pequeno quando se queira e como V^ε é por hipótese solução de viscosidade de (3.50), é possível encontrar para qualquer compacto $[0, T] \times \mathbb{K}$ de $[0, T] \times \mathbb{R}^n$, um par $(t_\varepsilon, x_\varepsilon)$ tendendo para (t_0, x_0) , tal que $\mathbb{V}^\varepsilon(t_\varepsilon, x_\varepsilon) - \varphi(t_\varepsilon, x_\varepsilon)$ é ponto de máximo local estrito em $(t_\varepsilon, x_\varepsilon)$ e tem-se

$$\begin{cases} -\frac{\partial \varphi}{\partial t}(t_\varepsilon, x_\varepsilon) + G^\varepsilon\left(t_\varepsilon, x_\varepsilon, -\frac{\partial \varphi}{\partial x}(t_\varepsilon, x_\varepsilon), -\frac{\partial^2 \varphi}{\partial x^2}(t_\varepsilon, x_\varepsilon)\right) \leq 0 \\ \mathbb{V}^\varepsilon(T, x) \leq h^\varepsilon(x). \end{cases}$$

Ora, G^ε é contínua por hipótese e φ tem primeira e segunda derivadas contínuas em x . Além disso, pela desigualdade triangular, resulta que

$$\begin{aligned} & \left| G^\varepsilon\left(t_\varepsilon, x_\varepsilon, -\frac{\partial \varphi}{\partial x}(t_\varepsilon, x_\varepsilon), -\frac{\partial^2 \varphi}{\partial x^2}(t_\varepsilon, x_\varepsilon)\right) - G\left(t_0, x_0, -\frac{\partial \varphi}{\partial x}(t_0, x_0), -\frac{\partial^2 \varphi}{\partial x^2}(t_0, x_0)\right) \right| \\ & \leq \left| G^\varepsilon\left(t_\varepsilon, x_\varepsilon, -\frac{\partial \varphi}{\partial x}(t_\varepsilon, x_\varepsilon), -\frac{\partial^2 \varphi}{\partial x^2}(t_\varepsilon, x_\varepsilon)\right) - G^\varepsilon\left(t_0, x_0, -\frac{\partial \varphi}{\partial x}(t_0, x_0), -\frac{\partial^2 \varphi}{\partial x^2}(t_0, x_0)\right) \right| \\ & + \left| G^\varepsilon\left(t_0, x_0, -\frac{\partial \varphi}{\partial x}(t_0, x_0), -\frac{\partial^2 \varphi}{\partial x^2}(t_0, x_0)\right) - G\left(t_0, x_0, -\frac{\partial \varphi}{\partial x}(t_0, x_0), -\frac{\partial^2 \varphi}{\partial x^2}(t_0, x_0)\right) \right|, \end{aligned}$$

donde por (3.51) e fazendo $\varepsilon \rightarrow 0$, resulta

$$\begin{cases} -\frac{\partial \varphi}{\partial t}(t_0, x_0) + G\left(t_0, x_0, -\frac{\partial \varphi}{\partial x}(t_0, x_0), -\frac{\partial^2 \varphi}{\partial x^2}(t_0, x_0)\right) \leq 0 \\ \mathbb{V}(T, x) \leq h(x). \end{cases}$$

Assim, a função valor $\mathbb{V}$ é subsolução de viscosidade de (3.52). $\square$

Observação 3.18. Os últimos dois resultados e a unicidade das soluções de viscosidade e clássica, permitem afirmar que, com $U = \{u_0\}$, se v^ε é solução clássica de (3.49) e v é solução de viscosidade de (3.6), temos

$$\lim_{\varepsilon \rightarrow 0} v^\varepsilon(t, x) = v(t, x) \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.53)$$

bastando para isso, considerar $\mathbb{V}^\varepsilon = v^\varepsilon$ e $\mathbb{V} = v$ em (3.53). Portanto, é perceptível que uma solução de viscosidade de uma PDE degenerada pode ser aproximada por uma solução clássica de uma PDE não degenerada. Este fato, será decisivo no próximo resultado.

Assuma a hipótese:

(HB) As funções $b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$; $\sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$; $F : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ são uniformemente contínuas e existe constantes L e λ positivas, tal que para $\varphi(t, x) := b(t, x)$; $\varphi(t, x) := \sigma(t, x)$; $\varphi(t, x) := F(t, x, y, z)$; $\varphi(t, x) := h(x)$, vem que

$$\begin{cases} |\varphi(t, x_1) - \varphi(t, x_2)| \leq L|x_1 - x_2| \\ \forall t \in [0, T]; \quad \forall x_1, x_2 \in \mathbb{R}^n \\ \\ |\varphi(t, 0)| \leq L \\ \forall t \in [0, T] \end{cases}$$

e

$$\begin{aligned} |F(t, x, y_1, z_1) - F(t, x, y_2, z_2)| &\leq \lambda(|y_1 - y_2| + |z_1 - z_2|) \\ \forall (t, x, y_1, z_1) \in [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m; \quad \forall (t, x, y_2, z_2) \in [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m. \end{aligned}$$

Fixe um espaço de probabilidade com filtração $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, sendo $W(\cdot)$ o movimento Browniano usual de dimensão m , com $t \in [0, T]$. Seja o par $(Y(\cdot), Z(\cdot))$ a solução adaptada da BSDE

$$\begin{cases} dY(s) = -F(s, X(s), Y(s), Z(s))ds + Z(s)dW(s) \\ Y(T) = h(X(T)) \end{cases} \quad s \in [t, T]; \quad P \text{ a.s.}, \quad (3.54)$$

com $Y(0) = 0$ e onde $X(\cdot) := X(\cdot, t, x)$ é a única solução forte de (3.28), com $(t, x) \in [0, T) \times \mathbb{R}^n$. Nestas condições, temos um importante resultado.

Teorema 3.19. Se vale a hipótese (HB) então v é a única solução de viscosidade de (3.37), tal que

$$v(t, x) = E[Y(t; t, x)] = Y(t; t, x). \quad (3.55)$$

Além disso, se (3.37) admite solução clássica então ela é única e é representada por (3.55).

Demonstração. Suponha que existe uma solução clássica de (3.37). Por (HB) , resulta do Teorema 2.16, a unicidade do par $(Y(\cdot), Z(\cdot))$ para a BSDE (3.54). Assim, aplicando a fórmula de Itô em $v(r, X(r))$ com $r \in [s, T]$ e $X(\cdot, t, x)$ $\mathcal{F}_s$ - progressivamente mensurável, fica

$$\begin{aligned} v(T, X(T)) - v(s, X(s)) &= \int_s^T \left\{ \frac{\partial v}{\partial r}(r, X(r)) + \left\langle \frac{\partial v}{\partial x}(r, X(r)), b(r, X(r)) \right\rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr} \left[\sigma(r, X(r)) \sigma^\top(r, X(r)) \frac{\partial^2 v}{\partial x^2}(r, X(r)) \right] \right\} dr \\ &\quad + \int_s^T \left\langle \sigma^\top(r, X(r)), \frac{\partial v}{\partial x}(r, X(r)) dW(r) \right\rangle. \end{aligned}$$

Logo, por (3.38) tem-se

$$\begin{aligned} v(s, X(s)) &= h(X(T)) + \int_s^T F \left[r, X(r), v(r, X(r)), \sigma^\top(r, X(r)) \frac{\partial v}{\partial x}(r, X(r)) \right] dr \\ &\quad - \int_s^T \left\langle \sigma^\top(r, X(r)), \frac{\partial v}{\partial x}(r, X(r)) dW(r) \right\rangle. \end{aligned} \quad (3.56)$$

Portanto, sendo $\left(v(r, X(r)), \sigma^\top(r, X(r)) \frac{\partial v}{\partial x}(r, X(r)) \right)$ um par de processos adaptados de (3.54), resulta necessariamente pela unicidade dessa solução que (compare com (3.13)),

$$\begin{cases} Y(s) = v(s, X(s)) \\ Z(s) = \sigma^\top(s, X(s)) \frac{\partial v}{\partial x}(s, X(s)) \end{cases} \quad \forall s \in [t, T]; \quad P \text{ a.s.} \quad (3.57)$$

Assim, por (3.28) e notando que o processo $Y(\cdot)$ é determinístico em $s = t$, segue que $v(t, x) = E[Y(t; t, x)] = Y(t; t, x)$ é a única solução clássica de (3.37). Suponha agora que a PDE (3.37) não admite uma solução clássica, isto é, que seja uma PDE possivelmente

degenerada. Pela hipótese (HB) , juntamente com os lemas (3.16) e (3.17), essa PDE admite uma única solução de viscosidade v , que é limite da convergência uniforme em (3.53), para qualquer compacto $[0, T] \times \mathbb{K}$, $\forall (t, x) \in [0, T] \times \mathbb{R}^n$, com $\varepsilon \in [0, 1]$. Assim, tem-se que v^ε será a solução clássica de

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + A^\varepsilon \frac{\partial^2 v}{\partial x^2}(t, x) + B^\varepsilon \frac{\partial v}{\partial x}(t, x) + F^\varepsilon\left(t, x, v(t, x), r^\varepsilon(t, x) \frac{\partial v}{\partial x}(t, x)\right) = 0 \\ v(T, x) = h^\varepsilon(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.58)$$

que é equivalente a escrever

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + \frac{1}{2} \text{tr} \left[\sigma^\varepsilon(t, x) (\sigma^\varepsilon)^\top(t, x) \frac{\partial^2 v}{\partial x^2}(t, x) \right] + \left\langle b^\varepsilon(t, x), \frac{\partial v}{\partial x}(t, x) \right\rangle + F^\varepsilon(t, x, \cdot, \cdot) = 0 \\ v(T, x) = h^\varepsilon(x) \end{cases} \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (3.59)$$

com

$$F^\varepsilon(t, x, \cdot, \cdot) := F^\varepsilon\left(t, x, v(t, x), (\sigma^\varepsilon)^\top(t, x) \frac{\partial v}{\partial x}(t, x)\right).$$

Assuma que b^ε , σ^ε , F^ε e h^ε convergem uniformemente para b , σ , F e h para qualquer compacto em seus domínios, impondo que $\sigma^\varepsilon (\sigma^\varepsilon)^\top$ é no mínimo $\sigma \sigma^\top + \varepsilon^2 I_n$, para $(t, x) \in [0, T] \times \mathbb{R}^n$. Assim, utilizando novamente a fórmula de Itô em $v^\varepsilon(s, X^\varepsilon(s))$ com $s \in [t, T]$ e $X^\varepsilon(\cdot, t, x)$ $\mathcal{F}_s$ - progressivamente mensurável, segue que

$$\begin{cases} Y^\varepsilon(s) = v^\varepsilon(s, X^\varepsilon(s)) \\ Z^\varepsilon(s) = (\sigma^\varepsilon)^\top(s, X^\varepsilon(s)) \frac{\partial v^\varepsilon}{\partial x}(s, X^\varepsilon(s)) \end{cases} \quad \forall s \in [t, T]; \quad P \text{ a.s.}, \quad (3.60)$$

onde $X^\varepsilon(\cdot)$ é a única solução forte de (3.28) e o par adaptado $(Y^\varepsilon(\cdot), Z^\varepsilon(\cdot)) \in B^2$ é a solução de

$$\begin{cases} dY^\varepsilon(s) = -F^\varepsilon(s, X^\varepsilon(s), Y^\varepsilon(s), Z^\varepsilon(s)) ds + Z^\varepsilon(s) dW(s) \\ Y^\varepsilon(T) = h^\varepsilon(X^\varepsilon(T)) \end{cases} \quad s \in [t, T]; \quad P \text{ a.s.}, \quad (3.61)$$

com $Y^\varepsilon(0) = 0$. Além disso, por um lado, vem por (3.47) com as devidas alterações na notação, que

$$\lim_{\varepsilon \rightarrow 0} E \left[\sup_{t \leq s \leq T} |X^\varepsilon(s) - X(s)|^2 \right] = 0. \quad (3.62)$$

Por outro lado, temos que garantir que

$$\lim_{\varepsilon \rightarrow 0} \left\{ E \left[\sup_{0 \leq t \leq T} |Y^\varepsilon(t) - Y(t)|^2 \right] + E \int_0^T |Z^\varepsilon(s) - Z(s)|^2 ds \right\} = 0. \quad (3.63)$$

Afirmção 3.19.1. Supondo a Hipótese (HB) , tem-se para $s \in [t, T]$

$$E \left[\sup_{0 \leq t \leq T} |\hat{Y}(t)|^2 \right] + E \int_0^T |\hat{Z}(s)|^2 ds \leq CE \int_0^T |\hat{F}(s)|^2 ds, \quad (3.64)$$

com $\hat{Y}(s) := Y^\varepsilon(s) - Y(s)$, $\hat{Z}(s) := Z^\varepsilon(s) - Z(s)$ e

$$\hat{F}(s) := F^\varepsilon(s, Y(s), Z(s)) - F(s, Y(s), Z(s)).$$

De fato, olhando para (HB) e aplicando a fórmula de Itô a $|Y(t)|^2$ para $t \in [0, T]$ (Compare com a prova do Teorema 2.16), fica

$$\begin{aligned} |\hat{Y}(t)|^2 &= 2 \int_0^t \langle \hat{Y}(s), [F^\varepsilon(s, Y^\varepsilon(s), Z^\varepsilon(s)) - F(s, Y(s), Z(s))] \rangle ds \\ &\quad + 2 \int_0^t \langle \hat{Y}(s), \hat{Z}(s) dW(s) \rangle + \int_0^t |\hat{Z}(s)|^2 ds \\ &\leq 2 \int_t^T |\hat{Y}(s)| |F^\varepsilon(s, Y^\varepsilon(s), Z^\varepsilon(s)) - F(s, Y(s), Z(s))| ds \\ &\quad - 2 \int_t^T \langle \hat{Y}(s), \hat{Z}(s) dW(s) \rangle - \int_t^T |\hat{Z}(s)|^2 ds \\ &\leq \int_t^T [2|\hat{Y}(s)| |F^\varepsilon(s, Y^\varepsilon(s), Z^\varepsilon(s)) - F^\varepsilon(s, Y(s), Z(s))| + 2|\hat{Y}(s)| |\hat{F}(s)|] ds \\ &\quad - 2 \int_t^T \langle \hat{Y}(s), \hat{Z}(s) dW(s) \rangle - \int_t^T |\hat{Z}(s)|^2 ds \\ &\leq \int_t^T [2\lambda |\hat{Y}(s)| (|\hat{Y}(s)| + |\hat{Z}(s)|) + 2|\hat{Y}(s)| |\hat{F}(s)|] ds \\ &\quad - 2 \int_t^T \langle \hat{Y}(s), \hat{Z}(s) dW(s) \rangle - \int_t^T |\hat{Z}(s)|^2 ds. \end{aligned}$$

Sabendo que $2ab \leq \varepsilon a^2 + \frac{b^2}{\varepsilon}$ para $\varepsilon > 0$ e qualquer número real a e b , resulta para $\lambda > 0$ e $\varepsilon = 2$,

$$\begin{aligned} |\hat{Y}(t)|^2 + \int_t^T |\hat{Z}(s)|^2 ds &\leq \int_t^T [2\lambda |\hat{Y}(s)| (|\hat{Y}(s)| + |\hat{Z}(s)|) + 2|\hat{Y}(s)| |\hat{F}(s)|] ds \\ &\quad - 2 \int_t^T \langle \hat{Y}(s), \hat{Z}(s) dW(s) \rangle \\ &\leq \int_t^T (1 + 2\lambda + 2\lambda^2) |\hat{Y}(s)|^2 ds + \int_t^T \frac{1}{2} |\hat{Z}(s)|^2 ds + \int_t^T |\hat{F}(s)|^2 ds \\ &\quad - 2 \int_t^T \langle \hat{Y}(s), \hat{Z}(s) dW(s) \rangle, \end{aligned}$$

donde aplicando a esperança no dois membros da desigualdade, tem-se

$$\begin{aligned} E|\hat{Y}(t)|^2 + E \int_t^T |\hat{Z}(s)|^2 ds &\leq \int_t^T (1 + 2\lambda + \lambda^2) E|\hat{Y}(s)|^2 ds \\ &\quad + E \int_0^T |\hat{F}(s)|^2 ds + E \int_t^T \frac{1}{2} |\hat{Z}(s)|^2 ds. \end{aligned}$$

Assim, a desigualdade de Gronwall em (1.4) fornece

$$E|\hat{Y}(t)|^2 + E \int_t^T |\hat{Z}(s)|^2 ds \leq \lambda_1 E \int_0^T |\hat{F}(s)|^2 ds, \quad \forall t \in [0, T]. \quad (3.65)$$

Além disso, pela estimativa (2.24) e tendo em conta que para $t = 0$ em (3.65),

$$E \int_0^T |\hat{Z}(s)|^2 ds \leq E \int_0^T |\hat{F}(s)|^2 ds, \quad (3.66)$$

vem que

$$\begin{aligned} E \left[\sup_{0 \leq t \leq T} |\hat{Y}(t)|^2 \right] &\leq E \int_0^T |\hat{F}(s)|^2 ds + K E \left[\int_0^T |\hat{Y}(s)|^2 |\hat{Z}(s)|^2 ds \right]^{\frac{1}{2}} \\ &\leq E \int_0^T |\hat{F}(s)|^2 ds + K E \left[\sup_{0 \leq t \leq T} |\hat{Y}(t)| \left(\int_0^T |\hat{Z}(s)|^2 ds \right)^{\frac{1}{2}} \right] \end{aligned}$$

$$\leq E \int_0^T |\hat{F}(s)|^2 ds + \frac{1}{2} E \left[\sup_{0 \leq t \leq T} |\hat{Y}(t)|^2 \right] + \frac{K^2}{2} E \int_0^T |\hat{Z}(s)|^2 ds.$$

Isto é, por (3.66) segue que

$$E \left[\sup_{0 \leq t \leq T} |\hat{Y}(t)|^2 \right] + E \int_0^T |\hat{Z}(s)|^2 ds \leq (K^2 + 3) E \int_0^T |\hat{F}(s)|^2 ds,$$

donde fazendo $C = K^2 + 3$, a afirmação 3.19.1 vale. Logo, como F^ε converge uniformemente para F dentro de qualquer compacto em seu domínio, tem-se

$$\lim_{\varepsilon \rightarrow 0} |F^\varepsilon(s, X(s), Y(s), Z(s)) - F(s, X(s), Y(s), Z(s))| = 0,$$

uniformemente. Assim, garante-se então por (3.62) e (3.64), que (3.63) é verdadeiro e portanto tem-se

$$\lim_{\varepsilon \rightarrow 0} \left\{ E \left[\sup_{t \leq r \leq T} |Y^\varepsilon(r) - Y(r)|^2 \right] \right\} = 0. \quad (3.67)$$

Finalmente pela desigualdade triangular e lembrando que $X(t) = X^\varepsilon(t) = x$ para $t \in [0, T]$, segue por (3.53) e (3.60) com $s = t$,

$$\begin{aligned} |v(t, x) - Y(t; t, x)| &\leq |v(t, x) - v^\varepsilon(t, x)| + |v^\varepsilon(t, x) - Y(t; t, x)| \\ &\leq |v(t, x) - v^\varepsilon(t, x)| + |Y^\varepsilon(t; t, x) - Y(t; t, x)| \\ &\leq |v(t, x) - v^\varepsilon(t, x)| + E \left[\sup_{t \leq r \leq T} |Y^\varepsilon(r) - Y(r)|^2 \right], \end{aligned} \quad (3.68)$$

donde passando ao limite quando $\varepsilon \rightarrow 0$ em ambos os membros da desigualdade (3.68), resulta que a única solução de viscosidade de (3.37) é dada por $v(t, x) = E[Y(t; t, x)] = Y(t; t, x)$, $\forall t \in [0, T]$. $\square$

Observação 3.20. Na análise dos Teoremas 3.15 e 3.19 foi perceptível no primeiro teorema, que a solução da PDE linear pode ser determinada, apenas utilizando uma SDE apropriada. Por outro lado, no segundo teorema, foi necessário utilizar uma BSDE pois nesse caso, a PDE que se apresentava não era necessariamente linear. Com efeito, para justificar a unicidade da solução de viscosidade em (3.32), a abordagem poderá ser feita via BSDE. Para isso, basta considerar o caso linear da função geradora F em (3.54), ou seja, faça

$$F(t, X(t), Y(t), Z(t)) = c(t, X(t))v(t, x) + g(t, X(t)) \quad \forall t \in [0, T]. \quad (3.69)$$

Ora, pelo Teorema 3.19, tem-se uma única solução de viscosidade v , dada por $v(t, x) = Y(t; t, x)$. Assim, pelos Teoremas 2.17 e 2.23 do *Capítulo 2*, existe uma solução explícita

para a BSDE linear que é dada por (3.33) (compare com (2.40)), ficando assim justificada a unicidade da solução de viscosidade em (3.32).

Como nota final, refira-se o fato das PDEs tratadas neste capítulo serem terminais, na medida em que o tempo final T está fixado na função h . No entanto, o leitor poderá questionar se toda esta teoria será válida para PDEs, fixando agora o tempo inicial em 0. A resposta é afirmativa assumindo algumas condições na função Hamiltoniana G , fazendo uma mudança de variável em $t \in [0, T]$ na função $v(t, x)$, isto é, $\tilde{v}(t, x) := v(T - t, x)$, com $(t, x) \in [0, T] \times \mathbb{R}^n$. Com esta alteração, ter-se-ia em analogia ao Teorema 3.19, $\tilde{v}(t, x) = v(T - t, x) = Y(T - t; T - t, x) = \tilde{Y}(0; T - t, x)$.

Referências Bibliográficas

- [1] BREZIS H., *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010.
- [2] DELLACHERIE C. & MEYER, P.A., *Probability and Potentials*, North Holland Publishing Company, 1978.
- [3] FLEMING W.H. & SONER, H.M., *Controlled Markov Processes and Viscosity Solutions*, Springer, 2006.
- [4] LAMBERTON D., *Optimal stopping and American options*, Ljubljana Summer School on Financial Mathematics, 2009.
- [5] KARATZAS I. & SHREVE S. E., *Brownian Motion and Stochastic Calculus*, Springer-Verlag New York, Second Edition, 1991.
- [6] KRYLOV N.V., *Controlled Diffusion Processes*, Springer, 2009.
- [7] OKSENDAL B., *Stochastic Differential Equations - An introduction with applications*, Springer - Verlag, Heidelberg London New York, Sixth Edition, 2010.
- [8] PHAM H., *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer-Verlag Berlin Heidelberg, Second Edition, 2009.
- [9] SHIRYAEV A. N., *Probability (Graduate Texts in Mathematics)*, Springer, Second Edition, 1995.
- [10] SMIRNOV G. V., *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, 2002.
- [11] YONG J. & ZHOU X.Y., *Stochastic Controls, Hamiltonian Systems and HJB Equations*, Springer, 1997.
- [12] ZEIDLER E., *Nonlinear Functional Analysis and Its Applications, I: Fixed-Point Theorems*, Springer-Verlag, 1986.