Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea Magistrale in Matematica

Bellman functions and their method in harmonic analysis

Tesi di Laurea in analisi armonica

Relatore:
Chiar.mo Prof.
Arcozzi Nicola

Presentata da:
Cavina Michelangelo

Quarta Sessione
Anno Accademico 2018/2019

Contents

Introduction

The goal of this work is to use the Bellman functions to prove theorems about inequalities over dyadic trees.

Bellman functions come from the theory of stochastic optimal control: they are solutions to the problem of maximizing the earned average payoff of a controlled stochastic process. Bellman functions are also solutions of the well known Hamilton-Bellman-Jacobi (HJB) equation, and a converse of this fact holds too, i.e. a smooth function that satisfies the HJB equation with proper boundary conditions is also a solution of the problem of optimal stochastic control associated to the HJB equation.

In the middle of the 90's F. Nazarov, S. Treil and A. Volberg, working over problems connected with harmonic analysis, developed a method to prove theorems about inequalities over dyadic trees which is connected to the theory of the Bellman functions. In principle they provide a unified approach to a vast class of integral inequalities. The restriction to trees is not so restrictive, since many inequalities in harmonic analysis and potential theories can be reduced to their dyadic counterparts.

The method consists of finding a mapping Φ from the data $(d_1, d_2, \ldots, d_m)$ of the theorem that we want to prove to a domain of points $D \subseteq \mathbb{R}^n$ and then find a function $g : D \to \mathbb{R}$ such that g is bounded and g satisfies a proper inequality, which we will refer to as main inequality. This strategy will allow us to compute the value of the composition of the function g and the mapping Φ and use the main inequality that g satisfies to gain the needed information to solve the problem.

The reason why this method is connected with the theory of the Bellman functions in stochastic optimal control is the fact that the main inequality that g satisfies represents an inequality of supersolution (or subsolution depending on the problem) for the HJB equation, so the function g that we are looking for to solve the problem is a supersolution of the HJB equation and, if we are lucky, the function g may even be a solution of the HJB equation, and this way we would find an expression for the solution of the stochastic optimal control problem associated to the HJB equation.

The first chapter of this work includes notions of stochastic analysis needed for the understanding of the theory behind the Bellman functions. The notations, definitions and theorems listed in this chapter are all thoroughly explained in the text [3] from B. Øksendal, in the chapters 1,2,3,4,5,7 and 9. The theory of the Bellman functions requires the notions of stochastic analysis necessary to define the Itô integral, which is used to define a Bellman function v as

$$v(x) = \sup_{\{u_t\}_{t\geq 0}} E^x \left[\int_s^{\widehat{T}} F(r, X_r, u_r)dr + K(\widehat{T}, X_{\widehat{T}})\chi_{\widehat{T}<+\infty} dB_r \right]$$

where X_t is a stochastic process solution of the stochastic differential equation

$$X_h = X_h^x = x + \int_s^h b(r, X_r, u_r)dr + \int_s^h \sigma(r, X_r, u_r)dB_r; \quad h \geq s$$

of coefficients b and σ, $\{B_t\}_{\geq 0}$ is a Brownian motion, $\widehat{T}$ is a proper stopping time and $\{u_t\}t \geq 0$ is an admissible control process. Here F is a profit density and K is a "bequest" function (gain at the moment of retirement). So a Bellman function v is the solution of a stochastic optimal control problem that consists of finding the maximum average gain over a trajectory of a controlled process $\{X_t\}_{t\geq 0}$.

This chapter also includes the notions needed to prove the theorem about the Bellman function being a solution of the HJB equation, and its converse, i.e. it contains the definitions and theorems about the strong Markov property, infinitesimal generator of a stochastic process, Dynkin's formula and the Dirichlet-Poisson problem.

The second chapter in this work will present the theory of the Bellman functions in stochastic control. It will include the statement of the problem of stochastic optimization and the definition of Bellman function, followed by the proof of theorem about the Bellman function being a solution of the HJB equation, and its converse. The topics of this chapter come from chapter 11 of the text [3] from B. Øksendal.

The third chapter in this work shows an example of a problem that can be solved using the Bellman function's method and explains how to use the method, and will also show some of the connections between the analytical and the stochastic aspects of the problem. Our notations for this chapter will be the same notations as the ones in the article [2], which analized this problem as well. and The chosen problem is an inequality over the dyadic tree about A_∞ weights, that is related to the characterizations of Carlseson measures, for more details over this topic see [6] and [7]. The problem will be solved step by step using the Bellman function method as an example about how the method works in a general case. The problem will be solved with a function g that is a solution of the HJB equation associated to the problem, we will also show with a heuristic argument that the function g found and used to solve the problem is actually the Bellman function associated to the problem.

The fourth chapter uses the Bellman function's method to prove Hardy's inequality for the dyadic tree in the general L^p case with $1 < p < +\infty$. This problem was already analyzed in the specific case where $p = 2$ by the paper [1] from N. Arcozzi, I. Holmes, P. Mozolyako and A. Volberg, and solved with the Bellman function method, so this work extends the previous proof to the general case. Hardy's inequality comes from Complex analysis and it is useful to characterize Carleson m easures.

The problem will be solved considering the function

$$\mathcal{B}(F, f, A, v) = F - \frac{f^p}{(A + v)^{p-1}}$$

defined over a proper convex domain. We will prove that the function B is a concave, bounded function that satisfies a proper main inequality, and we will show that this function with these properties allows us to prove Hardy's inequality.

Chapter 1

Preliminary notions of stochastic analysis

1.1 Notations and definitions

We will need the definitions of random variable and stochastic process, and we will use in most cases the same notations as the text [3] from B. Øksendal. We recommend to check a text of probability and measure theory for the basic notions of probability needed in this work.

Definition 1.1.1. We denote with $\mathscr{B}(\mathbb{R}^d)$ the σ-algebra over the set $\mathbb{R}^d$ generated by the Borel subsets of $\mathbb{R}^d$.
Given a probability space $(\Omega, \mathscr{F}, P)$, where Ω is a set, $\mathscr{F}$ is a σ-algebra over Ω and $P : \mathscr{F} \to \mathbb{R}$ is a probability measure over Ω, a random variable

$$Z : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$$

is an application

$$Z : \Omega \longrightarrow \mathbb{R}^d$$

measurable with respect to the σ-algebras $\mathscr{F}$ and $\mathscr{B}(\mathbb{R}^d)$.
Given two measurable spaces $(\Omega_1, \mathscr{F}_1)$, $(\Omega_2, \mathscr{F}_2)$ an application

$$Z : \Omega_1 \longrightarrow \Omega_2$$

is measurable if, for all $A \in \mathscr{F}_2$, $Z^{-1}(A) \in \mathscr{F}_1$.

Definition 1.1.2. Given a probability space $(\Omega, \mathscr{F}, P)$, given a set of times I, given for each $t \in I$ a random variable

$$X_t : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d)) \tag{1.1}$$

we define as a stochastic process in $\mathbb{R}^n$ over the probability space $(\Omega, \mathscr{F}, P)$ the collection of random variables $\{X_t\}_{t \in I}$.

Notation 1.1. Given a probability space $(\Omega, \mathscr{F}, P)$ and a set of times I, given a stochastic process $\{X_t\}_{t\in I}$

$$X_t : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$$

we will use the following notations to refer to $\{X_t\}_{t\in I}$:

1. $\{X_t\}_{t\in I}$ is the application X that to each time $t \in I$ associates the random variable X_t, i.e. for each $t \in I$

$$X : I \longrightarrow (R^d)^\Omega \tag{1.2}$$

$$X(t)(\omega) := X_t(\omega) \quad \forall \omega \in \Omega$$

2. $\{X_t\}_{t\in I}$ is the application X that to each $\omega \in \Omega$ associates the trajectory $t \mapsto X_t(\omega)$, i.e. for each $\omega \in \Omega$

$$X : \Omega \longrightarrow (R^d)^I \tag{1.3}$$

$$X(\omega)(t) := X_t(\omega) \quad \forall t \in I$$

3. $\{X_t\}_{t\in I}$ is the application X defined by

$$X : I \times \Omega \longrightarrow R^d \tag{1.4}$$

$$X(t, \omega) := X_t(\omega) \quad \forall (t, \omega) \in I \times \Omega$$

The notations in (1.1), (1.2), (1.3) and (1.4) are equivalent to each other, so we will use each one of them indiscriminately.
We will also denote a value $X_t(\omega) \in \mathbb{R}^d$ by

$$X_t(\omega) = X(\omega)(t) = X(t, \omega) = X(t)(\omega) \quad \forall \omega \in \Omega,\ \forall t \in I$$

We are going to enunciate the notions needed to define a Brownian motion. We recommend to check chapter 2 from [3] for a more detailed exposition.
The definition of finite-dimensional distributions is the core part in the construction of many stochastic processes, one example being the Brownian motion.

Definition 1.1.3. Given a stochastic process X=$\{X_t\}_{t\in T}$ in $\mathbb{R}^n$

$$X_t : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$$

where $T = [0, +\infty)$, we define as finite-dimensional distributions of the proccess X the measures $\mu_{t_1,\ldots,t_k}$ defined over the Borel σ-algebra $\mathscr{B}(\mathbb{R}^{nk})$, for $K = 1, 2, \ldots$, by

$$\mu_{t_1,\ldots,t_k}(F_1 \times F_2 \times \cdots \times F_k) = P[X_{t_1} \in F_1, \ldots, X_{t_k} \in F_k]; \quad t_i \in T$$

We recall the definition of expected value and conditional expectation from the basics of the theory of probability.

Definition 1.1.4. Given a random variable

$$X : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$$

we denote as expected value of X with respect to P the real number

$$E(X) = \int_\Omega X(\omega) dP(\omega)$$

Let $\mathscr{G}$ be a σ-algebra, $\mathscr{G} \subseteq \mathscr{F}$. Suppose that $E(|X|) < +\infty$. We denote as a realization of the conditional expectation of X given $\mathscr{G}$ (with respect to P) a random variable

$$Z : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$$

such that

1. Z is $\mathscr{G}$-measurable
2. $$\int_G Z(\omega) dP(\omega) = \int_G X(\omega) dP(\omega) \quad \text{for all } G \in \mathscr{G}$$

We will write $Z = E[X|\mathscr{G}]$ to denote that Z is a realization of the conditional expectation of X given $\mathscr{G}$. For all Z_1, Z_2 realizations of the conditional expectation of X given $\mathscr{G}$, then $Z_1 = Z_2$ almost surely with respect to P, so we will sometimes just write $E[X|\mathscr{G}]$ in expressions to denote a realization Z of the conditional expectation when the expression is true for every possible choice of Z realization of the conditional expectation. We recommend to check appendix B from Øksendal [3] or a text of probability for an exposition over the conditional expectation.

The definition of filtration is needed for the theory of stochastic processes, and it represents the amount of "information" we know at each time t about the configuration of the stochastic process. The concept or martingale is a key element in the theory of stochastic processes, and it is also needed for the definition of the Itô integral. It represents a stochastic process $\{X_t\}_{t\geq 0}$ such that X_t can be estimated at a time $s < t$ by considering X_s.

Definition 1.1.5. Given a measurable space $(\Omega, \mathscr{F})$, a filtration of said space is a family $\mathcal{M} = \{\mathcal{M}_t\}_{t\geq 0}$ of σ-algebras $\mathcal{M}_t \subseteq \mathscr{F}$ such that

$$0 \leq s < t \Longrightarrow \mathcal{M}_s \subseteq \mathcal{M}_t$$

A n-dimensional stochastic process $\{M_t\}_{t\geq 0}$ on a probability space $(\Omega, \mathscr{F}, P)$ is called a martingale with respect to a filtration $\{\mathcal{M}_t\}_{t\geq 0}$ (and with respect to P) if

(i) M_t is $\mathcal{M}_t$-measurable for all t

(ii) $E[|M_t|] < +\infty$ for all t

(iii) $E[M_s \mid \mathcal{M}_t] = M_t$ for all $s \geq t$

Stopping times are a key element in the theory of stochastic processes. They are random times with good properties that make them usable to replace deterministic times in most of the theorems about stochastic analysis.

Definition 1.1.6. Let $(\Omega, \mathscr{F}, P)$ be a probability space, let $\{\mathcal{N}_t\}_{t\geq 0}$ be a filtration. A function

$$\tau : \Omega \longrightarrow [0, +\infty]$$

is called a (strict) stopping time with respect to $\{\mathcal{N}_t\}_{t\geq 0}$ if

$$\{\omega \in \Omega \mid \tau(\omega) \leq t\} \in \mathcal{N}_t \quad \text{for all } t \geq 0$$

Let $\mathcal{N}_\infty$ be the smallest σ-algebra containing $\mathcal{N}_t$ for all $t \geq 0$. Then we define by $\mathcal{N}_\tau$ the σ-algebra of all sets $N \in \mathcal{N}_\infty$ such that

$$N \cap \{\tau \leq t\} \in \mathcal{N}_t \quad \text{for all } t \leq 0$$

A stochastic process $\{X_t\}_{t\geq 0}$ over $(\Omega, \mathscr{F}, P)$

$$X_t : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$$

is called adapted to the fitration $\{\mathcal{N}_t\}_{t\geq 0}$ if X_t is $\mathcal{N}_t$-measurable for all $t \geq 0$.
We define X_τ as the random variable

$$X_\tau : (\Omega, \mathcal{N}_\tau, P) \longrightarrow (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$$

$$X_\tau(\omega) = \begin{cases} X_{\tau(\omega)}(\omega) & \text{if } \tau(\omega) < +\infty \\ 0 & \text{if } \tau(\omega) = +\infty \end{cases}$$

Infact it can be proved that X_τ defined in this way is measurable with respect to $\mathcal{N}_\tau$ and $\mathscr{B}(\mathbb{R}^n)$.

The Brownian motion is the starting point for the theory of Itô processes. It is the most important stochastic process for this work, we recommend to check a textbook about stochastic processes to get a thorough explanation of this topic, Øksendal explains this topic on chapter 2 from [3].

Definition 1.1.7 (Brownian motion)**.** Let $x \in \mathbb{R}^n$, $s \in \mathbb{R}$ be a fixed point and a fixed time. Define

$$p(t_1, y_1, t_2, y_2) = (2\pi(t_2 - t_1))^{-\frac{n}{2}} \cdot \exp(-\frac{|y_2 - y_1|^2}{2(t_2 - t_1)}) \quad \text{for } y \in \mathbb{R}^n,\ t > 0$$

For $k = 1, 2, \dots$, for $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$ define a measure $\nu_{t_1,\dots,t_k}$ on $\mathscr{B}(\mathbb{R}^{nk})$ by

$$\nu_{t_1,\dots,t_k}(F_1 \times \dots \times F_k) = \tag{1.5}$$

$$= \int\limits_{F_1\times\dots\times F_k} p(s, x, t_1, x_1)p(t_1, x_1, t_2, x_2) \dots p(t_{k-1}, x_{k-1}, t_k, x_k)dx_1 \dots dx_k$$

where we use the convention that $p(t, y, t, z)dz = d\delta_y(z)$, the Dirac delta measure centered at y.
We define a (version of) n-dimensional Brownian motion starting from x at the time s as a stochastic process $B = \{B_t\}_{t\geq s}$ on a probability space $(\Omega, \mathscr{F}, P^x)$

$$B_t : (\Omega, \mathscr{F}, P^x) \longrightarrow (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$$

such that the finite-dimensional of B are given by (1.5), i.e.

$$P^x(B_{t_1} \in F_1, \dots, B_{t_k} \in F_k) =$$

$$= \int\limits_{F_1 \times \dots \times F_k} p(s, x, t_1, x_1) p(t_1, x_1, t_2, x_2) \dots p(t_{k-1}, x_{k-1}, t_k, x_k) dx_1 \dots dx_k$$

for all $k \in \mathbb{N}$, $t_1, \dots, t_k \in [s, +\infty)$, for all $F_1, \dots, F_k$ Borel subsets of $\mathbb{R}^n$.
The existence of a process with such properties is guaranteed by Kolmogorov's extension Theorem 1.2.1.
A Brownian motion $B = \{B_t\}_{t \geq 0}$ satisfies the condition (1.13) in Kolmogorov's continuity Theorem 1.2.2 with $\alpha = 4$, $\beta = 1$, $D = n(n+2)$, so the Theorem guarantees that there exists a continuous modification of B.

The concept of modification of a stochastic process is needed to understand Kolmogorov's continuity Theorem that proves that the Brownian motion can be considered a continuous process.

Definition 1.1.8. Let $X = \{X_t\}_{t \in I}$, $Y = \{Y_t\}_{t \in I}$ be stochastic processes on the same probability space $(\Omega, \mathscr{F}, P)$. We say that X is a version (or a modification) of Y if, for all $t \in I$

$$P(\{\omega \in \Omega \mid X_t(\omega) = Y_t(\omega)\}) = 1$$

We need a notation for the σ-algebra generated by a Brownian motion at a time t for many propositions, expecially for the important Markov property.

Definition 1.1.9. Let $\{B_t\}_{t \geq 0}$ be a n-dimensional Brownian motion. We define $\mathcal{F}_t = \mathcal{F}_t^{(n)}$ to be the σ-algebra generated by the collection of random variables

$$\{B_s \mid 0 \leq s \leq t\}$$

We are going to enunciate the definitions and theorems needed to define the Itô integral. We recommend to check a textbook about stochastic analysis for a thorough explanation of the topic. The notations and definitions are taken from chapter 3 of [3] from Øksendal.
The construction of the Itô integral begins with the construction of the Itô integral over elementary processes as a Riemann-Stieltjes integral and then it extends the definition to a bigger class $\mathcal{V}$ of processes. We begin with the 1-dimensional case for the Itô integral.

Definition 1.1.10. Let $(\Omega, \mathscr{F}, P)$ be a probability space, let $0 \leq S < T$, let $\mathcal{V} = \mathcal{V}(S, T)$ be the class of functions

$$f : [0, +\infty) \times \Omega \to \mathbb{R}$$

such that

(i) $(t, \omega) \mapsto f(t, \omega)$ is $\mathscr{B}([0, +\infty)) \times \mathscr{F}$-measurable

(ii) $(t, \omega) \mapsto f(t, \omega)$ is $\mathcal{F}_t$-adapted

(iii) $E[\int_S^T f(t, \omega)^2 dt] < +\infty$

Definition 1.1.11. A function $\phi \in \mathcal{V}$ is called elementary if it has the form

$$\phi(t,\omega) = \sum_j e_j(\omega) \cdot \chi_{[t_j,t_{j+1})}(t) \tag{1.6}$$

Here e_j are functions

$$e_j : \Omega \longrightarrow \mathbb{R}$$

that must be $\mathcal{F}_{t_j}$-measurable since $\phi \in \mathcal{V}$.
Let $B = \{B_t\}_{t\geq 0}$ be a 1-dimensional Brownian motion over Ω, we define the Itô integral (with respect to B) for an elementary function ϕ, with the form written in (1.6), by

$$\int_S^T \phi(t,\omega) dB_t(\omega) = \sum_{j\geq 0} e_j(\omega)[B_{\tilde{t}_{j+1}} - B_{\tilde{t}_j}](\omega) \tag{1.7}$$

where $\tilde{t}_j$ are the points

$$\tilde{t}_j = \begin{cases} t_j & \text{if} \quad S \leq t_j \leq T \\ S & \text{if} \quad t_j < S \\ T & \text{if} \quad t_j > t \end{cases}$$

Definition 1.1.12. (The Itô integral) Let $(\Omega, \mathscr{F}, P)$ be a probability space. Let $0 \leq S < T$. Let $f \in \mathcal{V}(S,T)$. Let $B = \{B_t\}_{t\geq 0}$ be a 1-dimensional Brownian motion over Ω. Then the Itô integral of f from S to T (with respect to B) is defined by

$$\int_S^T f(t,\omega) dB_t(\omega) = \lim_{t\to+\infty} \int_S^T \phi_n(t,\omega) dB_t(\omega) \quad \text{(limit in } L^2(P)) \tag{1.8}$$

where $\{\phi_n\}_{n\in\mathbb{N}}$ is a sequence of elemntary functions such that

$$E\bigg[\int_S^T (f(t,\omega) - \phi_n(t,\omega))^2 dt\bigg] \to 0 \quad \text{as } n \to +\infty \tag{1.9}$$

here the right hand side of (1.8) is defined by (1.7).
Such a sequence $\{\phi_n\}_{n\in\mathbb{N}}$ exists because of Lemma (1.2.4).

We are going to enunciate the definitions needed to define the Itô integral in the n-dimensional case.

Definition 1.1.13 (The n-dimensional Itô integral)**.** Let $(\Omega, \mathscr{F}, P)$ be a probability space, let $\{B_t\}_{t\geq 0} = B = (B^1, B^2, \ldots, B^n)$ be a n-dimensional Brownian motion of components

$$B_t^k : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}, \mathscr{B}(\mathbb{R})) \quad \text{for } k = 1, 2, \ldots, n$$

Then we denote by $\mathcal{V}_{\mathcal{H}}^{m\times n}(S,T)$ the set of matrices $v = [v_{i,j}(t,\omega)]_{i,j=1,\ldots,n}$ where each entry

$$\begin{aligned} v_{i,j} : [0,+\infty) \times \Omega &\longrightarrow \mathbb{R} \\ (t,\omega) &\longmapsto v_{i,j}(t,\omega) \end{aligned}$$

satisfies conditions (i) and (iii) in definition (1.1.10) and satisfies the condition

(ii)' There exists an increasing family of σ-algebras $\{\mathcal{H}_t\}_{t\geq 0}$ such that

a) $\{B_t\}$ is a martingale with respect to $\{\mathcal{H}_t\}$

b) $\{v_{i,j}\}_{t\geq 0}$ is $\{H_t\}$-adapted

It is possible to construct the Itô integral for the functions $f \in \mathcal{V}_{\mathcal{H}}^{m\times n}(S,T)$ in the same way as it is done in (1.1.12).
If $v \in \mathcal{V}_{\mathcal{H}}^{m\times n}(S,T)$ we define, using matrix notation

$$\int_S^T v dB = \int_S^T \begin{pmatrix} v_{1,1} & \cdots & v_{1,n} \\ \vdots & & \vdots \\ v_{n,1} & \cdots & v_{n,n} \end{pmatrix} \begin{pmatrix} dB_1 \\ \vdots \\ dB_n \end{pmatrix}$$

to be the $m \times 1$ matrix whose i-th component is the following sum

$$\sum_{j=i}^{n} \int_S^T v_{i,j}(s,\omega) dB_j(s,\omega)$$

Definition 1.1.14. Under the same notations as the previous definition, $\mathcal{W}_{\mathcal{H}}(S,T)$ denotes the class of processes

$$f : [0,+\infty) \longrightarrow \mathbb{R}$$

satisfying the conditions (i), (ii)' and condition

(iii)' $P\left[\int_S^T f(s,\omega)^2 ds < +\infty\right] = 1$

We also define $\mathcal{W}_{\mathcal{H}} = \bigcap_{T>0} \mathcal{W}_{\mathcal{H}}(0,T)$

The Itô process is the basic example of solution of a stochastic differential equation and is the key element in the definition of the Bellman function.

Definition 1.1.15. Let B_t be a m-dimensional Brownian motion on $(\Omega, \mathscr{F}, P)$. A Itô process (or stochastic integral) is a stochastic process $\{X_t\}_{t\geq 0}$ on $(\Omega, \mathscr{F}, P)$ of the form

$$X_t = X_0 + \int_0^t u(s,\omega) ds + \int_0^t v(s,\omega) dB_s \tag{1.10}$$

here the coefficients

$$u : [0,+\infty) \times \Omega \longrightarrow \mathbb{R}^n; \quad v : [0,+\infty) \times \Omega \longrightarrow \mathbb{R}^{n\times m}$$

have good properties to guarantee that the object (1.10) is well defined, i.e. $v \in \mathcal{W}_{\mathcal{H}}$, so that

$$P\left[\int_0^t v(s,\omega)^2 ds < +\infty \text{ for all } t \geq 0\right] = 1$$

We also assume that u is $\mathcal{H}_t$-adapted and

$$P\left[\int_0^t |u(s,\omega)| ds < +\infty \text{ for all } t \geq 0\right] = 1$$

If $\{X_t\}_{t\geq 0}$ is a Itô process of the form (1.10), the equation (1.10) can be denoted by the differential expression

$$dX_t = udt + vdB_t$$

The Itô diffusion is an example of Itô process where the coefficients of the associated stochastic differential equation don't depend on the time variable. This processes are very important for the proofs in this work because for a process of this kind the Markov property holds.

Definition 1.1.16 (Itô diffusion)**.** A (time-homogeneous) Itô diffusion is a stochastic process

$$\begin{aligned} X : [0, +\infty[\times \Omega &\longrightarrow \mathbb{R}^n \\ (t, \omega) &\longmapsto X_t(\omega) \end{aligned}$$

satisfying a stochastic differential equation of the form

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad t \geq s;\ X_s = x \tag{1.11}$$

where $\{B_t\}_{t\geq 0}$ is a m-dimensional Brownian motion and the coefficients

$$b : \mathbb{R}^n \longrightarrow \mathbb{R}^n; \quad \sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m}$$

satisfy the conditions in Theorem (1.2.6), which in this case simplify to

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq D|x - y| \quad \text{for some } D \in \mathbb{R},\ \forall x, y \in \mathbb{R}^n$$

We denote the unique solution of (1.11) by $X_t = X_t^{s,x}$; $t \geq s$. If $s = 0$ we write X_t^x for $X_t^{0,s}$.

We need a notation to denote the expected value of a Ito diffusion $\{X_t\}_{t\in I}$ at the time t for the theorems about Itô diffusions, like the Markov property.

Definition 1.1.17. Given a Itô diffusion $\{X_t\}_{t\geq 0} = \{X_t^y\}_{t\geq 0}$, for $y \in \mathbb{R}^n$, over the probability space $(\Omega, \mathscr{F}, P)$, solution of the equation

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t; \quad X_0 = y$$

we denote with $\mathcal{M}_\infty$ the σ-algebra (of subsets of Ω) generated by the collection of random variables

$$\{\omega \mapsto X_t^y(\omega) \mid t > 0,\ y \in \mathbb{R}^n\}$$

For each $x \in \mathbb{R}^n$ we define a measure Q^x over the members of $\mathcal{M}_\infty$ by

$$Q^x[X_{t_1} \in E_1, \dots, X_{t_k} \in E_k] = P[X_{t_1}^x \in E_1, \dots, X_{t_k}^x \in E_k]$$

where $E_i \subseteq \mathbb{R}^n$ are Borel sets; $k \in \mathbb{N}$.
Q^x are the probability laws of $\{X_t\}_{t\geq 0}$ for $x \in \mathbb{R}^n$. Q^x gives the distribution of $\{X_t\}_{t\geq 0}$ assuming that $X_0 = x$.
We denote by $E^x[X_t]$ the "expected value of X_t with respect to the measure Q^x", i.e. the expected value the random variable $\omega \mapsto X_t^x(\omega)$ with respect to the measure P, similarly we denote by $E^x[X_t \mid \mathscr{G}]$ the conditional expectation of $\omega \mapsto X_t^x(\omega)$ with respect to the measure P given a σ-algebra $\mathscr{G} \subseteq \mathcal{M}_\infty$.

The infinitesimal generator is a key element to connect the theory of stochastic analysis with the theory of differential problems, allowing for example to solve problems like Dirichlet's problem using the tools from stochastic analysis.

Definition 1.1.18. Let $X = \{X_t\}_{t\geq 0}$ be an Itô diffusion in $\mathbb{R}^n$. We denote by $\mathcal{D}_A(x)$ the set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that it exists the limit

$$\lim_{t\downarrow 0} \frac{E^x[f(X_t)] - f(x)}{t} \tag{1.12}$$

We define the infinitesimal generator of $\{X_t\}_{t\geq 0}$ in x as the operator

$$A : \mathcal{D}_A(x) \to \mathbb{R}$$
$$Af(x) = \lim_{t\downarrow 0} \frac{E^x[f(X_t)] - f(x)}{t}$$

We denote by D_A the set of functions for which the limit (1.12) exists for all $x \in \mathbb{R}^n$.

The exit time of a process $\{X_t\}_{t\in I}$ from a Borel set U is one of the most important examples of exit times, and it is used in important theorems about solving differential problems like the Dirichlet problem using the theory of stochastic analysis.

Definition 1.1.19. Let $U \in \mathbb{R}^n$ be a Borel set, let $X = \{X_t\}_{t\geq 0}$ be a Itô diffusion

$$X_t : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)), \quad t \geq 0$$

We define the exit time for X from the set U as

$$\tau_U : \Omega \longrightarrow \mathbb{R}$$
$$\tau_U(\omega) = \inf\{t > 0 \mid X_t(\omega) \notin U\}$$

The definition of regular point of the boundary of a domain D is a very important definition in the theory about the Dirichlet problem, and there is the analogous version for the theory about stochastic resolution of Dirichlet problems.

Definition 1.1.20. Under the same hypotheses as the previous definition, the point $y \in \partial U$ is called regular for X if

$$Q^y[\tau_U = 0] = 1$$

Otherwise the point y is called irregular.
The boundary set ∂U is called regular for X if all the points $y \in \partial D$ are regular for X.

The Dirichlet-Poisson problem is used to prove the important theorem about the Bellman function being the solution of the HJB equation.

Definition 1.1.21. Let $D \subseteq \mathbb{R}^n$ be a domain, let L denote a semi-elliptic partial differential operator on $C^2(\mathbb{R}^n)$ of the form

$$L = \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i} + \sum_{i,j=1}^{n} a_{i,j}(x)\frac{\partial^2}{\partial x_i \partial x_j}$$

where the functions b_i and $a_{i,j} = a_{j,i}$ are continuous functions.
Let $\phi \in C(\partial D)$ and $g \in C(D)$ be given functions. A function $w \in C^2(D)$ is called a solution of the Dirichlet-Poisson problem (over D, associated to L, ϕ, g) if

(I) $Lw = -g \quad \text{in } D$

(II) $\lim_{\substack{x\to y\\ x\in D}} w(x) = \phi(y) \quad \text{for all } y \in \partial D$

1.2 Theorems

The proofs of these theorems can be found in [3] from Øksendal, in chapters 1,2,3,4,5,7 and 9. We will enunciate the theorems needed for the construction of the Bellman function and to prove the theorem about the HJB equation.

Kolmogorov's extension theorem is one of the fondamental results in the theory of stochastic processes, and it allows to prove the existence of stochastic processes having given finite-dimensional distributions, like the Brownian motion.

Theorem 1.2.1 (Kolmogorov's extension theorem)**.** *Let T be a set of times. For all $k \in \mathbb{N}$, $t_1, \ldots, t_k \in T$, let $\nu_{t_1,\ldots,t_k}$ be probability measures on $\mathbb{R}^{nk}$ such that, for all $F_1, \ldots, F_k$ Borel subsets of $\mathbb{R}^n$*

$$\nu_{t_{\sigma(1)},\ldots,t_{\sigma(k)}}(F_1 \times \cdots \times F_k) = \nu_{t_1,\ldots,t_k}(F_{\sigma^{-1}(1)} \times \cdots \times F_{\sigma^{-1}(k)})$$

for all permutations σ on $\{1, 2, \ldots, k\}$ and

$$\nu_{t_1,\ldots,t_k}(F_1 \times \cdots \times F_k) = \nu_{t_1,\ldots,t_k,t_{k+1}}(F_1 \times \cdots \times F_k \times \mathbb{R}^n)$$

Then there exists a complete probability space $(\Omega, \mathscr{F}, P)$ and a stochastic process $\{X_t\}_{t\in T}$

$$X_t : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$$

such that

$$\nu_{t_1,\ldots,t_k}(F_1, \times, F_k) = P[X_{t_1} \in F_1, \ldots, X_{t_k} \in F_k]$$

for all $t_i \in T$, $k \in \mathbb{N}$, and for all F_i Borel subsets of $\mathbb{R}^n$.

Kolmogorov's continuity theorem is another fondamental result in the theory of stochastic processes, and it is used to prove that the Brownian motion can be considered a continuous process.

Theorem 1.2.2 (Kolmogorov's continuity theorem)**.** *Let $X = \{X_t\}_{t\geq 0}$ be a stochastic process such that for all $T > 0$ there exist positive constants α, β, D such that*

$$E[|X_t - X_s|^\alpha] \leq D \cdot |t - s|^{1+\beta}; \quad \text{for } 0 \leq s, t \leq T \tag{1.13}$$

Then there exists a continuous version of X.

The Itô isometry is one of the most important results in the theory of stochastic differential equations and it's one of the key elements used in the construction of the Itô integral. The Itô isometry for elementary functions is used to define the Itô integral, and then using the Itô integral we can extend the Itô isometry to all the Itô integrable processes.

Lemma 1.2.3 (Itô isometry for elementary functions)**.** *Let $(\Omega, \mathscr{F}, P)$ be a probability space. If*

$$\phi : [0, +\infty) \times \Omega \to \mathbb{R}^n$$

$$\phi(t, \omega) = \sum_j e_j(\omega) \cdot \chi_{[t_j,t_j+1)}(t)$$

is a bounded elementary function, then

$$E\Big[\Big(\int_S^T \phi(t,\omega)dB_t(\omega)\Big)^2\Big] = E\Big[\int_S^T \phi(t,\omega)^2 dt\Big] \tag{1.14}$$

where

$$\int_S^T \phi(t,\omega)dB_t(\omega) = \sum_{j\geq 0} e_j(\omega)[B_{t_{j+1}} - B_{t_j}](\omega)$$

This Lemma proves the three statements needed in the construction of the Itô integral.

Lemma 1.2.4. *The following three statements hold true:*

1. *Let $g \in \mathcal{V}$ be a bounded and such that $t \mapsto g(t,\omega)$ is continuous for each $\omega \in \Omega$. Then there exist a sequence of elementary functions $\phi_n \in \mathcal{V}$ such that*

$$E\Big[\int_S^T (g-\phi_n)^2 dt\Big] \to 0 \quad \textit{for } n \to +\infty$$

2. *Let $h \in \mathcal{V}$ be bounded. Then there exist a sequence of bounded functions $g_n \in \mathcal{V}$ such that $t \mapsto g_n(t,\omega)$ is continuous for all $\omega \in \Omega$ and for all n, and*

$$E\Big[\int_S^T (h-g_n)^2 dt\Big] \to 0 \quad \textit{for } n \to +\infty$$

3. *Let $f \in \mathcal{V}$. Then there exist a sequence of functions $h_n \in \mathcal{V}$ such that h_n is bounded for each n and*

$$E\Big[\int_S^T (f-h_n)^2 dt\Big] \to 0 \quad \textit{for } n \to +\infty$$

Theorem 1.2.5 (Itô isometry)**.**

$$E\Big[\Big(\int_S^T f(t,\omega)dB_t(\omega)\Big)^2\Big] = E\Big[\int_S^T f(t,\omega)^2 dt\Big] \quad \textit{for all } f \in \mathcal{V}(S,T) \tag{1.15}$$

The following theorem allows us to prove the existence and uniqueness of solutions for stochastic differential equations, which is needed to guarantee that the Bellman function is well defined.

Theorem 1.2.6 (Existence and uniqueness theorem for solutions of stochastic differential equations)**.** *Given $T > 0$, let*

$$b : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
$$\sigma : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m}$$

be measurable functions. Suppose that it exists a constant $C > 0$ such that

$$|b(t,x)| + |\sigma(t,x)| \leq C(1+|x|); \quad \forall x \in \mathbb{R}^n,\ \forall t \in [0,T]$$

Suppose that it exists a constant $D > 0$ such that

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x-y|; \quad \forall x,y \in \mathbb{R}^n,\ \forall t \in [0,T]$$

Let $\{B_t\}_{t\geq 0}$ be a m-dimensional Brownian motion, let Z be a random variable wich is independent of the σ-algebra $\mathcal{F}_\infty^{(m)}$ generated by the collection of random variables $\{B_s(\cdot) \mid s \geq 0\}$, and such that

$$E[|Z|^2] < +\infty$$

Then the stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad 0 \leq t \leq T,\ X_0 = Z \tag{1.16}$$

has a unique t-continuous solution $\{X_t\}_{0\leq t\leq T}$ with the property that:

1. *$\{X_t\}_{0\leq t\leq T}$ is adapted to the filtration $\{\mathcal{F}_t^Z\}_{0\leq t\leq T}$, where $\mathcal{F}_t^Z$ is the σ-algebra generated by the collection of random variables $\{Z, B_s(\cdot) \mid 0 \leq s \leq t\}$.*

2. $$E\left[\int_0^T |X_t|^2 dt\right] < +\infty$$

The strong Markov property is the most important result for the theory of the Bellman functions and it allows to prove important propositions like the Bellman principle and the theorem about the HJB equation. The Markov property basically states that what happens to an Itô diffusion $\{X_t\}_{t\in I}$ after a time t only depends on X_t and doesn't depend on X_s for $s < t$.

Theorem 1.2.7 (The strong Markov property for Itô diffusions)**.** *Let $\{X_t\}_{t\geq 0}$ be a Itô diffusion in $\mathbb{R}^n$. Let f be a bounded Borel function $f : \mathbb{R}^n \to \mathbb{R}$, let $\{B_t\}_{t\geq 0}$ be a m-dimensional Brownian motion, let τ be a stopping time with respect to the σ-algebra $\mathcal{F}_t^{(m)}$ generated by $\{B_t\}_{t\geq 0}$, suppose $\tau < +\infty$ almost surely. Then*

$$E^x[f(X_{\tau+h}) \mid \mathcal{F}_\tau^{(m)}] = E^{X_\tau}[f(X_h)] \quad \forall h \geq 0$$

The following theorem is very important for the work and for the general theory. It characterizes the infinitesimal generator of an Itô diffusion

Theorem 1.2.8. *Let X_t be the Itô diffusion*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

If $f \in C_0^2(\mathbb{R}^n)$ then $f \in \mathcal{D}_A$ and

$$Af(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} \tag{1.17}$$

The following lemma is used to prove Dynkin's formula, a result which is a very important result to understand the evolution of the composition between a smooth function and a stochastic process.

Lemma 1.2.9. *Let $\{X_t\}_{t\geq 0} = \{X_t^x\}_{t\geq 0}$ be an Itô diffusion in $\mathbb{R}^n$ of the form*

$$X_t^x(\omega) = x + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s(\omega)$$

where $\{B\}_{t\geq 0}$ is a m-dimentional Brownian motion. Let $f \in C_0^2(\mathbb{R}^n)$, let τ be a stopping time with respect to the filtration $\{\mathcal{F}_t^{(m)}\}$, and assume that $E^x[\tau] < +\infty$.

Assume that $u(t,\omega)$ and $v(t,\omega)$ are bounded on the set of (t,ω) such that $X(t,\omega)$ belongs to the support of f. Then

$$E[f(X_t)] = f(x)+E^x\left[\int_0^\tau\left(\sum_i u_i(s,\omega)\frac{\partial f}{\partial x_i}(X_s)+\frac{1}{2}\sum_{i,j}(vv^T)_{i,j}(s,\omega)\frac{\partial^2 f}{\partial x_i\partial x_j}(X_s)\right)ds\right]$$

Theorem 1.2.10 (Dynkin's formula)**.** *Under the same assumptions of Lemma (1.2.9) it follows that*

$$E^x[f(X_\tau)] = f(x) + E^x\left[\int_0^\tau Af(X_s)dx\right] \tag{1.18}$$

The following lemma is used in the proof of the theorem about the HJB equation and it allows to calculate the time shift of a process stopped on an exit time from a Borel set.

Lemma 1.2.11. *Let $H \subseteq \mathbb{R}^n$ be measurable, let $X = \{X_t\}_{t\geq 0}$ be a Itô diffusion in $\mathbb{R}^n$*

$$X_t : (\Omega, \mathscr{F}, P) \longrightarrow (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$$

Let τ_H be the first exit time from H for X. Let α be another stopping time, g be a bounded continuous function on $\mathbb{R}^n$. Let θ_t be the shift operator

$$\theta_t : \mathcal{H} \longrightarrow \mathcal{H}$$

defined by, given $\nu = g_1(X_{t_1})\ldots g_k(X_{t_k})$, the expression

$$\theta_t\nu = g_1(X_{t_1+t})\ldots g_k(X_{t_k+t})$$

extended over all functions in $\mathcal{H}$ by taking limits of sums of such functions. Consider

$$\eta = g(X_{\tau_H})\cdot\chi_{\{\tau_H<+\infty\}}$$

$$\tau_H^\alpha(\omega) = \inf\{t > \alpha \mid X_t(\omega) \notin H\}, \quad \omega \in \Omega$$

then

$$\theta_\alpha\eta\cdot\chi_{\{\alpha<+\infty\}} = g(X_{\tau_H^\alpha})\cdot\chi_{\{\tau_H^\alpha<+\infty\}}$$

where

$$(\theta_\alpha\eta)(s,\omega) = \eta(s+\alpha,\omega)$$

The Dirichlet-Poisson problem is used in the theorem about the HJB equation to prove that the Bellman function is solution of the HJB equation with boundary values equal to the value of the bequest function.

Theorem 1.2.12. *Let $D \subseteq \mathbb{R}^n$ be a domain. Let $X = \{X_t\}_{t\geq 0}$ be a Itô diffusion in $\mathbb{R}^n$. Let $\mathcal{A}$ be the infinitesimal generator of X. Let Q^x be the probability law of X starting at $X_0 = x$, for $x \in \mathbb{R}^n$. Let τ_D be the stopping time*

$$\tau : \Omega \longrightarrow \mathbb{R}$$

$$\tau(\omega) = \inf\{t > 0 \mid X_t(\omega) \notin D\}$$

Suppose that $\tau_D < +\infty$ *almost surely with respect to* Q^x *for all* $x \in D$. *Let* $\phi \in C(\partial D)$ *be bounded and let* $g \in C(D)$ *satisfy*

$$E^x\left[\int_0^{\tau_D} |g(X_s)|ds\right] < +\infty \quad \textit{for all } x \in D \tag{1.19}$$

Define

$$w(x) = E^x[\phi(X_{\tau_D})] + E^x\left[\int_0^{\tau_D} g(X_s)ds\right], \quad x \in D \tag{1.20}$$

Then the following two statements hold true

a)

$$\mathcal{A}w = -g \quad \textit{in } D \tag{1.21}$$

and

$$\lim_{t\uparrow\tau_D} w(X_t) = \phi(X_{\tau_D}) \tag{1.22}$$

almost surely with respect to Q^x, *for all* $x \in D$

b) If there exists a function $w_1 \in C^2(D)$ *and a constant* C *such that*

$$|w_1(x)| < C\left(1 + E^x\left[\int_0^{\tau_D} |g(X_s)|ds\right]\right), \quad x \in D \tag{1.23}$$

and w_1 *satisfies (1.21) and (1.22), then* $w_1 = w$.

Chapter 2

Bellman functions in stochastic control

2.1 Bellman functions

Let $\{X_t\}_{t\geq 0}$ be an Itô process described by the stochastic differential equation

$$dX_t = dX_t^u = b(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dB_t \tag{2.1}$$

where $X_t(\omega) \in \mathbb{R}^n$, and the coefficients b and σ are

$$b : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}^n, \quad \sigma : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$$

and B_t is an m-dimensional Brownian motion. Let. Here U is a given Borel set $U \subseteq \mathbb{R}^k$ and $\{u_t\}_{t\geq 0}$ is the control, i.e. a stochastic process whose values are $u_t(\omega) \in U$ that is adapted to the σ-algebra $\mathcal{F}_t^{(m)}$, i.e. for all $t \geq 0$ the random variable u_t is measurable with respect to the σ-algebra $\mathcal{F}_t^{(m)}$.
Let $\{X_h^{s,x}\}_{h\geq s}$ the solution of (2.1) such that $X_s^{s,x} = x$, i.e.

$$X_h^{s,x} = x + \int_s^h b(r, X_r^{s,x}, u_r)dr + \int_s^h \sigma(r, X_r^{s,x}, u_r)dB_r; \quad h \geq s \tag{2.2}$$

Let the probability law of X_t starting at x for $t = s$ be denoted by $Q^{s,x}$, i.e.

$$Q^{s,x}[X_{t_1} \in F_1, \dots, X_{t_k} \in F_k] = P[X_{t_1}^{s,x} \in F_1, \dots, X_{t_k}^{s,x} \in F_k] \tag{2.3}$$

for all $s \leq t_i$, F_i measurable subset of R^n; for all $1 \leq i \leq k$, $k = 1, 2, \dots$

Let F and K be two continuos functions

$$F : \mathbb{R} \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}, \quad K : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

here F is the "utility rate" function, and K is the "bequest" function.
Let G be a fixed domain in $\mathbb{R} \times \mathbb{R}^n$ and let $\widehat{T}$ be the first exit time after s from G for the process $\{X_r^{s,x}\}_{r\geq s}$, i.e.

$$\widehat{T} = \widehat{T}^{s,x}(\omega) = \inf\{r > s \mid (r, X_r^{s,x}(\omega)) \in G\} \leq +\infty \tag{2.4}$$

Let $F^u(r,z) = F(r,z,u)$. Suppose that

$$E^{s,x}\left[\int_s^{\widehat{T}} |F^{u_r}(r,X_r)|dr + |K(\widehat{T},X_{\widehat{T}})| \cdot \chi_{\{\widehat{T}<+\infty\}}\right] < +\infty \quad \text{for all } s,x,u \tag{2.5}$$

We define the performance function $J^u(s,x)$ by

$$J^u(s,x) = E^{s,x}\left[\int_s^{\widehat{T}} F^{u_r}(r,X_r)dr + K(\widehat{T},X_{\widehat{T}}) \cdot \chi_{\{\widehat{T}<+\infty\}}\right] \tag{2.6}$$

In order to get a simpler notation we define

$$Y_t = (s+t, X_{s+t}^{s.x}) \in \mathbb{R}^{n+1} \qquad \text{for } t \geq 0,\ Y_0 = (s,x) \tag{2.7}$$

and we substitute Y_t in (2.1) to get the equation

$$dY_t = dY_t^u = b(Y_t.u_t)dt + \sigma(Y_t,u_t)dB_t \tag{2.8}$$

We denote by $Q^{s,x} = Q^y$ the probability of Y_t starting at $y=(s,x)$ for $t=0$. We observe that

$$\int_s^{\widehat{T}} F^{u_r}(r,X_r)dr = \int_0^{\widehat{T}-s} F^{u_{s+t}}(s+t,X_{s+t})dt = \int_s^T F^{u_{s+t}}(Y_t)dt$$

where

$$T := \inf\{t>0 \mid Y_t \notin G\} = \widehat{T} - s \tag{2.9}$$

We also observe that

$$K(\widehat{T},X_{\widehat{T}}) = K(Y_{\widehat{T}-s}) = K(Y_T)$$

so the performance function may be written in terms of Y as follows, with y=(s,x),

$$J^u(y) = E^y\left[\int_0^T F^{u_t}(Y_t)dt + K(Y_T) \cdot \chi_{\{T<+\infty\}}\right] \tag{2.10}$$

here u_t is a time shift of the u_t in (2.8).

Definition 2.1.1. Given a Borel set $U \subseteq \mathbb{R}^{n+1}$, given two continuous functions

$$F : \mathbb{R}\times\mathbb{R}^n\times U \longrightarrow \mathbb{R}, \quad K : \mathbb{R}\times\mathbb{R}^n \longrightarrow \mathbb{R}$$

and given the stochastic differential equation (2.1)

$$dX_t = dX_t^u = b(t,X_t,u_t)dt + \sigma(t,X_t,u_t)dB_t$$

associated to the coefficients

$$b : \mathbb{R}\times\mathbb{R}^n\times U \to \mathbb{R}^n, \quad \sigma : \mathbb{R}\times\mathbb{R}^n\times U \to \mathbb{R}^{n\times m}$$

we denote with Bellman function associated to the equation (2.1), to the functions F and K, over a set of admissible controls $\mathcal{C}$, a function

$$\mathcal{B} : U \longrightarrow \mathbb{R}$$

$$\mathcal{B}(y) = \sup_{\{u_t\}_{t\geq 0}\in\mathcal{C}} J^u(y)$$

here J^u is the performance function defined in (2.10), and the supremum is taken over the set $\mathcal{C}$ of admissible controls. Here $\mathcal{C}$ is a set of controls $\{u_t\}_{t\geq 0}$ that are $\mathcal{F}_t^{(m)}$-adapted, whose values are $u_t(\omega) \in U$.
If a control $\{u_t^*\}_{t\geq 0}$ such that

$$\mathcal{B}(y) = \sup_{\{u_t\}_{t\geq 0}\in\mathcal{C}} J^u(y) = J^{u^*}(y)$$

exists then $\{u_t^*\}_{t\geq 0}$ is called optimal control.

We may take into consideration different types of control functions. The set of control functions that we will look into is the set of Markov controls, which is the set $\mathcal{C}$ defined by the set of the processes

$$\mathcal{C} := \left\{ u(t,\omega) = u_0(t, X_t(\omega)) \;\mid\; \text{for } u_0 : \mathbb{R}^{n+1} \to U, \quad u_0 \text{ measurable} \right\} \tag{2.11}$$

2.2 The Hamilton-Jacobi-Bellman Equation

According to the definitions in the previous section, we consider the set $\mathcal{C}$ of Markov controls

$$u(t,\omega) = u_0(t, X_t(\omega))$$

defined in (2.11), and, after introducing $Y_t = (s+t, X_{s+t})$ as explained in (2.7), the system equation becomes

$$dY_t = b(Y_t, u_0(Y_t))dt + \sigma(Y_t, u_0(Y_t))dB_t \tag{2.12}$$

For every $v \in U$ and $f \in C_0^2(\mathbb{R}\times\mathbb{R}^n)$ we define the operator

$$(L^v f)(y) = \frac{\partial f}{\partial s}(y) + \sum_{i=1}^n b_i(y,v)\frac{\partial f}{\partial x_i}(y) + \sum_{i,j=1}^n a_{i,j}(y,v)\frac{\partial^2 f}{\partial x_i \partial x_j}(y), \quad \forall y \in \mathbb{R}\times\mathbb{R}^n \tag{2.13}$$

here $a_{i,j} = \frac{1}{2}(\sigma\sigma^T)_{i,j}$, $y = (s,x)$ and $x = (x_1, \dots, x_n)$. Then, by Theorem (1.2.8), for each choice of the function u_0 (that defines the control u), the solution $Y_t = Y_t^u$ is an Itô diffusion with infinitesimal generator A given by

$$(Af)(y) = (L^{u_0(y)}f)(y) \quad \text{for } f \in C_0^2(\mathbb{R}\times\mathbb{R}^n),\ y \in G$$

For every $v \in U$ define $F^v(y) = F(y,v)$. The first fondamental result in stochastic control theory is the following:

Theorem 2.2.1 (The Hamilton-Jacobi-Bellman (HJB) equation (I))**.** *Under the notations of the previous section, consider the Bellman function*

$$\mathcal{B}(y) = \sup\{J^u(y) \mid u = u_0(Y) \ \ \textit{Markov control}\ \}$$

Suppose that $\mathcal{B}$ satisfies

$$E^y\left[|\mathcal{B}(Y_\alpha)| + \int_0^\alpha |L^v\mathcal{B}(Y_t)|dt\right] < +\infty$$

for all bounded stopping times $\alpha < T$, for all $y \in G$ and for all $v \in U$. Suppose that the stopping time T is $T < +\infty$ almost surely with respect to Q^y for all $y \in G$, and *suppose that a optimal Markov control $u^* = u_0^*(Y)$ exists. Suppose ∂G is regular for Y^{u^*}. Then*

$$\sup_{v \in U}\{F^v(y) + (L^v \mathcal{B})(y)\} = 0 \quad \text{for all } y \in G \tag{2.14}$$

and

$$\mathcal{B}(y) = K(y) \quad \text{for all } y \in \partial G \tag{2.15}$$

The supremum in (2.14) is obtained if $v = u_0^*(y)$, where $u^* = u_0^*(Y_t)$ is an optimal control. In other words

$$F(y, u_0^*(y)) + (L^{u_0^*(y)}\mathcal{B})(y) = 0 \quad \text{for all } y \in G \tag{2.16}$$

Proof. We will begin the proof by proving the statements (2.15) and (2.16). Since u^* is optimal we have

$$\mathcal{B}(y) = J^{u^*}(y) = E^y\Big[\int_0^T F(Y_s, u^*(Y_s))ds + K(Y_t)\Big]$$

If $y \in \partial G$ then $T = 0$ almost surely with respect to Q^y (since ∂G is regular) and (2.15) follows. By Theorem (1.2.12) the solution of the Dirichlet-Poisson problem we get

$$(L^{u^*}(y)\mathcal{B})(y) = -F(y, u^*(y)) \quad \text{for all } y \in G$$

which is (2.16). Now we will prove (2.14).
Choose U open, $y = (s, x) \in U$, $U \subset\subset G$. Define

$$\eta = \int_0^{\tau_G} g(Y_s)ds$$

and $\tau = \tau_U$.
Now, since the stopping time T is actually the exit time τ_G, apply Lemma (1.2.11) to η to get

$$\theta_\tau \eta = \int_\tau^{\tau_G} g(X_s)ds \tag{2.17}$$

almost surely with respect to Q^y because $T < +\infty$ almost surely with respect to Q^y for hypothesis.
Fix $y = (s, x) \in G$ and choose a Markov control $u = u_0(Y)$. Let $\alpha \le T$ be a stopping time. Since

$$J^u(y) = E^y\Big[\int_0^T F^u(Y_r)dr + K(Y_t)\Big]$$

we get by the strong Markov property, combined with Lemma (1.2.11), that

$$\begin{aligned}
E^y[J^u(Y_\alpha)] =& E^y\Big[E^{Y_\alpha}\Big[\int_0^T F^u(Y_r)dr + K(Y_T)\Big]\Big] = \\
=& E^y\Big[E^y\Big[\theta_\alpha\Big(\int_0^T F^u(Y_r)dr + K(Y_T)\Big) \Big| \mathcal{F}_\alpha\Big]\Big] = \\
=& E^y\Big[E^y\Big[\int_\alpha^T F^u(Y_r)dr + K(Y_T) \Big| \mathcal{F}_\alpha\Big]\Big] = \\
=& E^y\Big[\int_0^T F^u(Y_r)dr + K(Y_T) - \int_0^\alpha F^u(Y_r)dr\Big] = \\
=& J^u(y) - E^y\Big[\int_0^\alpha F^u(Y_r)dr\Big]
\end{aligned}$$

So we get

$$J^u(y) = E^y\left[\int_0^\alpha F^u(Y_r)dr\right] + E^y[J^u(Y_\alpha)] \tag{2.18}$$

Now let $W \subseteq Q$ be of the form $W = \{(r,z) \in G \mid r < t_1\}$, where $s < t_1$. Put $\alpha = \inf\{t \geq 0 \mid Y_t \notin W\}$. Suppose an optimal control $u^*(y) = u^*(r,z)$ exists and choose

$$u(r,z) = \begin{cases} v & \text{if } (r,z) \in W \\ u^*(r,z) & \text{if } (r,z) \notin W \end{cases}$$

where $v \in U$ is chosen arbitrarily. Then

$$\mathcal{B}(Y_\alpha) = J^{u^*}(Y_\alpha) = J^u(Y_\alpha) \tag{2.19}$$

So, by combining (2.18) and (2.19) we get

$$\mathcal{B}(y) \geq J_u(y) = E^y\left[\int_0^\alpha F^v(Y_r)dr\right] + E^y[\mathcal{B}(Y_\alpha)] \tag{2.20}$$

Since $\mathcal{B} \in C^2(G)$ we get by Dynkin's formula

$$E^y[\mathcal{B}(Y_\alpha)] = \mathcal{B}(y) + E^y\left[\int_0^\alpha (L^u\mathcal{B})(Y_r)dr\right]$$

which substituted in (2.20) gives

$$\mathcal{B}(y) \geq E^y\left[\int_0^\alpha F^v(Y_r)dr\right] + \mathcal{B}(y) + E^y\left[\int_0^\alpha (L^u\mathcal{B})(Y_r)dr\right]$$

that is equivalent to

$$E^y\left[\int_0^\alpha (F^v(Y_r) + (L^v\mathcal{B})(Y_r))dr\right] \leq 0 \quad \text{for all such W}$$

So, by letting $t_1 \downarrow s$ we obtain, since $F^v(\cdot)$ and $(L^v\mathcal{B})(\cdot)$ are continuous at y, that $F^v(y) + (L^v\mathcal{B})(y) \leq 0$, which combined with (2.16) gives (2.14). □

Theorem 2.2.2 (A converse of the HJB equation (I))**.** *Let ϕ be a function in $C^2(G) \cap C(\overline{G})$ such that, for all $v \in U$,*

$$F^v(y) + (L^v\phi)(y) \leq 0; \quad y \in G \tag{2.21}$$

with boundary values

$$\lim_{t\to T} \phi(Y_t) = K(Y_T) \cdot \chi_{\{T<+\infty\}} \tag{2.22}$$

almost surely with respect to Q^y and such that

$$\{\phi(Y_\tau)\}_{\tau \leq T} \quad \textit{is uniformly } Q^y\textit{-integrable} \tag{2.23}$$

for all Markov controls u and all $y \in G$. Then

$$\phi(y) \geq J^u(y) \quad \textit{for all Markov controls } u \textit{ and all } y \in G. \tag{2.24}$$

Moreover, if for each $y \in G$ we have found $u_0^*(y)$ such that

$$F^{u_0^*(y)}(y) + (L^{u_0^*(y)}\phi)(y) = 0 \tag{2.25}$$

then $u_0 = u_0^*(Y)$ is a Markov control such that

$$\phi(y) = J^{u_0}(y)$$

and hence u_0 must be a optimal control and $\phi(y) = \mathcal{B}(y)$.

Proof. Let $\phi \in C^2(G) \cap C(\overline{G})$ such that it satisfies (2.21) and (2.22). Let u be a Markov control. Since $L^u\phi \leq -F^u$ in G we have by Dynkin's formula

$$\begin{aligned} E^y[\phi(T_{T_R})] =& \phi(y) + E^y\left[\int_0^{T_R} (L^u\phi)(Y_r)dr\right] \\ &\leq \phi(y) - E^y\left[\int_0^{T_R} F^u(Y_r)dr\right] \end{aligned}$$

where

$$T_R = \min\{R, T, \inf\{t > 0 \mid |Y_t| \geq R\}\} \tag{2.26}$$

for all $R < +\infty$. This gives, by (2.1), (2.21) and (2.22)

$$\begin{aligned} \phi(y) \geq E^y&\left[\int_0^{T_R} F^u(Y_r)dr + \phi(Y_{T_R})\right] \\ \to E^y&\left[\int_0^{T_R} F^u(Y_r)dr + K(Y_T) \cdot \chi_{\{T<+\infty\}}\right] = J^u(y) \end{aligned}$$

as $R \to +\infty$, which proves (2.24). If u_0 is such that (2.24) holds, then the calculations above give equality, completing the proof. □

The last result that we are going to mention is that, under suitable conditions on $b, \sigma, F, \partial G$ and assuming that the set of control values is compact, it is possible to show that is exists a smooth function ϕ such that

$$\sup_v \{F^v(y) + (L^v\phi)(y)\} = 0 \quad \text{for } y \in G$$

and

$$\phi(y) = K(y) \quad y \in \partial G$$

Moreover, by a measurable selection theorem one can find a measurable function u_0 that defines a Markov control $u_t^*(\omega) = u_0(X_t(\omega))$ such that

$$F^{u_0(y)}(y) + (L^{u_0(y)}\phi)(y) = 0$$

for almost all $y \in G$ with respect to Lebesgue measure in $\mathbb{R}^{n+1}$, and that the solution $X_t = X_t^{u^*}$ exists. For details see Øksendal [3, pg. 241].
Moreover, it is always possible to get as good as a performance with Markov controls as with arbitrary $\mathcal{F}_t^{(m)}$-adapted controls as long as some extra conditions are satisfied, as stated in the next theorem.

Theorem 2.2.3. *Let*

$$\Phi_M(y) = \sup\{J^u(y) \mid u = u_0(Y) \text{ Markov control}\}$$

and

$$\Phi_a(y) = \sup\{J^u(y) \mid u = u(t, \omega)\ \mathcal{F}_t^{(m)}\text{-adapted control}\}$$

Suppose there exists an optimal Markov control $u^* = u_0(Y)$ for the Markov control problem

$$\Phi_M(y) = J^{u_0}(y) \quad \text{for all } y \in G$$

such that all the boundary points of G are regular with respect to $Y_t^{u^*}$ and that Φ_M is a function in $C^2(G) \cap C(\overline{G})$ satisfying

$$E^y\left[|\Phi_M(Y_\alpha)| + \int_0^\alpha |L^u \Phi_M(Y_t)|dt\right] < +\infty$$

for all bounded stopping times $\alpha \leq T$, all adapted controls u and all $y \in G$. Then

$$\Phi(M)(y) = \Phi_a(y) \quad \textit{for all } y \in G$$

For the proof of this theorem see Øksendal [3, pg. 232-233].
So we will use both the general $\mathcal{F}_t^{(m)}$-adapted controls and the Markov controls to solve the problems in the next chapter.

Chapter 3

Bellman function's method

We will state again the Bellman function problem in this chapter to adapt it to a simpler case (where there are no restriction over the time values) and we will use an example to illustrate the Bellman function's method to prove theorems over dyadic trees.

3.1 Bellman equation for stochastic optimal control

3.1.1 Bellman function

Let $X = \{X_t\}_{t\geq 0}$ be a stochastic process in $\mathbb{R}^d$ satisfying the following integral equation

$$X_t = x + \int_0^t \sigma(\alpha_s, X_s)dW_s + \int_0^t b(\alpha_s, X_s)ds \tag{3.1}$$

where $x \in \mathbb{R}^d$ is the starting point, $t \in [0, +\infty[$ is the time, $W = \{W_t\}_{t\geq 0}$ is a d_1-dimensional Brownian motion

$$W_t : (\Omega, \mathscr{F}) \longrightarrow (\mathbb{R}^{d_1}, \mathscr{B}(\mathbb{R}^{d_1}))$$

$\alpha = \{\alpha_t\}_{t\geq 0}$ is a d-dimentional stochastic process

$$\alpha_t : (\Omega, \mathscr{F}) \rightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$$

σ is a measurable function that runs its values in the matrix space $\mathbb{R}^d \times \mathbb{R}^{d_1}$

$$\sigma : (\mathbb{R}^d \times \mathbb{R}^d, \mathscr{B}(\mathbb{R}^d \times \mathbb{R}^d)) \longrightarrow (\mathbb{R}^d \times \mathbb{R}^{d_1}, \mathscr{B}(\mathbb{R}^d \times \mathbb{R}^{d_1}))$$

b is a measurable function that runs its values in the vector space $\mathbb{R}^d$

$$b : (\mathbb{R}^d \times \mathbb{R}^d, \mathscr{B}(\mathbb{R}^d \times \mathbb{R}^d)) \longrightarrow (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$$

The process α is supposed to be a control of our choice. We denote by $A \subseteq \mathbb{R}^d$ the set of admissible values, which is the set where the vector of control parameters α is allowed to run. We denote by admissible control a control α such as $\alpha_t(\omega) \in A$ for all $t \in [0, +\infty]$, for all $\omega \in \Omega$ and such that α_t is measurable with respect to the σ-algebra generated by the random variables x_s for $0 \leq s \leq t$. Since α_t is measurable with respect to the variables x_s for $0 \leq s \leq t$, we will also denote α by

$\alpha = \{\alpha_t(X_{[0,t]})\}_{t\geq 0}$.
We also denote by profit density function associated to α a measurable function f

$$f : (A \times \mathbb{R}^d, \mathscr{B}(\mathbb{R}^d \times \mathbb{R}^d)) \longrightarrow (\mathbb{R}, \mathscr{B}(\mathbb{R}))$$

$$(a, x) \longmapsto f(a, x) =: f^a(x)$$

and the bonus function as a measurable function F

$$F : (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d)) \longrightarrow (\mathbb{R}, \mathscr{B}(\mathbb{R}))$$

Here f represents the density function of the profit gain over a trajectory, while F represents a bonus function that expresses the profit gain at retirement. Given a trajectory

$$\gamma : [0, t] \longrightarrow \mathbb{R}^d$$

$$s \longmapsto X_s$$

we define the profit on the trajectory γ as the random variable

$$\Omega \ni \omega \longmapsto \int_0^t f^{\alpha_s(\omega)}(X_s(\omega))ds \in \mathbb{R} \tag{3.2}$$

here X is the stochastic process satisfying (1).
The goal of the problem is to choose the control $\alpha = \{\alpha_s(X_{[0,t]})\}_{t\geq 0}$ to maximize the average profit

$$v^\alpha(x) = E^x \int_0^{+\infty} f^{\alpha_t}(X_t)dt + \lim_{t\to+\infty} E^x(F(X_t)) \tag{3.3}$$

here $v^\alpha(x)$ is the average profit gain over the random trajectory $t \mapsto X_t$ starting from the point $x \in \mathbb{R}^d$.
We define the Bellman function for stochastic control as a function

$$v : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$x \longmapsto \sup_\alpha v^\alpha(x)$$

here the supremum is taken over all admissible controls α.

3.1.2 Bellman's principle and Bellman's equation

A Bellman function satisfies an equation known as Bellman's principle. The Bellman's principle states that

$$v(x) = \sup_\alpha E^x \left[\int_0^t f^{\alpha_s}(X_s)ds + v(X_t) \right] \tag{3.4}$$

for each $t > 0$. Here the supremum is taken over all admissible control processes α.

Proof. To prove this let α be a Markov admissible control chosen arbitrarily. We have mentioned that it is sufficient to choose Markov controls to get the same Bellman function you would get by choosing any admissible control. Let $t > 0$ be a fixed time. We denote with X^y where $y \in \mathbb{R}^d$ a stochastic process satisfying the equation

$$X_t^y = y + \int_0^t \sigma(\alpha_s, X_s^y)dW_s + \int_0^t b(\alpha_s, X_s^y)ds$$

Let $\mathcal{F}_t^{(m)}$ be the σ-algebra generated by the variables $\{W_s \mid 0 \leq s \leq t\}$.
We will use a more precise notation to allow us to manage properly all the aleatory variables in the following equations. We denote by $\mathbb{E}^{\omega}\big[g(\omega)\big]$ the expected value of the aleatory variable $\omega \longrightarrow g(\omega)$. For example, the thesis of the Bellman principle can be rewritten (using this new notation) as

$$v(x) = \sup_{\alpha} \mathbb{E}^{\omega_1}\left[\int_0^t f^{\alpha_s}(X_s^x(\omega_1))ds + v(X_t^x)\right]$$

The following equations prove the theorem:

$$v(x) = \sup_{\alpha}\left\{\mathbb{E}^{\omega_1}\left[\int_0^{+\infty} f^{\alpha}(X_s^x(\omega_1))ds\right] + \lim_{s\to+\infty}\mathbb{E}^{\omega_3}(F(X_s^x(\omega_3)))\right\} =$$

additivity of integrals on integral and expectation

$$= \sup_{\alpha}\left\{\mathbb{E}^{\omega_1}\left[\int_0^t f^{\alpha}(X_s^x(\omega_1))ds\right] + \right.$$
$$\left. + \mathbb{E}^{\omega_1}\left[\int_t^{+\infty} f^{\alpha}(X_s^x(\omega_1))ds\right] + \lim_{s\to+\infty}\mathbb{E}^{\omega_3}(F(X_s^x(\omega_3)))\right\} =$$

last addend isn't aleatory

$$= \sup_{\alpha}\left\{\mathbb{E}^{\omega_1}\left[\int_0^t f^{\alpha}(X_s^x(\omega_1))ds\right] + \right.$$
$$\left. + \mathbb{E}^{\omega_1}\left[\int_t^{+\infty} f^{\alpha}(X_s^x(\omega_1))ds + \lim_{s\to+\infty}\mathbb{E}^{\omega_3}(F(X_s^x(\omega_3)))\right]\right\} =$$

total probability formula

$$= \sup_{\alpha}\left\{\mathbb{E}^{\omega_1}\left[\int_0^t f^{\alpha}(X_s^x(\omega_1))ds\right] + \right.$$
$$\left. + \mathbb{E}^{\omega_1}\left[\mathbb{E}^{\omega_2}\left[\int_t^{+\infty} f^{\alpha}(X_s^x(\omega_2))ds + \lim_{s\to+\infty}\mathbb{E}^{\omega_3}(F(X_s^x(\omega_3)))\middle|\mathcal{F}_t^{(m)}\right](\omega_1)\right]\right\} =$$

time shift change of variables

$$= \sup_{\alpha}\left\{\mathbb{E}^{\omega_1}\left[\int_0^t f^{\alpha}(X_s^x(\omega_1))ds\right] + \right.$$
$$\left. + \mathbb{E}^{\omega_1}\left[\mathbb{E}^{\omega_2}\left[\int_0^{+\infty} f^{\alpha}(X_{s+t}^x(\omega_2))ds + \lim_{s\to+\infty}\mathbb{E}^{\omega_3}(F(X_{s+t}^x(\omega_3)))\middle|\mathcal{F}_t^{(m)}\right](\omega_1)\right]\right\}$$

Now we apply the strong Markov property (we are allowed to use it because the control α is a Markov control, so the process X^y is an Itô diffusion) so we get

$$v(x) = \sup_{\alpha}\left\{\mathbb{E}^{\omega_1}\left[\int_0^t f^{\alpha}(X_s^x(\omega_1))ds\right] + \mathbb{E}^{\omega_1}\left[\mathbb{E}^{\omega_2}\left[\int_0^{+\infty} f^{\alpha}(\tilde{X}_s^{X_t^x(\omega_1)}(\omega_2))ds\right] + \right.\right.$$
$$\left.\left. + \lim_{s\to+\infty}\mathbb{E}^{\omega_3}(F(\tilde{X}_s^{X_t^x(\omega_1)}(\omega_3)))\right]\right\}$$

Here the notation used for the process $\tilde{X} = \{\tilde{X}_s^{X_t^x(\omega_1)}\}_{s\geq 0}$ means that $\tilde{X}$ is the process solution of the equation

$$\tilde{X}_s = y + \int_0^s \sigma(\beta_r, \tilde{X}_r)dW_r + \int_0^s b(\beta_r, \tilde{X}_r)dr$$

where we set $y = X_t^x(\omega_1)$
We observe that technically speaking the process $\{\tilde{X}_s^{X_t^x(\omega_1)}\}_{s\geq 0}$ depends on a time shifted version of the control α.
Now it is possible to choose controls such that they optimize independently the supremum of the last two addends (conditioned to the aleatory starting point $X_t^x(\omega_1)$) and the supremum of the first addend (see Krylov [4]) and we get

$$v(x) = \sup_\alpha \left\{ \mathbb{E}^{\omega_1}\left[\int_0^t f^\alpha(X_s^x(\omega_1))ds\right] + \mathbb{E}^{\omega_1}\left[\sup_\beta \left\{ \mathbb{E}^{\omega_2}\left[\int_0^{+\infty} f^\beta(\tilde{X}_s^{X_t^x(\omega_1)}(\omega_2))ds\right] + \right.\right.\right.$$
$$\left.\left.\left. + \lim_{s\to+\infty} \mathbb{E}^{\omega_3}(F(\tilde{X}_s^{X_t^x(\omega_1)}(\omega_3)))\right\}\right]\right\}$$

Now we observe that the second expectation value is exactly the definition of the Bellman function of aleatory starting point $X_t^x(\omega_1)$.

$$v(X_t^x(\omega_1)) = \sup_\beta v^\beta(X_t^x(\omega_1))$$

So the expression that we get is

$$v(x) = \sup_\alpha \left\{ \mathbb{E}^{\omega_1}\left[\int_0^t f^\alpha(X_s^x(\omega_1))ds\right] + \mathbb{E}^{\omega_1}\left[v(X_t^x(\omega_1))\right]\right\}$$

Which is exactly the Bellman principle. □

We recall from Chapter 2 that the Bellman function v is solution of the Hamilton-Jacobi-Bellman equation. Given the operator

$$L^a(x) = \sum_{i=1}^n b_i(x,a)\frac{\partial}{\partial x_i}(x) + \sum_{i,j=1}^n \frac{1}{2}(\sigma\sigma^T)_{i,j}(x,a)\frac{\partial^2}{\partial x_i \partial x_j}(x), \quad \forall x \in \mathbb{R}^d \tag{3.5}$$

then by Theorem (1.2.8), as long as we assume that α is a Markov control $\alpha(t,\omega) = \alpha_0(X_t(\omega))$, then the process $\{X_t\}_{t\geq 0}$ is an Itô diffusion with infinitesimal generator given by the operator

$$A(x) = L^{\alpha_0(x)}(x)$$

so we can use Theorem (2.2.1), with notations adapted for this specific case, to prove that the equation (2.14) holds, which means that

$$\sup_{a\in A}\{f^a(x) + (L^a v)(x)\} = 0 \quad \text{for all } x \in \mathbb{R}^n \tag{3.6}$$

For the applications to the Harmonic Analysis we will be interested in supersolutions of the HJB equation (3.6), i.e. functions V that sastify the inequality

$$\sup_{a\in A}[L^a(x)V(x) + f^a(x)] \leq 0 \tag{3.7}$$

Observation 3.1.1. If the bequest function F is identically equal to 0 and the profit density f is non negative, then the Bellman function v is trivially non negative. Moreover, a function $V : \mathbb{R}^d \to \mathbb{R}$ such that $V \geq 0$ and such that it has continuous derivatives up to the second order and such that V satisfies the inequality (3.7) majorates v, where v is the Bellman function associated to the problem, i.e.

$$V(x) \geq v(x) = \sup_{\{\alpha_t\}_{t\geq 0}} v^{\alpha}(x)$$

Proof. To prove this result all we need to do is to apply Dynkin's formula (1.18) to the function V and the process $\{X_t\}_{t\geq 0}$ and to a deterministic time t, so we get

$$V(x) = E^x[V(X_t)] - E^x \int_0^t (L^{\alpha_s}(X_s)V)(X_s)ds$$

The inequality (3.7) implies that $-L^a(x)V(x) \geq f^a(x)$, and $V \geq 0$ so we get

$$V(x) \geq E^x \int_0^t f^{\alpha_s}(X_s)ds$$

So by taking the limit for $t \to +\infty$ and the supremum over all controls α we get

$$V(x) \geq \sup_{\alpha} E^x \int_0^{+\infty} f^{\alpha_s}(X_s)ds = v(x)$$

□

3.2 An example of the Bellman function's method

3.2.1 A_∞ weights and their associated Carleson measures

We will now take into consideration a theorem about an inequality over a dyadic tree as an example of a theorem that can be proved using the Bellman's function method. This theorem comes from complex analysis and it's useful to study the characterization of Carleson measures in Hardy spaces. Details about these topics can be found in [5].

Definition 3.2.1. A function

$$\omega : \mathbb{R} \longrightarrow \mathbb{R}^+$$

is called an A_∞ weight (we will write $\omega \in A_\infty$) if

$$\langle \omega \rangle_J \leq c_1 \, e^{\langle \log \omega \rangle_J}, \quad \forall J \in \mathcal{D} \tag{3.8}$$

where

$$\langle f \rangle_J = \frac{1}{|J|}\int_J f(x)dx$$

The Bellman function's method allows us to prove the following theorem

Theorem 3.2.1. *Let $\omega \in A_\infty$. Then*

$$\frac{1}{|I|}\sum_{l\subseteq I}\left(\frac{\langle \omega \rangle_{l_+} - \langle \omega \rangle_{l_-}}{\langle \omega \rangle_l}\right)^2 |l| \leq c_2, \quad \forall I \in \mathcal{D} \tag{3.9}$$

where c_2 is a real constant depending only on c_1 in (3.8).

3.2.2 The Bellman function method

In order to prove the theorem 3.2.1 we will start by showing a way to approach the problem with "common sense" and gather information about the Bellman function needed to solve this problem.
We need to find a domain $D \subseteq \mathbb{R}^n$ for some appropriate $n \in \mathbb{N}$, a map

$$\Phi : (\omega, I) \longmapsto (x_1, x_2, \dots, x_n) \in D$$

where $\omega \in A_\infty$ and $I \in \mathcal{D}$ are the variables in the problem (3.2.1), and a function (which will be a supersolution of the bellman equation, or eventually it could even be the Bellman function associated with this problem)

$$g : D \longrightarrow \mathbb{R} \tag{3.10}$$

such that g satisfies a property that allows us to prove the thesis of the theorem 3.2.1 when we compute $g(\Phi(\omega, I))$. This property will usually be satisfying an inequality, which will be named principal inequality.
For this problem we observe that the inequality (3.9) is rescaling invariant, so the function g computed in $\Phi(\omega, I)$ will not depend on the choice of the interval I.
We also observe that, for all $\omega \in A_\infty$, the inequality (3.9) and the A_∞ condition depend on $\langle\omega\rangle_J$ and $\langle\log\omega\rangle_J$ for $J \in \mathcal{D}$ and on $|I|$.
Based on these observation, we assume that the map $\Phi(\omega, I)$ will only need to keep track of $\langle\omega\rangle_I$, $\langle\log\omega\rangle_I$ and $|I|$. So we define

$$\Phi : (\omega, I) \longmapsto (u, w, |I|) := (\langle\log\omega\rangle_I, \langle\omega\rangle_I, |I|) \in \mathbb{R}^3 \tag{3.11}$$

Now that we found a candidate for the map Φ, we can use a simple approach to find a candidate for the Bellman function g.
We will now define an "ausiliary" function $\mathcal{B} : D \to \mathbb{R}$ with a standard procedure that will allow us to find the main inequality associated to the problem. Since we have to prove the inequality (3.9), we will take into consideration the function $\mathcal{B}$ defined as the supremum of the possible values that the left hand side of the thesis (3.9), so in order to prove the thesis it will be sufficient to prove that the function $\mathcal{B}$ is bounded.
We define

$$\mathcal{B} : D \longrightarrow \mathbb{R}$$
$$\mathcal{B}(u, w, |I|) = \mathcal{B}_I(u, w) := \sup_\omega \frac{1}{|I|} \sum_{J \subseteq I} \left(\frac{\langle\omega\rangle_{J_+} - \langle\omega\rangle_{J_-}}{\langle\omega\rangle_J} \right)^2 |J| \tag{3.12}$$

here the supremum is taken over all the A_∞ weights ω such that $\langle\log\omega\rangle_I = u$ and $\langle\omega\rangle_I = w$.
We define this function over the domain $D \subseteq \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+$ defined by

$$D = \{\Phi(\omega, I) = (\langle\log\omega\rangle_I, \langle\omega\rangle_I, |I|) \in \mathbb{R}^3 \mid \omega \in A_\infty,\ I \in \mathcal{D}\}$$

We observe that the function $\mathcal{B}$ does not depend on the third variable because, if we consider two intervals $I_1, I_2 \in \mathcal{D}$, then for all $u, v \in \mathbb{R} \times \mathbb{R}^+$ and for each function $\omega_1 \in A_\infty$ such that $\langle\log\omega_1\rangle_{I_1} = u$, $\langle\omega_1\rangle_{I_1} = w$ then it exists a function $\omega_2 \in A_\infty$

such that $\langle\log\omega_2\rangle_{I_2} = u$, $\langle\omega_2\rangle_{I_2} = w$ obtained by rescaling the function $\omega_{1|_{I_1}}$ over the interval I_2, so we will actually treat the function $\mathcal{B}$ as a function of two variables and we will just write

$$\Phi : A_\infty \times \mathcal{D} \longrightarrow D \subseteq \mathbb{R}^2$$

$$\Phi(\omega, I) = (u, w) := (\langle\log\omega\rangle_I, \langle\omega\rangle_I)$$

where D is the domain

$$D = \{(u, w) \in \mathbb{R} \times \mathbb{R}^+ \mid \exists\omega \in A_\infty, \exists I \in \mathcal{D} \text{ such that } \langle\log\omega_1\rangle_I = u,\ \langle\omega_1\rangle_I = w\}$$

and

$$\mathcal{B} : D \longrightarrow \mathbb{R}$$

$$\mathcal{B}(u, v) = \sup_{\omega, I} \frac{1}{|I|} \sum_{J \subseteq I} \left(\frac{\langle\omega\rangle_{J_+} - \langle\omega\rangle_{J_-}}{\langle\omega\rangle_J} \right)^2 |J|$$

here the supremum is taken over all $\omega \in A_\infty$ and $I \in \mathcal{D}$ such that $\langle\log\omega_1\rangle_I = u$, $\langle\omega_1\rangle_I = w$.

Now, in order to properly approach this problem and find a function g with a usable expression as mentioned earlier, we will first study the properties of the domain D and the properties of $\mathcal{B}$.

1. It is important to understand how big the domain is. We observe that, for all ω, the condition $\omega \in A_\infty$ tells us that $\langle\omega\rangle_I \leq c_1\ e^{\langle\log\omega\rangle_I}$, but the domain D is defined as the set of points (u, w) such that it exists $\omega \in A_\infty$ such that $\langle\log\omega_1\rangle_{I_1} = u$, $\langle\omega_1\rangle_{I_1} = w$, so this entails that $w \leq c_1\ e^u$. Moreover, we recall that the logarithmic function is concave, so by Jensen's inequality we get that $\langle\log\omega\rangle_I \leq \log\langle\omega\rangle_I$, which means $e^{\langle\log\omega\rangle_I} \leq \langle\omega\rangle_I$, so we get $e^u \leq w$.
   We have proved that $D \subseteq \{(u, w) \in \mathbb{R} \times \mathbb{R}^+ \mid e^u \leq w \leq c_1\ e^w\}$. It is possible to prove that $D = \{(u, w) \in \mathbb{R} \times \mathbb{R}^+ \mid e^u \leq w \leq c_1\ e^w\}$, but it's not necessary to find out the exact minimal domain of the function $\mathcal{B}$ as long as we can find a function $g : D \to \mathbb{R}$ with a good expression that satisfies the main inequality.

2. It is important to understand the boundaries of the function $\mathcal{B}$. We observe that

$$\mathcal{B}(u, v) = \sup_{\omega, I} \frac{1}{|I|} \sum_{J \subseteq I} \left(\frac{\langle\omega\rangle_{J_+} - \langle\omega\rangle_{J_-}}{\langle\omega\rangle_J} \right)^2 |J| \geq 0$$

   because it is the supremum of a collection of non negative values. On the other hand, in order to prove the thesys (3.9), since we defined the function $\mathcal{B}$ as the supremum of the left hand side of the thesis, it is necessary that $\mathcal{B}(u, v) \leq c_2$. So we proved that, in order to prove the theorem, our function g must be such that $0 \leq g(u, w) \leq c_2$ for all $(u, w) \in D$.

3. The function $\mathcal{B}$ satisfies a main inequality. To find this inequality the procedure is to take into consideration $(u, w) \in D$, $\omega \in A_\infty$ and $I \in \mathcal{D}$ such that

$\langle \log \omega \rangle_I = u$, $\langle \omega \rangle_I = w$. We also define

$$w_+ = \langle \omega \rangle_{I_+} = \frac{1}{|I_+|} \int\limits_{I_+} u(x)dx, \qquad w_- = \langle \omega \rangle_{I_-} = \frac{1}{|I_-|} \int\limits_{I_-} u(x)dx$$

$$u_+ = \langle \log \omega \rangle_{I_+} = \frac{1}{|I_+|} \int\limits_{I_+} \log u(x)dx, \quad u_- = \langle \log \omega \rangle_{I_-} = \frac{1}{|I_-|} \int\limits_{I_-} \log u(x)dx$$

We observe that it follows from these definitions that

$$w = \frac{1}{2}[w_+ + w_-], \quad u = \frac{1}{2}[u_+ + u_-]$$

Now we take into consideration the following sum

$$S(\omega, I) = \frac{1}{|I|} \sum_{J \subseteq I} \left(\frac{\langle \omega \rangle_{J_+} - \langle \omega \rangle_{J_-}}{\langle \omega \rangle_J} \right)^2 |J|$$

we write this sum as

$$\begin{aligned} S(\omega, I) =& \frac{1}{|I|} \left(\frac{\langle \omega \rangle_{I_+} - \langle \omega \rangle_{I_-}}{\langle \omega \rangle_I} \right)^2 |I| + \frac{1}{|I|} \sum_{J \subseteq I_+} \left(\frac{\langle \omega \rangle_{J_+} - \langle \omega \rangle_{J_-}}{\langle \omega \rangle_J} \right)^2 |J| + \\ &+ \frac{1}{|I|} \sum_{J \subseteq I_-} \left(\frac{\langle \omega \rangle_{J_+} - \langle \omega \rangle_{J_-}}{\langle \omega \rangle_J} \right)^2 |J| = \\ &= \left(\frac{\langle \omega \rangle_{I_+} - \langle \omega \rangle_{I_-}}{\langle \omega \rangle_I} \right)^2 + \frac{1}{2} \cdot \frac{1}{|I_+|} \sum_{J \subseteq I_+} \left(\frac{\langle \omega \rangle_{J_+} - \langle \omega \rangle_{J_-}}{\langle \omega \rangle_J} \right)^2 |J| + \\ &+ \frac{1}{2} \cdot \frac{1}{|I_-|} \sum_{J \subseteq I_-} \left(\frac{\langle \omega \rangle_{J_+} - \langle \omega \rangle_{J_-}}{\langle \omega \rangle_J} \right)^2 |J| = \\ &= \left(\frac{w_+ - w_-}{w} \right)^2 + \frac{1}{2} \Big[S(\omega, I_+) + S(\omega, I_-) \Big] \end{aligned}$$

Now we take the supremum of $S(\omega, I)$ over all the possible $\omega \in A_\infty$ and $I \in \mathcal{D}$ taken into consideration and, using the definition of $\mathcal{B}$ and the fact that it does not depend on the choice of I, we get

$$\sup_{\omega, I} S(\omega, I) = \mathcal{B}(u, w) = \sup_{\omega} S(\omega, I)$$

so this entails that

$$\mathcal{B}(u, w) \geq S(\omega, I) = \left(\frac{w_+ - w_-}{w} \right)^2 + \frac{1}{2} \Big[S(\omega, I_+) + S(\omega, I_-) \Big]$$

we notice that ω and I_+ are such that $w_+ = \langle \omega \rangle_{I_+}$ and $u_+ = \langle \log \omega \rangle_{I_+}$, and on the other hand ω and I_- are such that $w_- = \langle \omega \rangle_{I_-}$ and $u_- = \langle \log \omega \rangle_{I_-}$. Since we can define independently a function ω over I_+ and I_- such that $w_+ = \langle \omega \rangle_{I_+}$, $u_+ = \langle \log \omega \rangle_{I_+}$, $w_- = \langle \omega \rangle_{I_-}$, $u_- = \langle \log \omega \rangle_{I_-}$ for all choices of w_+, u_+, w_-, u_-, such that

$$w = \frac{1}{2}[w_+ + w_-], \quad u = \frac{1}{2}[u_+ + u_-]$$

we can consider the supremum over all possible ω and get

$$\begin{aligned} \mathcal{B}(u, w) \geq& \left(\frac{w_+ - w_-}{w} \right)^2 + \frac{1}{2} \Big[\sup_{\omega} [S(\omega, I_+)] + S(\omega, I_-) \Big] = \\ =& \left(\frac{w_+ - w_-}{w} \right)^2 + \frac{1}{2} \Big[\mathcal{B}(u_+, v_+) + S(\omega, I_-) \Big] \end{aligned}$$

using the same argument over $S(\omega, I_-)$ we finally get the main inequality

$$\mathcal{B}(u,w) \geq \left(\frac{w_+ - w_-}{w}\right)^2 + \frac{1}{2}\Big[\mathcal{B}(u_+,v_+) + \mathcal{B}(u_-,v_-)\Big] \tag{3.13}$$

for all choices of $(u,w),(u_+,w_+),(u_-,w_-) \in D$ such that

$$w = \frac{1}{2}[w_+ + w_-], \quad u = \frac{1}{2}[u_+ + u_-]$$

The main inequality is the tool that allows us to solve the problem. Infact, given a function $g : D \to \mathbb{R}$, $0 \leq g \leq c_2$, that satisfies the main inequality (3.13), i.e.

$$g(\Phi(\omega,I)) \geq \left(\frac{\langle\omega\rangle_{I_+} - \langle\omega\rangle_{I_-}}{\langle\omega\rangle_I}\right)^2 + \frac{1}{2}\Big[g(\Phi(\omega,I_+)) + g(\Phi(\omega,I_-))\Big]$$

we can compute for all $\omega \in A_\infty$ and for all $I \in \mathcal{D}$

$$|I|\cdot g(\Phi(\omega,I)) \geq |I|\cdot\left(\frac{\langle\omega\rangle_{I_+} - \langle\omega\rangle_{I_-}}{\langle\omega\rangle_I}\right)^2 + \frac{1}{2}\,|I|\cdot[g(\Phi(\omega,I_+)) + g(\Phi(\omega,I_-))]$$

We observe that $\frac{1}{2}|I| = |I_+| = |I_-|$ so we get that

$$|I|\cdot g(\Phi(\omega,I)) \geq |I|\cdot\left(\frac{\langle\omega\rangle_{I_+} - \langle\omega\rangle_{I_-}}{\langle\omega\rangle_I}\right)^2 + |I_+|\cdot g(\Phi(\omega,I_+)) + |I_-|\cdot g(\Phi(\omega,I_-)) \tag{3.14}$$

so we compute the main inequality (3.13) for $g(\phi(\omega,I_+))$ and $g(\phi(\omega,I_-))$ and substitute it in (3.14) and iterate the procedure an infinite amount of times and, using the fact that

$$\left(\frac{\langle\omega\rangle_{J_+} - \langle\omega\rangle_{J_-}}{\langle\omega\rangle_J}\right)^2 \geq 0 \quad \forall J \in \mathcal{D}$$

and the fact that $g \geq 0$ it follows that

$$|I|\cdot g(\Phi(\omega,I)) \geq \sum_{J\subseteq I}\left(\frac{\langle\omega\rangle_{J_+} - \langle\omega\rangle_{J_-}}{\langle\omega\rangle_J}\right)^2 |J|$$

so by using the fact that $g \leq c_2$ it follows that

$$c_2 \geq g(\Phi(\omega,I)) \geq \frac{1}{|I|}\sum_{J\subseteq I}\left(\frac{\langle\omega\rangle_{J_+} - \langle\omega\rangle_{J_-}}{\langle\omega\rangle_J}\right)^2 |J|$$

which is exactly the thesis (3.9) that we wanted to prove.
So in order to solve the problem all we need to do is to find a function

$$g : D \longrightarrow \mathbb{R}$$

defined over the domain

$$D = \{(u,w) = (\langle\log\omega\rangle_I, \langle\omega\rangle_I) = \Phi(\omega,I) \mid \omega \in A_\infty,\ I \in \mathcal{D}\}$$

such that $0 \leq g \leq c_2$ and such that it satisfies the inequality

$$g(u,w) \geq \left(\frac{w_+ - w_-}{w}\right)^2 + \frac{1}{2}\Big[g(u_+,v_+) + g(u_-,v_-)\Big]$$

for all $(u, w), (u_+, w_+), (u_-, w_-) \in D$ such that

$$w = \frac{1}{2}[w_+ + w_-], \quad u = \frac{1}{2}[u_+ + u_-]$$

We also observed that $D \subseteq \{(u, w) \in \mathbb{R}^2 \mid e^u \leq w \leq c_1 \, e^u\}$, so we will just try to find a function g with the mentioned properties over the set $\{(u, w) \in \mathbb{R}^2 \mid e^u \leq w \leq c_1 \, e^u\}$ (although we mentioned earlier that this step is necessary because it can be proved that $D = \{(u, w) \in \mathbb{R}^2 \mid e^u \leq w \leq c_1 \, e^u\}$, but we do not need to prove it as long as we prove that D is a subset of the domain of definition of g).
Now it is time to find a candidate function g

$$g : D = \{(u, w) \in \mathbb{R}^2 \mid e^u \leq w \leq c_1 \, e^u\} \longrightarrow \mathbb{R}$$

with the mentioned properties.
There is no formula to immediatly find a function g with such properties, so we have to use our information about the function to make an educated guess. The domain of the function can suggest us some natural guesses for the function g. It may be interesting to see a wrong guess first. We observe that, since $e^u \leq w \leq c_1 \, e^u$, then the function

$$g_1 : D \longrightarrow \mathbb{R}$$

$$g_1(u, w) = \tilde{c}[w - e^u]$$

where $\tilde{c} \in \mathbb{R}$ is a constant, is such that, for $(u, w) \in D$, it follows that $0 \leq g_1(u, w)$, however this function can't be a good candidate because we also notice that $g_1(u, w) \leq \tilde{c}(c_1 - 1)e^u$, moreover, as long as we choose $w = c_1 \, e^u - \epsilon$ for $\epsilon > 0$ a small real number, we get that $g_1(u, w) = \tilde{c}[(c_1 - 1)e^u - \epsilon]$, so $g_1(u, v)$ can be arbitrarily big as long as we chose a big enough u, so it is not true that $g_1(u, w) \leq c_2$.
We can also observe that, from the definition of D, for all $(u, w) \in D$ it follows that $u \leq \log w$ and $\log w \leq u + \log c_2$, so we can make another educated guess and consider the function

$$g_2 : D \longrightarrow \mathbb{R}$$

$$g_2(u, w) = \tilde{c}[\log(w) - u] \tag{3.15}$$

We observe that, for all $(u, w) \in D$, then $g_2(u, w) = \tilde{c}[\log w - u] \geq \tilde{c}[u - u] = 0$ and $g_2(u, w) \leq \tilde{c}[\log(c_1 \, e^u) - u] = \tilde{c} \, c_1$. So, after considering $c_2 = \tilde{c} \, c_1$, we get $0 \leq g_2(u, w) \leq c_2$, which is the first property we needed. We will prove that g_2 is actually the function we are looking for with $\tilde{c} = 8$, moreover we will also show with a heuristic proof that it actually is the Bellman function associated to this problem. Before doing that wee may get another hint that g_2 could be the right guess by using the (discrete) main inequality (3.13) to get a continuous version of the inequality for the Bellman function.

3.2.3 Continuous version of the main inequality

Let $x = (u, w)$ be a point of the interior of the domain D, let $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ be a fixed point such that the segment of points $\{(u + \alpha_1 t, w + \alpha_2 t) \in \mathbb{R}^2 \mid -1 \leq t \leq 1\}$

is a subset of the domain D. Let $x_+ = (u + \alpha_1, w + \alpha_2)$, $x_- = (u - \alpha_1, w - \alpha_2)$. Consider the functions ϕ, ψ defined by

$$\phi(t) = \mathcal{B}(x + \alpha t); \quad -1 \le t \le 1$$

$$\psi(t) = \mathcal{B}(x) - \frac{1}{2}\Big(\mathcal{B}(x + \alpha t) + \mathcal{B}(x - \alpha t)\Big) - \left(\frac{2\alpha_2 t}{w}\right)^2; \quad -1 \le t \le 1$$

If the function $\mathcal{B}$ is C^2-smooth then we get

$$\lim_{t\to 0} \frac{\phi(t) + \phi(-t) - 2\phi(0)}{t^2} = \phi''(0) = \sum_{j,k=1}^{2} \frac{\partial^2 \mathcal{B}(x)}{\partial x_j \partial x_k} \alpha_j \alpha_k \tag{3.16}$$

Since x, x_+ and x_- are points of the domain D, then by the main inequality (3.13) we get that $\psi(t) = 0\ \forall t \in [-1, 1]$.
So we may consider the inequality $\psi(t)/t^2 \ge 0$ and take the limit for $t \to 0$ and, taking into account (3.16), we get

$$\begin{aligned} 0 \le \lim_{t\to 0} \frac{\psi(t)}{t^2} &= \lim_{t\to 0} \Bigg[- \frac{\mathcal{B}(x + \alpha t) + \mathcal{B}(x - \alpha t) - 2\mathcal{B}(x)}{2t^2} - \left(\frac{2\alpha_2 t}{wt}\right)^2 \Bigg] = \\ &= \lim_{t\to 0} \Bigg[-\frac{1}{2} \frac{\phi(t) + \phi(-t) - 2\phi(0)}{t^2} \Bigg] - \left(\frac{2\alpha_2}{w}\right)^2 = \\ &= -\frac{1}{2}\phi''(0) - \left(\frac{2\alpha_2}{w}\right)^2 = \\ &= -\frac{1}{2} \sum_{j,k=1}^{2} \frac{\partial^2 \mathcal{B}(x)}{\partial x_j \partial x_k} \alpha_j \alpha_k - \left(\frac{2\alpha_2}{w}\right)^2 \end{aligned}$$

So the inequality we got is

$$-\frac{1}{2} \sum_{j,k=1}^{2} \frac{\partial^2 \mathcal{B}(x)}{\partial x_j \partial x_k} \alpha_j \alpha_k \ge 4\left(\frac{\alpha_2}{w}\right)^2 \tag{3.17}$$

Since the inequality (3.17) is homogeneous of degree 2, then it holds for all choices of x, x_+, x_- in the domain D. This is the continuous version of the main inequality (3.13).
We observe now that, if we consider a controlled process $\{X_t\}_{t\ge 0}$ defined by

$$X_t = x + \int_0^t \alpha_s \, dW_s; \quad x_s = (x_{s,1}, x_{s,2}), \quad \alpha_s = (\alpha_{s,1}, \alpha_{s,2})$$

where $\{W_s\}_{s\ge 0}$ is a Brownian motion, then this process is solution of the equation (3.1) where $b(a, x) = 0$ and $\sigma(a, x) = a$. So, if we consider the profit density f and the bequest function F defined by

$$f^{\alpha}(x) = 4\left(\frac{\alpha_2}{x_2}\right)^2, \quad F(x) = 0$$

Then we can consider the Hamilton-Jacobi-Bellman equation associated to the Bellman function v defined by

$$v(x) = \sup_{\alpha} E^x \Bigg[\int_0^{+\infty} f^{\alpha_t}(X_t) dt \Bigg] \tag{3.18}$$

which is, by using (3.6) and (3.5), the equation

$$\sup_{\alpha\in A}\left\{\frac{1}{2}\sum_{j,k=1}^{2}\frac{\partial^2 v(x)}{\partial x_j\partial x_k}\alpha_j\alpha_k+4\left(\frac{\alpha_2}{x_2}\right)^2\right\}=0 \tag{3.19}$$

We observe that earlier we proved that the function $\mathcal{B}$ satisfies the inequality

$$\frac{1}{2}\sum_{j,k=1}^{2}\frac{\partial^2 \mathcal{B}(x)}{\partial x_j\partial x_k}\alpha_j\alpha_k+4\left(\frac{\alpha_2}{x_2}\right)^2\leq 0$$

so this proves that the function $\mathcal{B}$ is a supersolution of the HJB equation (3.19).
By observation 3.1.1 we know that, since the bequest function is $F=0$ and the profit density is $f^\alpha(x)=4(\alpha_2/x_2)^2\geq 0$, then it follows that $\mathcal{B}\geq v$, moreover it follows that any smooth function g that satisfies the main inequality (3.13) majorates the Bellman function v.
We mentioned earlier that the function g_2 defined in (3.15) is the function we are looking for (with $\tilde{c}=8$). Another way to get a hint about this fact is to observe that $g_2(u,w)=\tilde{c}[\log(w)-u]$ is a solution of the HJB equation (3.19) when $\tilde{c}=8$. Infact we compute the HJB equation for the function g_2 and, for $\alpha=(\alpha_1,\alpha_2)\in A$ and $\tilde{c}=8$, we get

$$\begin{aligned}
f^\alpha(u,w)+(L^\alpha g_2)(u,w) &= \frac{1}{2}\sum_{j,k=1}^{2}\frac{\partial^2 g_2(u,w)}{\partial x_j\partial x_k}\alpha_j\alpha_k+4\left(\frac{\alpha_2}{w}\right)^2=\\
&=\frac{1}{2}\left[\frac{\partial^2 g_2(u,w)}{\partial u^2}\alpha_1^2+\frac{\partial^2 g_2(u,w)}{\partial w^2}\alpha_2^2+2\frac{\partial^2 g_2(u,w)}{\partial u\partial w}\alpha_1\alpha_2\right]+4\left(\frac{\alpha_2}{w}\right)^2=\\
&=\frac{1}{2}\left[0-\tilde{c}\frac{1}{w^2}\alpha_2^2+2\cdot 0\right]+4\left(\frac{\alpha_2}{w}\right)^2=\\
&=-\frac{1}{2}\cdot 8\cdot\left(\frac{\alpha_2}{w}\right)^2+4\left(\frac{\alpha_2}{w}\right)^2=0
\end{aligned}$$

So by taking the supremum over all $\alpha\in A$ we get

$$\sup_{\alpha\in A}\left\{\frac{1}{2}\sum_{j,k=1}^{2}\frac{\partial^2 g_2(u,w)}{\partial x_j\partial x_k}\alpha_j\alpha_k+4\left(\frac{\alpha_2}{w}\right)^2\right\}=0\quad\forall(u,w)\in D$$

which means that $g_2(u,w)=8[\log w-u]$ is a solution of the HJB equation (3.19).

3.2.4 Solution of the problem

We mentioned earlier that the function $g_2(u,w)=8[\log w-u]$ is the function we are looking for in subsection 3.2.2, now we will prove it. We already proved that $0\leq g_2(u,w)\leq c_2$ where $c_2=8\cdot c_1$, so we only need to prove that g_2 satisfies the main inequality (3.13), so we need to prove that

$$g_2(u,w)-\frac{1}{2}\Big[g_2(u_+,v_+)+g_2(u_-,v_-)\Big]\geq\left(\frac{w_+-w_-}{w}\right)^2$$

for all choices of $(u,w),(u_+,w_+),(u_-,w_-)\in D$ such that

$$w=\frac{1}{2}[w_++w_-],\quad u=\frac{1}{2}[u_++u_-]$$

We denote $(u_+, w_+) = (u + \alpha_1, w + \alpha_2)$ and $(u_-, w_-) = (u - \alpha_1, w - \alpha_2)$. Then

$$\begin{aligned} & g_2(u,w) - \frac{1}{2}\Big[g_2(u_+,w_+) + g_2(u_-,w_-)\Big] = \\ &= 8\Big[\log w - u - \frac{1}{2}(\log(w+\alpha_2) - u - \alpha_1 + \log(w - \alpha_2) - u + \alpha_1)\Big] = \\ &= 8\Big[\log w - \frac{1}{2}\log\Big((w+\alpha_2)(w-\alpha_2)\Big)\Big] = \\ &= 8\Big[\log w - \frac{1}{2}\Big(\log(w^2) + \log\Big(1 - \Big(\frac{\alpha_2}{w}\Big)^2\Big)\Big)\Big] = \\ &= -4\log\Big(1 - \Big(\frac{\alpha_2}{w}\Big)^2\Big) \end{aligned}$$

Since we chose $(u,w), (u_+,w_+), (u_-,w_-) \in D$, then we have $w - \alpha_2 \geq e^u > 0$, so we have $0 \leq (\alpha_2/w)^2 < 1$. Now it is easy to prove that, for all $0 \leq x < 1$, then $\log(1-x) \leq -x$, which is equivalent to $-\log(1-x) \geq x$, which means that, when $x = (\alpha_2/w)^2$, we get

$$g_2(u,w) - \frac{1}{2}\Big[g_2(u_+,w_+) + g_2(u_-,w_-)\Big] \geq \Big(\frac{2\alpha_2}{w}\Big)^2$$

which is exactly the main inequality

$$g_2(u,w) - \frac{1}{2}\Big[g_2(u_+,w_+) + g_2(u_-,w_-)\Big] \geq \Big(\frac{w_+ - w_-}{w}\Big)^2$$

This completes the proof of theorem 3.2.1 as we observed earlier in subsection 3.2.2. To conclude this analysis over the problem 3.2.1 we will show a heuristic proof of the fact that the function g_2 is actually the Bellman function.
To show this we will use discrete stochastic processes. Let us consider the process $\{\zeta_n\}_{n\in\mathbb{N}} = \{\sum_{k=0}^{n}\xi_k\}_{n\in\mathbb{N}}$, where ξ_k for $k \in \mathbb{N}$ are independent random variables taking values 1 and -1 with probabilities $1/2$. This process is the discrete analogous to the Brownian motion. We consider a control $\{\alpha_k\}_{k\in\mathbb{N}}$, $\alpha_k = (\alpha_{k,1}, \alpha_{k,2}) \in \mathbb{R}^2$. We define a process

$$X_{n+1} = X_n + \alpha_n\xi_n, \quad X_0 = (u,w) \in \mathbb{R}^2$$

This is the discrete version of the equation (3.1), it can be interpreted as

$$X_n = X_0 + \int_0^n \alpha_k \, d\zeta_k$$

Whith this procedure it is possible to find a correspondence between A_∞ weights and processes $X = \{(u_k, w_k)\}_{k\in\mathbb{N}}$ controlled by $\alpha = \{\alpha_k\}_{k\in\mathbb{N}}$. The way to get this correspondence is by starting from the interval I and going down the dyadic tree to its subintervals in the following way: at the time n we are on a subinterval J of length $2^{-n}|I|$ and the position is determined by the sequence of coin tosses ξ_k, i.e. we move from a interval K to a interval K_+ if $\xi_k = 1$, and we move to K_- if $\xi_k = -1$. This way for every $\omega \in A_\infty$ we get that a process $X_n = (u_n, w_n)$ is the vector of averages over an appropriate interval J, starting from the initial state $X_0 = (\langle\log\omega\rangle_I, \langle\omega\rangle_I)$ and it is controlled by a control $\alpha = \{\alpha_k\}_{k\in\mathbb{N}}$ defined by $\alpha_n = (X_{n+1} - X_n)/\xi_n$, i.e.

$$\alpha_{n,1} = \frac{1}{2}\Big(\langle\log\omega\rangle_{J_+} - \langle\log\omega\rangle_{J_-}\Big), \quad \alpha_{n,2} = \frac{1}{2}\Big(\langle\omega\rangle_{J_+} - \langle\omega\rangle_{J_-}\Big)$$

This is the way to associate a weight ω to a control α.
On the other hand if we have a process $X = \{X_k\}_{k\in\mathbb{N}}$ defined by an initial state (u, w) and a control α such that

$$\lim_{n\to+\infty} e^{u_n} = \lim_{n\to+\infty} w_n$$

with probability 1, then this process defines a unique A_∞ weight. The values of X are vectors of averages

$$X_k = (\langle \log\omega\rangle_J, \langle\omega\rangle_J) = (u_k, w_k)$$

for a proper interval J determined by the sequence of the coin tosses. So we just define

$$\omega = \lim_{n\to+\infty} \sum_{\substack{J\subseteq I \\ |J|=2^{-n}|I|}} w_J \chi_J$$

for every $n \in \mathbb{N}$ the element of the succession

$$\sum_{\substack{J\subseteq I \\ |J|=2^{-n}|I|}} w_J \chi_J$$

is an A_∞ weight, so the limit of this succession will also satisfy the A_∞ condition, so ω is the weight we were looking for.
Now, given $\epsilon > 0$, we will construct a control α (that corresponds to a weight $\omega \in A_\infty$ based on the last argument) such that

$$E^x\left[\sum_{n=0}^{+\infty} 4\left(\frac{\alpha_{n,2}}{w_n}\right)^2\right] \geq (1-\epsilon)g_2(x) \tag{3.20}$$

This will be enough to prove that $g_2 = v$, where v is the Bellman function defined in (3.18). This follows because

$$v(x) = \sup_\alpha v^\alpha(x) = \sup_\alpha E^x\left[\int_0^{+\infty} f^{\alpha_t}(X_t)dt\right]$$

and, when we consider our case with a discrete set of times $t \in \mathbb{N}$, the expression of the payoff becomes

$$v^\alpha(x) = E^x\left[\sum_{n=0}^{+\infty} f^{\alpha_n}(X_n)\right]$$

and since the profit density is

$$f^{(\alpha_1,\alpha_2)}(x_1, x_2) = 4\left(\frac{\alpha_2}{x_2}\right)^2$$

then, if (3.20) holds, we get

$$v(x) \geq v^\alpha(x) = E^x\left[\sum_{n=0}^{+\infty} 4\left(\frac{\alpha_{n,2}}{w_n}\right)^2\right] \geq (1-\epsilon)g_2(x)$$

for all $\epsilon > 0$, so $v(x) \geq g_2(x)$, but we already proved that $v(x) \leq g_2(x)$ earlier at the end of section 3.1, so this would mean that $v = g_2$ completing the heuristic argument.

To construct this process α let us fix a line segment $L = [A, B]$ in the domain D such that $x \in L$ and the endpoints A and B are on the lower bound of the domain $\{(u, w) \mid w = e^u\}$. Our process will be a discrete random walk on the line L with small steps. Let $\delta > 0$ be a fixed number we will choose later. Let l be a unit vector parallel to L. We define

$$\alpha_n = \min\{\delta, \operatorname{dist}(x_{n-1}, \{A, B\})\} \cdot l$$

this means that when we consider the process

$$X_n = x + \int_0^n \alpha_k \, d\zeta_k$$

then X_{n+1} will be equal to $X_n + \delta \cdot l$ if the $n - th$ coin toss is heads (i.e. $\xi_n = 1$), or will be equal to $X_n - \delta \cdot l$ if it is tails instead ($\xi_n = 1$), unless X_n is very close to the boundary, in that case the point will either land on the boundary point (A or B) or it will move on the other direction of the distance between X_n and the boundary point. A consequence of this fact is that if the point X_n lands on the boundary points in $\{A, B\}$ then the process will never move from there, and $X_m = X_n$ for all $m > n$.

Using the properties of the random walk it is easy to see that all trajectories hit the boundary with probability 1. So, applying Taylor's formula and using the compactness of L we can choose a small enough δ such that

$$g_2(x) \leq \frac{1}{2}\Big(g_2(x+\alpha) + g_2(x-\alpha)\Big) + (1+\epsilon)\cdot 4 \cdot \left(\frac{\alpha_2}{w}\right)^2 \tag{3.21}$$

Now we iterate the inequality (3.21) by computing the same inequality for $g_2(x+\alpha)$ and $g_2(x-\alpha)$ and substitute it in the right hand side of (3.21) and we repeat the procedure N times, and then we compute the limit for $N \to +\infty$. After iterating the procedure N times we get

$$g_2(x) \leq (1+\epsilon) \sum_{n=1}^{N+1} \sum_{i=1}^{2^n} 4\left(\frac{\alpha_{n,2}^{(i)}}{w_n^{(i)}}\right)^2 \frac{1}{2^n} + \sum_{i=1}^{2^{N+1}} \frac{1}{2^{N+1}} g_2(X_{N+1}^{(i)}) \tag{3.22}$$

Where. given a process $\{Z_t\}$, $Z_t^{(i)}$ for $i = 1, 2, \ldots, M$ is a notation to enumerate the M possible values of Z_t.

Now we observe that, because each trajectory up to the time n has probability $1/2^n$ to happen for the definition of the process $X_n = (u_n, w_n)$, then

$$\sum_{i=1}^{2^n} 4\left(\frac{\alpha_{n,2}^{(i)}}{w_n^{(i)}}\right)^2 \frac{1}{2^n} = E^x\left[4\left(\frac{\alpha_{n,2}}{w_n}\right)^2\right]$$

$$\sum_{i=1}^{2^N} \frac{1}{2^N} g_2(X_N^{(i)}) = E^x\Big[g_2(X_N)\Big]$$

Now we observe that

$$\lim_{N\to+\infty} E^x\Big[g_2(X_N)\Big] = 0$$

because $X_N \in \{A, B\}$ almost surely for $N \to +\infty$ and $g_2(x) = 0$ when $x \in \{A, B\}$ and g_2 is bounded in his domain, so it is bounded on the trajectories that don't end

on the boundary points $\{A, B\}$.
So we can compute the limit for $N \to +\infty$ of (3.22) and we get

$$E^x\left[\sum_{n=0}^{+\infty} 4\left(\frac{\alpha_{n,2}}{w_n}\right)^2\right] \geq \frac{1}{1+\epsilon} g_2(x) \geq (1-\epsilon) g_2(x)$$

which is exactly (3.20), completing the heuristic argument.

Chapter 4

A new Bellman function on a tree

In this chapter we will prove Hardy's inequality using the Bellman's function method. Hardy's inequality comes from harmonic analysis and it allows to characterize Carleson measures for Besov spaces. For more details we suggest to check [8] and [9] for a proof that Hardy's inequality allows us to characterize Carleson measures.

4.1 Bellman function on a tree

Let $p \in \mathbb{R}$, $1 < p < +\infty$. We consider the function

$$\mathcal{B}(F, f, A, v) = F - \frac{f^p}{(A+v)^{p-1}} \tag{4.1}$$

defined over the domain

$$\mathcal{D} := \left\{ (F, f, A, v) \in \mathbb{R}^4 \;\middle|\; F \geq 0,\ f \geq 0,\ A > 0,\ v > 0,\ v \geq A,\ f^p \leq F v^{p-1} \right\}$$

This function is the function used in the article [1, pg. 3], in the general case $p \neq 2$.

Observation 4.1.1. The function $\mathcal{B}$ has the following properties:

1) $\mathcal{B}$ is a concave function defined over a convex domain

2) $F \geq \mathcal{B}(F, f, A, v) \geq 0$

Proof. 1) We write the domain $\mathcal{D}$ as

$$\mathcal{D} = \{v \geq A\} \cap \mathcal{A}$$

here $\mathcal{A}$ is the set

$$\mathcal{A} = \{(F, f, A, v) \in \mathbb{R}^4 \mid F \geq 0, f \geq 0, v > 0, f^p \leq F v^{p-1}\}$$

To prove that the domain is convex we just need to prove that it is a intersection of convex sets.

The set $\{v \geq A\}$ is trivially convex because it is a half-plane. Since $\frac{1}{p} + \frac{1}{q} = 1$ and $\frac{p-1}{p} = \frac{1}{q}$, the set $\mathcal{A}$ can be written as

$$\mathcal{A} = \mathcal{S} \cap \{f \geq 0\}$$

here $\{f \geq 0\}$ is another half-plane (a convex set), while $\mathcal{S}$ is the set

$$\mathcal{S} = \{(F, f, A, v) \in \mathbb{R}^4 \mid F \geq 0, v > 0,\ f \leq F^{\frac{1}{p}} v^{\frac{1}{q}}\}$$

The set $\mathcal{S}$ is the subgraph of the function

$$\begin{aligned} h : \mathbb{R}_0^+ \times \mathbb{R} \times \mathbb{R}^+ &\longrightarrow \mathbb{R}_0^+ \\ (F, A, v) &\longmapsto F^{\frac{1}{p}} v^{\frac{1}{q}} \end{aligned}$$

To prove that $\mathcal{S}$ is convex, all we need to do is to prove that h is a concave function (since h is defined over a convex domain).
Since h does not depend on the variable A, we will treat it as a function over the other two variables only:

$$\begin{aligned} h : \mathbb{R}_0^+ \times \mathbb{R}^+ &\longrightarrow \mathbb{R}_0^+ \\ (F, v) &\longmapsto F^{\frac{1}{p}} v^{\frac{1}{q}} \end{aligned}$$

We compute the Hessian matrix of the function h: for all $F > 0$, $v > 0$

$$\frac{\partial h}{\partial F}(F,v) = \frac{1}{p} F^{\frac{1}{p}-1} v^{\frac{1}{q}}, \quad \frac{\partial h}{\partial v}(F,v) = \frac{1}{q} F^{\frac{1}{p}} v^{\frac{1}{q}-1}$$

$$\begin{aligned} \frac{\partial^2 h}{\partial F^2}(F,v) &= \frac{1-p}{p^2} F^{\frac{1-2p}{p}} v^{\frac{1}{q}}, & \frac{\partial^2 h}{\partial v \partial F}(F,v) &= \frac{1}{pq} F^{\frac{1-p}{p}} v^{\frac{q-1}{q}} \\ \frac{\partial^2 h}{\partial F \partial v}(F,v) &= \frac{1}{pq} F^{\frac{1-p}{p}} v^{\frac{q-1}{q}}, & \frac{\partial^2 h}{\partial v^2}(F,v) &= \frac{1-q}{q^2} F^{\frac{1}{p}} v^{\frac{1-2q}{q}} \end{aligned}$$

$$H_h(F,v) = \begin{bmatrix} \frac{1-p}{p^2} F^{\frac{1}{p}-2} v^{\frac{1}{q}} & \frac{1}{pq} F^{\frac{1}{p}-1} v^{\frac{1}{q}-1} \\ \\ \frac{1}{pq} F^{\frac{1}{p}-1} v^{\frac{1}{q}-1} & \frac{1-q}{q^2} F^{\frac{1}{p}} v^{\frac{1}{q}-2} \end{bmatrix} \tag{4.2}$$

So as long as the Hessian matrix of h has non positive eigenvalues then the function h is concave.
Now we compute the eigenvalues of the Hessian matrix (4.2):

$$\det(H_h(F,v) - \lambda I) = \det \begin{bmatrix} \frac{1-p}{p^2} F^{\frac{1}{p}-2} v^{\frac{1}{q}} - \lambda & \frac{1}{pq} F^{\frac{1}{p}-1} v^{\frac{1}{q}-1} \\ \\ \frac{1}{pq} F^{\frac{1}{p}-1} v^{\frac{1}{q}-1} & \frac{1-q}{q^2} F^{\frac{1}{p}} v^{\frac{1}{q}-2} - \lambda \end{bmatrix} =$$

$$\begin{aligned} =& \frac{(1-p)(1-q)}{(pq)^2} F^{\frac{1}{p}+\frac{1}{p}-2} v^{\frac{1}{q}+\frac{1}{q}-2} - \frac{1}{(pq)^2} F^{2(\frac{1}{p}-1)} v^{2(\frac{1}{q}-1)} \\ & - \lambda \left[\frac{1-p}{p^2} F^{\frac{1}{p}-2} v^{\frac{1}{q}} + \frac{1-q}{q^2} F^{\frac{1}{p}} v^{\frac{1}{q}-2} \right] + \lambda^2 \end{aligned}$$

Now we recall that $pq = p+q$, so $(1-p)(1-q) = 1-p-q+pq = 1-p-q+p+q = 1$, so we get that

$$\det(H_h(F,v) - \lambda I) = \lambda^2 - \lambda\left[\frac{1-p}{p^2}F^{\frac{1}{p}-2}v^{\frac{1}{q}} + \frac{1-q}{q^2}F^{\frac{1}{p}}v^{\frac{1}{q}-2}\right]$$

The eigenvalues of $H_h(F,v)$ are the solutions of the equation of variable λ

$$\det(H_h(F,v) - \lambda I) = 0$$

that are the two values

$$\lambda_1 = 0, \quad \lambda_2 = \frac{1-p}{p^2}F^{\frac{1}{p}-2}v^{\frac{1}{q}} + \frac{1-q}{q^2}F^{\frac{1}{p}}v^{\frac{1}{q}-2}$$

Now we observe that $1-p < 0$, $1-q < 0$ and $F > 0$, $v > 0$, so the second eigenvalue is $\lambda_2 < 0$, so the Hessian matrix $H_h(F,v)$ is negative semi-definite for all $F > 0$ and $v > 0$ and, since h is continuos up to the boundary of its domain, this entails that h is concave and the subgraph $\mathcal{S}$ is a convex set. So the domain $\mathcal{D}$ of the function $\mathcal{B}$ in (4.1) is a convex set since it's a intersection of convex sets.

Now we need to prove that the function $\mathcal{B}$ is concave.
We observe that the function $\mathcal{B}$ is the sum of the functions $(F,f,A,v) \mapsto F$ and $g : (F,f,A,v) \mapsto -\frac{f^p}{(A+v)^{p-1}}$, and the first function is linear (so it is concave) so we only need to prove that the function g is concave. We also observe that $g(F,f,A,v) = -h(f,A+v) = -\tilde{h}(f,A,v)$, here h is the function

$$\begin{aligned} h : \mathbb{R}_0^+ \times \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ (f,Z) &\longmapsto \frac{f^p}{Z^{p-1}} \end{aligned}$$

and $\tilde{h}$ is the function

$$\begin{aligned} h : \mathbb{R}_0^+ \times \mathbb{R}^+ \times \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ (f,A,v) &\longmapsto \frac{f^p}{(A+v)^{p-1}} \end{aligned}$$

So, in order to prove that g is concave, all we need to do is to prove that $\tilde{h}$ is convex. Now we observe that if h is convex then $\tilde{h} : (f,A,v) \mapsto h(f,A+v)$ is convex. Infact, if h is convex, for all $t \in [0,1]$, for all $f_1, f_2 \in \mathbb{R}_0^+$, $A_1, A_2 \in \mathbb{R}^+$, $v_1, v_2 \in \mathbb{R}^+$we get

$$\begin{aligned} &\tilde{h}(tf_1 + (1-t)f_2, tA_1 + (1-t)A_2, tv_1 + (1-t)v_2) = \\ &\quad = h(tf_1 + (1-t)f_2, (tA_1 + (1-t)A_2) + (tv_1 + (1-t)v_2)) = \\ &\quad = h(tf_1 + (1-t)f_2, t(A_1 + v_1) + (1-t)(A_2 + V_2)) \leq \\ &\quad \leq th(f_1, A_1 + v_1) + (1-t)h(f_2, A_2 + v_2) = \\ &\quad = t\tilde{h}(f_1, A_1, v_1) + (1-t)\tilde{h}(f_2, A_2, v_2) \end{aligned}$$

So if h is convex then $\tilde{h}$ is convex, proving that the subgraph $\mathcal{S}$ is a convex set, finishing the proof. All that is left to do is to prove that h is convex.

We compute the Hessian matrix of h. For all $f \in \mathbb{R}^+$, $Z \in \mathbb{R}^+$

$$\frac{\partial h}{\partial f}(f,Z) = p\frac{f^{p-1}}{Z^{p-1}}, \quad \frac{\partial h}{\partial Z}(f,Z) = -(p-1)\frac{f^p}{Z^p}$$

$$\frac{\partial^2 h}{\partial f^2}(f,Z) = p(p-1)\frac{f^{p-2}}{Z^{p-1}}, \quad \frac{\partial^2 h}{\partial Z \partial f}(f,Z) = -p(p-1)\frac{f^{p-1}}{Z^p}$$
$$\frac{\partial^2 h}{\partial f \partial Z}(f,Z) = -p(p-1)\frac{f^{p-1}}{Z^p}, \quad \frac{\partial^2 h}{\partial Z^2}(f,A) = p(p-1)\frac{f^p}{Z^{p+1}}$$

$$H_h(f,Z) = \begin{bmatrix} p(p-1)\frac{f^{p-2}}{Z^{p-1}} & -p(p-1)\frac{f^{p-1}}{Z^p} \\ \\ -p(p-1)\frac{f^{p-1}}{Z^p} & p(p-1)\frac{f^p}{Z^{p+1}} \end{bmatrix} \tag{4.3}$$

Now we compute the eigenvalues of the Hessian matrix (4.3):

$$\begin{aligned} \det(H_h(f,Z) - \lambda I) = &\det \begin{bmatrix} p(p-1)\frac{f^{p-2}}{Z^{p-1}} - \lambda & -p(p-1)\frac{f^{p-1}}{Z^p} \\ \\ -p(p-1)\frac{f^{p-1}}{Z^p} & p(p-1)\frac{f^p}{Z^{p+1}} - \lambda \end{bmatrix} = \\ =&p^2(p-1)^2\frac{f^{2p-2}}{Z^{2p}} - p^2(p-1)^2\frac{f^{2p-2}}{Z^{2p}} \\ &- \lambda\left[p(p-1)\frac{f^{p-2}}{Z^{p-1}} + p(p-1)\frac{f^p}{Z^{p+1}}\right] + \lambda^2 = \\ =&\lambda^2 - \lambda\left[p(p-1)\frac{f^{p-2}}{Z^{p-1}} + p(p-1)\frac{f^p}{Z^{p+1}}\right] \end{aligned}$$

So the eigenvalues of the function h are $\lambda_1 = 0$ and $\lambda_2 = p(p-1)\frac{f^{p-2}}{Z^{p-1}} + p(p-1)\frac{f^p}{Z^{p+1}}$. Since $p > 1$, $f > 0$, $Z > 0$, the eigenvalue λ_2 is greater than 0. So the Hessian matrix $H_h(f,Z)$ is positive semi-definite for every $f > 0$, $Z > 0$. h is continuous up to the boundary of its domain, so h is a convex function, completing the proof that $\mathcal{B}$ is concave.

2) This is trivially true because of the conditions over the domain of the function. Since $p > 1$, $F \geq 0$, $f \geq 0$, $A > 0$, $v \geq A$ and $f^p \leq Fv^{p-1}$, then

$$\mathcal{B}(F,f,A,v) = F - \frac{f^p}{(A+v)^{p-1}} \geq F - \frac{f^p}{(2v)^{p-1}} \geq F - \frac{Fv^{p-1}}{(2v)^{p-1}} \geq F - \frac{F}{2^{p-1}} \geq 0$$

and

$$\mathcal{B}(F,f,A,v) = F - \frac{f^p}{(A+v)^{p-1}} \leq F$$

□

Lemma 4.1.1. *The function $\mathcal{B}$ satisfies*

$$\mathcal{B}(F,f,A,v) - \frac{1}{2}\Big[\mathcal{B}(F_-,f_-,A_-,v_-) + \mathcal{B}(F_+,f_+,A_+,v_+)\Big] \geq \left(\frac{p-1}{2^p}\right)\frac{f^p}{v^p}c$$

where the inequality holds for all

$$F = \tilde{F} + b^p, \qquad f = \tilde{f} + ab$$
$$v = \tilde{v} + a^q, \qquad A = \tilde{A} + c$$

and

$$\tilde{F} = \frac{1}{2}(F_- + F_+), \qquad \tilde{f} = \frac{1}{2}(f_- + f_+)$$
$$\tilde{v} = \frac{1}{2}(v_- + v_+), \qquad \tilde{A} = \frac{1}{2}(A_- + A_+)$$

for every choice of $a \geq 0$, $b \geq 0$, $c \geq 0$. *Here* q *is the real number such as* $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. We start by considering the telescopic sum

$$\mathcal{B}(F, f, A, v) - \mathcal{B}(\tilde{F}, \tilde{f}, A - c, \tilde{v}) = \mathcal{B}(F, f, A, v) - \mathcal{B}(F, f, A - c, v) + \tag{4.4}$$
$$+\mathcal{B}(F, f, A - c, v) - \mathcal{B}(\tilde{F}, \tilde{f}, A - c, \tilde{v})$$

Since the function $\mathcal{B}$ is concave and differentiable over a convex domain, we recall that a concave differentiable function's values are lower or equal to the values of any of its tangent hyperplanes. This entails that, for every g concave and differentiable, for every choice of x, x^* in the domain of the function g:

$$g(x) - g(x^*) \leq \sum_{i=1}^{4} \frac{\partial g(x^*)}{dx_i}(x_i - x_i^*) \tag{4.5}$$

by changing the sign of (4.5) we get

$$g(x^*) - g(x) \geq \sum_{i=1}^{4} \frac{\partial g(x^*)}{dx_i}(x_i^* - x_i) \tag{4.6}$$

So when $g = \mathcal{B}$, $x = (F, f, A, v)$, $x^* = (F, f, \tilde{A}, v) = (F, f, A - c, v)$, the inequality (4.6) becomes

$$\mathcal{B}(F, f, A, v) - \mathcal{B}(F, f, A - c, v) \geq (p - 1)\frac{f^p}{(A + v)^p}c$$

now, since $v \geq A$ by definition of the domain of $\mathcal{B}$, then

$$\mathcal{B}(F, f, A, v) - \mathcal{B}(F, f, A - c, v) \geq \frac{p - 1}{2^p}\frac{f^p}{v^p}c \tag{4.7}$$

By combining (4.7) with (4.4) we get

$$\mathcal{B}(F, f, A, v) - \mathcal{B}(\tilde{F}, \tilde{f}, A - c, \tilde{v}) \geq \mathcal{B}(F, f, A - c, v) - \mathcal{B}(\tilde{F}, \tilde{f}, A - c, \tilde{v}) + \frac{p - 1}{2^p}\frac{f^p}{v^p}c \tag{4.8}$$

Now we consider $g = \mathcal{B}$, $x = (\tilde{F}, \tilde{f}, A - c, \tilde{v})$, $x^* = (F, f, A - c, v)$, the inequality (4.6) becomes

$$\mathcal{B}(F, f, A - c, v) - \mathcal{B}(\tilde{F}, \tilde{f}, A - c, \tilde{v}) \geq b^p - p\left(\frac{f}{A + v - c}\right)^{p-1} ab + (p - 1)\left(\frac{f}{A + v - c}\right)^{p} a^q$$

Now let $y \in \mathbb{R}$ be

$$y = \frac{f}{A + v - c}$$

we observe that $y \geq 0$ because $f \geq 0$, $v > 0$, $A - c > 0$ by definition of the domain of $\mathcal{B}$. So the last inequality can be rewritten as

$$\mathcal{B}(F, f, A - c, v) - \mathcal{B}(\tilde{F}, \tilde{f}, A - c, \tilde{v}) \geq b^p - py^{p-1}ab + (p - 1)y^p a^q = \phi(y)$$

Now we prove that $\phi(y) \geq 0$ for all $y \geq 0$.
We observe that $\phi(y) = b^p \geq 0$ trivially when $a = 0$, so we assume $a > 0$.
We compute the first derivative of the function ϕ:

$$\begin{aligned}\phi'(y) =& p(p-1)a^q y^{p-1} - p(p-1)ab\, y^{p-2} \\ =& p(p-1)y^{p-2}\big(a^q y - ab\big)\end{aligned}$$

So the derivative $\phi'(y)$ is $\phi'(y) \leq 0$ for $0 \leq y \leq \frac{b}{a^{q-1}}$ and $\phi'(y) \geq 0$ for $y \geq \frac{b}{a^{q-1}}$, so $\tilde{y} = \frac{b}{a^{q-1}}$ is a point of absolute minimum, so as long as $\phi(\tilde{y}) \geq 0$ the inequality holds for all $y \geq 0$.

$$\begin{aligned}\phi(\tilde{y}) =& b^p - p\tilde{y}^{p-1}ab + (p-1)\tilde{y}^p a^q = \\ =& b^p - p\left(\frac{b}{a^{q-1}}\right)^{p-1} ab + (p-1)\left(\frac{b}{a^{q-1}}\right)^{p} a^q = \\ =& b^p - p\frac{b^{p-1}}{a^{(q-1)(p-1)}}ab + (p-1)\frac{b^p}{a^{p(q-1)}}a^q = \\ =& b^p - p\frac{b^p}{a^{(pq-p-q)}} + (p-1)\frac{b^p}{a^{(pq-p-q)}}\end{aligned}$$

now we recall that

$$\frac{1}{p} + \frac{1}{q} = 1$$

$$pq = p + q$$

so we get

$$\begin{aligned}\phi(\tilde{y}) =& b^p - p\frac{b^p}{a^0} + (p-1)\frac{b^p}{a^0} = \\ =& \ b^p(1 - p + (p-1)) = 0\end{aligned}$$

So the inequality $\phi(y) \geq 0$ holds for all $y \geq 0$, for every choice $a \geq 0$, $b \geq 0$, therefore the inequality (4.8) becomes

$$\mathcal{B}(F, f, A, v) - \mathcal{B}(\tilde{F}, \tilde{f}, A - c, \tilde{v}) \geq \frac{p-1}{2^p}\frac{f^p}{v^p}c \tag{4.9}$$

now $(\tilde{F}, \tilde{f}, A - c, \tilde{v}) = (\tilde{F}, \tilde{f}, \tilde{A}, \tilde{v}) = \frac{1}{2}((F_+, f_+, A_+, v_+) + (F_-, f_-, A_-, v_-))$, so for the last step we use the fact that $\mathcal{B}$ is concave

$$\mathcal{B}(F, f, A, v) - \frac{1}{2}\Big[\mathcal{B}(F_+, f_+, A_+, v_+) + \mathcal{B}(F_-, f_-, A_-, v_-)\Big] \geq \frac{p-1}{2^p}\frac{f^p}{v^p}c$$

4.2 Hardy's inequality

Let $\mathcal{D}(I_0)$ be the dyadic tree over $I_0 = [0, 1]$, let Λ be a positively valued measure over the dyadic tree defined as follows: for each node $I \in \mathcal{D}(I_0)$

$$\mathcal{D}(I_0) \ni I \longmapsto \lambda_I \in \mathbb{R}^+$$

now we define the following objects as follows:

$$\Lambda(I) = \sum_{K\subseteq I} \lambda_K$$
$$(\Lambda)_I = \frac{1}{|I|}\sum_{K\subseteq I} \lambda_K = \frac{1}{|I|}\Lambda(I)$$
$$\int_I \phi \, d\Lambda = \sum_{K\subseteq I} \phi(K)\lambda_K$$
$$(\phi\Lambda)_I = \frac{1}{|I|}\sum_{K\subseteq I} \phi(K)\lambda_K = \frac{1}{|I|}\int_I \phi \, d\Lambda$$

Now we are going to prove the theorem (1.3) in the article [1, pg. 4] in the general case $p \neq 2$.

Theorem 4.2.1 (Hardy's inequality). *Let $\mathcal{D}(I_0)$ be the dyadic tree originating at I_0 with notations as above, let $\{\alpha_I\}_{I\subseteq I_0}$ be a sequance of positive numbers. Let $\Lambda : \mathcal{D}(I_0) \to \mathbb{R}^+$ be a positive measure over the dyadic tree. Let $\phi : \mathcal{D}(I_0) \to \mathbb{R}_0^+$ be a non-negative function. Let p be a real number $1 < p < +\infty$. Then if the inequality*

$$\frac{1}{|I|}\sum_{K\subseteq I} \alpha_K(\Lambda)_K^p \leq (\Lambda)_I \quad \forall I \in \mathcal{D}(I_0) \tag{4.10}$$

is satisfied, then

$$\frac{1}{|I_0|}\sum_{I\subseteq I_0} \alpha_I(\phi\Lambda^{\frac{1}{q}})_I^p \leq C(p)(\phi^p)_{I_0} \tag{4.11}$$

Here $\frac{1}{p}+\frac{1}{q} = 1$, $C(p) = \frac{2^p}{p-1}$ is a constant depending only on p, and

$$(\phi\Lambda^{\frac{1}{q}})_I = \frac{1}{|I|}\sum_{K\subseteq I} \phi(K)\lambda_K^{\frac{1}{q}}, \quad (\Lambda^p)_{I_0} = \frac{1}{|I_0|}\sum_{I\subseteq I_0} \lambda_I^p$$

Before proving the theorem we are going to mention the reason why Hardy's inequality is useful to study Carleson measures on Besov spaces.
Let $\mathbb{D} = \{z \in \mathbb{C} \mid |z| \leq 1\}$ be the unitary disk in the set of complex numbers $\mathbb{C}$. Let $1 < q < +\infty$ be a real number. Then we define the Besov space $\mathcal{D}_q$ as the set

$$\mathcal{D}_q = \{f : \mathbb{D} \longrightarrow \mathbb{C} \mid f \text{ is holomorphic, } ||f||_{\mathcal{D}_q} < +\infty\}$$

where the norm $||\cdot||_{\mathcal{D}_q}$ is defined by

$$||f||_{\mathcal{D}_q}^q = |f(0)|^q + \int_{\mathbb{D}} |(1-|z|^2)f'(x)|^q \cdot \frac{1}{(1-|z|^2)^2}dz$$

We define the Carleson measures over $\mathbb{D}$ in the following way

Definition 4.2.1. A Carleson measure over $(\mathbb{D}, \mathscr{B}(\mathbb{D}))$ is a measure $\mu : \mathscr{B}(\mathbb{D}) \to \mathbb{R}_0^+$ such that

$$\int_{\mathbb{D}} |f|^q d\mu \leq C(\mu)\cdot ||f||_{\mathcal{D}_q}^q \qquad \text{for all } f \in \mathcal{D}_q$$

where $C(\mu)$ is a real constant depending only on μ.

Carleson measures can be characterized using a discretization theorem in the following way: let

$$Q_{n,j} = \{r \cdot e^{it} \in \mathbb{D} \mid 2^{-n-1} < 1 - r \leq 2^n,\ 2\pi \frac{j-1}{2^n} \leq t < 2\pi \frac{j}{2^n}\}$$

for $n = 0, 1, 2, \dots$ and $j = 1, 2, \dots, 2^n$. The sets $Q_{n,j}$ allow to build a tree structure in this way: $Q_{0,1}$ is the root of the tree, $Q_{1,1}$ is the first son of $Q_{0,1}$ while $Q_{1,2}$ is the second son of $Q_{0,1}$, $Q_{2,1}$ is the first son of $Q_{1,1}$, $Q_{2,2}$ is the second son of $Q_{1,1}$, $Q_{2,3}$ is the first son of $Q_{1,2}$ and so on, increasing the index n increases the generation of the node $Q_{n,j}$. We will denote this dyadic tree by $\mathbb{T}$.
Given a measure μ on $\mathbb{D}$ we may now define a measure on $\mathbb{T}$ by

$$\tilde{\mu} : \mathbb{T} \longrightarrow \mathbb{R}_0^+$$

$$\tilde{\mu}(Q_{n,j}) = \int_{Q_{n,j}} d\mu$$

The following discretization theorem holds true:

Theorem 4.2.2. *Let μ be a measure on $\mathbb{D}$. Let I be the Hardy operator defined by, given $\phi : \mathbb{T} \to \mathbb{R}_0^+$:*

$$I\phi : \mathbb{T} \longrightarrow \mathbb{R}_0^+$$

$$I\phi(Q) = \sum_{R \supseteq Q} \phi(R)$$

Then the following two statemens are equivalent:

1)

$$\int_{\mathbb{D}} |f|^q d\mu \leq C(\mu) \cdot ||f||^q_{\mathcal{D}_q} \quad \text{for all } f \in \mathcal{D}_q$$

2)

$$\sum_{Q \in \mathbb{T}} I\phi(Q)^q \tilde{\mu}(Q) \leq C'(\mu) \cdot ||\phi||^q_{l^q(\mathbb{T})} \quad \text{for all } \phi \in l^q(\mathbb{T})$$

where $C'(\mu)$ is a constant depending only on μ.

It can also be proved that the statement 2) is equivalent to its dual version 3), i.e. 2) is equivalent to

3)

$$\sum_{R \subseteq Q} \Big(\sum_{P \subseteq R} \tilde{\mu}(P) \Big)^p \leq C''(\mu) \sum_{R \subseteq Q} \tilde{\mu}(R) \quad \forall Q \in \mathbb{T}$$

where $C''(\mu)$ is a constant depending only on μ, and $\frac{1}{p} + \frac{1}{q} = 1$.

Details about these topics can be found in [8] and [9].
We observe that Hardy's inequality (4.11) entails the statement 3), because you can prove 3) by considering $\alpha_I = |I|^p$ in (4.11).
This is why Hardy's inequality is useful to study the characterization of Carleson measures for Besov spaces.

Now we will prove theorem 4.2.1 using the Bellman function's method.

Proof. Let $I \in \mathcal{D}(I_0)$, we denote with $I_- \in \mathcal{D}(I_0)$ and $I_+ \in \mathcal{D}(I_0)$ the two children of the node I.
For every $I \in \mathcal{D}(I_0)$ we define

$$
\begin{aligned}
I &\longmapsto v_I \in \mathbb{R}^+\\
I &\longmapsto F_I \in \mathbb{R}_0^+\\
I &\longmapsto f_I \in \mathbb{R}_0^+\\
I &\longmapsto A_I \in \mathbb{R}^+
\end{aligned}
$$

as follows

$$
\begin{aligned}
v_I :=& (\Lambda)_I = \frac{1}{|I|}\lambda_I + \frac{1}{2}(V_{I_-} + V_{I_+}) = \\
=& a_I^q + \tilde{v}_I, \quad \text{where } a_I := \left(\frac{\lambda_I}{|I|}\right)^{\frac{1}{q}}\\
F_I :=& (\phi^p)_I = \frac{1}{|I|}\phi(I)^p + \frac{1}{2}(F_{I_-} + F_{I_+}) = \\
=& b_I^p + \tilde{F}_I, \quad \text{where } b_I := \frac{\phi(I)}{|I|^{\frac{1}{p}}}\\
f_I :=& (\phi\Lambda^{\frac{1}{q}})_I = \frac{\phi(I)\lambda_I^{\frac{1}{q}}}{|I|} + \frac{1}{2}(f_{I_-} + f_{I_+}) = a_I b_I + \tilde{f}_I\\
A_I :=& \frac{1}{|I|}\sum_{K\subseteq I}\alpha_K(\Lambda)_K^p = \frac{\alpha_I(\Lambda)_I^p}{|I|} + \frac{1}{2}(A_{I_-} + A_{I_+}) = \\
=& c_I + \tilde{A}_I, \quad \text{where } c_I := \frac{\alpha_I(\Lambda)_I^p}{|I|}
\end{aligned}
$$

This mapping is the mapping Φ mentioned in subsection 3.2.2.
We observe that the hypothesis (4.10) is exactly $A_I \leq v_I$, and we also observe that, by applying Hölder's inequality to f_I, we get

$$
\begin{aligned}
f_I =& \frac{1}{|I|}\sum_{K\subseteq I}\phi_I\lambda_Y^{\frac{1}{q}} \leq \\
\leq& \frac{1}{|I|^{\frac{1}{p}}}\left(\sum_{K\subseteq I}\phi_I^p\right)^{\frac{1}{p}}\frac{1}{|I|^{\frac{1}{q}}}\left(\sum_{K\subseteq I}\lambda_I\right)^{\frac{1}{q}} = \\
=& (\phi^p)_I^{\frac{1}{p}}(\Lambda)_I^{\frac{1}{q}} = F_I^{\frac{1}{p}}v_I^{\frac{1}{q}}
\end{aligned}
$$

So, for all choices of $\phi : \mathcal{D}(I_0) \to \mathbb{R}_0^+$, $\alpha : \mathcal{D}(I_0) \to \mathbb{R}^+$, $\Lambda : \mathcal{D}(I_0) \to \mathbb{R}^+$, $I \in \mathcal{D}(I_0)$, the vectors

$$
x_I := (F_I, f_I, A_I, v_I), \quad x_{I_-} := (F_{I_-}, f_{I_-}, A_{I_-}, v_{I_-}), \quad x_{I_+} := (F_{I_+}, f_{I_+}, A_{I_+}, v_{I_+})
$$

are elements of the domain of the function $\mathcal{B}$ defined in (4.1). So we can compute the value of the function $\mathcal{B}$ over x_I, x_{I_-}, x_{I_+} for all $I \in \mathcal{D}(I_0)$. We observe that

$$
\mathcal{B}(F_I, f_I, A_I, v_I) = \mathcal{B}(\tilde{F}_I + b_I^p, \tilde{f}_I + a_I b_I, \tilde{A}_I + c_I, \tilde{v}_I + a_I^q)
$$

where

$$\tilde{F} = \frac{1}{2}(F_- + F_+), \qquad \tilde{f} = \frac{1}{2}(f_- + f_+)$$
$$\tilde{v} = \frac{1}{2}(v_- + v_+), \qquad \tilde{A} = \frac{1}{2}(A_- + A_+)$$

so we can apply Lemma 4.1.1 to get

$$|I|\frac{(p-1)}{2^p}\frac{f_I^p}{v_I^p}c_I \leq |I|\left(\mathcal{B}(x_I) - \frac{1}{2}(\mathcal{B}(x_{I_-}) + \mathcal{B}(x_{I_+})\right)$$
$$|I|\frac{(p-1)}{2^p}\frac{f_I^p}{(\lambda)_I^p}\frac{\alpha_I(\lambda)_I^p}{|I|} \leq |I|\mathcal{B}(x_I) - |I_-|\mathcal{B}(x_{I_-}) - |I_+|\mathcal{B}(x_{I_+})$$
$$\frac{(p-1)}{2^p}\alpha_I f_I^p \leq |I|\mathcal{B}(x_I) - |I_-|\mathcal{B}(x_{I_-}) - |I_+|\mathcal{B}(x_{I_+})$$

Summing over all $I \in \mathcal{D}(I_0)$ and using the telescopic nature of the sum we get

$$\frac{(p-1)}{2^p}\sum_{I\subseteq I_0}\alpha_I f_I^p \leq |I_0|\mathcal{B}(F_{I_0}, f_{I_0}, A_{I_0}, v_{I_0}) \leq |I_0|F_{I_0} \tag{4.12}$$

Infact, after ordering the nodes of the dyadic tree in the following way: $I_0 \equiv I_0$, $I_1 := (I_0)_-$, $I_2 := (I_0)_+$, $I_3 := ((I_0)_-)_-$, $I_4 := ((I_0)_-)_+$, $I_5 := ((I_0)_+)_-$, $I_6 := ((I_0)_+)_+$, $I_7 := (((I_0)_-)_-)_-$ and so on, we observe that

$$0 \leq \frac{(p-1)}{2^p}\alpha_I f_I^p \leq |I|\mathcal{B}(x_I) - |I_-|\mathcal{B}(x_{I_-}) - |I_+|\mathcal{B}(x_{I_+})$$

so the sums

$$\frac{(p-1)}{2^p}\sum_{j=0}^{+\infty}\alpha_{I_j} f_{I_j}^p$$

$$\sum_{j=0}^{+\infty}\left(|I_j|\mathcal{B}(x_{I_j}) - |(I_j)_-|\mathcal{B}(x_{(I_j)_-}) - |(I_j)_+|\mathcal{B}(x_{(I_j)_+})\right)$$

converge to a limit in $\mathbb{R}_0^+ \cup \{+\infty\}$ and this limit does not depend of the order of the addends in the sum because all the addends are greater or equal to 0. So, by considering a partial sum of the first $N+1$ addends

$$\frac{(p-1)}{2^p}\sum_{j=0}^{N}\alpha_{I_j} f_{I_j}^p \leq \sum_{j=0}^{N}\left(|I_j|\mathcal{B}(x_{I_j}) - |(I_j)_-|\mathcal{B}(x_{(I_j)_-}) - |(I_j)_+|\mathcal{B}(x_{(I_j)_+})\right)$$

we observe that the last summing is telescopic, and for each $j > 0$ the term $|I_j|\mathcal{B}(x_{I_j})$ addend $|I_j|\mathcal{B}(x_{I_j}) - |(I_j)_-|\mathcal{B}(x_{(I_j)-}) - |(I_j)_+|\mathcal{B}(x_{(I_j)+})$ is simplified with one of the negative terms in the previous addends, so the sum can be written as

$$\frac{(p-1)}{2^p}\sum_{j=0}^{N}\alpha_{I_j} f_{I_j}^p \leq |I_0|\mathcal{B}(x_{I_0}) - \sum_{j\in\mathcal{L}}|I_j|\mathcal{B}(x_{I_j})$$

where $\mathcal{L} \subset \mathbb{N}$ is the (finite) set of the indexes of the negative terms that are not simplified in the partial sum. However we recall that $F \geq \mathcal{B}(F, f, A, v) \geq 0$, so the inequality becomes

$$\frac{(p-1)}{2^p}\sum_{j=0}^{N}\alpha_{I_j} f_{I_j}^p \leq |I_0|\mathcal{B}(x_{I_0}) - \sum_{j\in\mathcal{L}}|I_j|\mathcal{B}(x_{I_j}) \leq |I_0|\mathcal{B}(x_{I_0}) \leq |I_0|F_{I_0}$$

By letting $N \to +\infty$ we get the inequality (4.12). Now we recall that $F_{I_0} = (\phi^p)_{I_0}$ and $f_I = (\phi\Lambda^{\frac{1}{q}})_I$, so we get

$$\frac{1}{|I_0|}\sum_{I\subseteq I_0} \alpha_I (\phi\Lambda^{\frac{1}{q}})_I^p \leq \frac{2^p}{(p-1)}(\phi^p)_{I_0}$$

which is exactly (4.11) , where $C(p) = \frac{2^p}{(p-1)}$, ending the proof. $\square$

Bibliography

[1] Arcozzi N., Holmes I., Mozolyako P., Volberg A. *Bellman function sitting on a tree.* arXiv:1809.03397v2 [math.CA] 18 Dec 2018

[2] Nazarov F., Treil S., Volberg A. *Bellman function in stochastic control and harmonic analysis.* Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000), 393423, Oper. Theory Adv. Appl., 129, Birkhuser, Basel, 2001.

[3] Øksendal B. *Stochastic differential equations.* An introduction with applications. Fifth edition. Universitext. Springer-Verlag, Berlin, 1998.

[4] Krylov, N. V. *Controlled diffusion processes.* Translated from the 1977 Russian original by A. B. Aries. Reprint of the 1980 edition. Stochastic Modelling and Applied Probability, 14. Springer-Verlag, Berlin.

[5] Duren, P. L. *Theory of Hp spaces.* Pure and Applied Mathematics, Vol. 38 Academic Press, New York-London 1970.

[6] Buckley S., *Summation conditions on weights.* Mich. Math. J., 40 (1993), 153-170.

[7] Fefferman R., Kenig C., Pipher J. *The theory of weights and the Dirichlet problem for elliptic equations.* Annals of Math., 134, (1991), 65-124.

[8] Arcozzi N., Rochberg R., Sawyer E. *Carleson measures for analytic Besov spaces.* Rev. Mat. Iberoamericana 18 (2002), no. 2, 443510.

[9] Arcozzi N., Rochberg R., Sawyer E. *The characterization of the Carleson measures for analytic Besov spaces: a simple proof.* Complex and harmonic analysis, 167177, DEStech Publ., Inc., Lancaster, PA, 2007.