Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Control òptim estocàstic

Autor: Armak Karimi González

Director: Dr. Josep Vives i Santa-Eulàlia

Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 24 de gener de 2021

Abstract

The main goal of this project is to study the relationship between two classic methods that are used in solving optimal control problems: Pontryagin's maximum principle and Bellman's dynamic programming. Throughout the project, two different cases are always laid out at the same time: the deterministic case and the stochastic case. Beginning with a stochastic calculus introduction and an exhaustive description of the optimal control problem, the two solving methods are studied separately in order to conclude with a comparison between both of them.

Resum

L'objectiu principal d'aquest treball és estudiar la relació entre els dos mètodes clàssics de resolució d'un problema de control òptim: el principi del màxim de Pontryagin i el principi de la programació dinàmica de Bellman. Al llarg de tot el treball, sempre s'exposen paral·lelament dos casos: el cas determinista i el cas estocàstic. Començant amb una introducció del càlcul estocàstic i una descripció exhaustiva del problema de control òptim, s'estudien separadament els dos camins de resolució per acabar fent una comparació entre ambdós mètodes.

2010 Mathematics Subject Classification. 49K15, 49K20, 49K45, 93E20.

Agraïments

Voldria començar donant les gràcies al meu tutor, el Dr. Josep Vives i Santa-Eulàlia, pel seu compromís amb aquest treball, que ha sigut clau en l'elecció del tema i en les últimes revisions.

En segon lloc, voldria agrair a la meva família, sobretot a la meva mare, el suport rebut en aquests mesos de tanta feina. Aquest treball, que culmina el Grau de Matemàtiques, va dedicat a ella.

Per últim, i no menys important, vull donar-te les gràcies a tu, Mariona, perquè els teus ànims perseverants han sigut el motor més útil en els moments d'absoluta foscor.

Sense tots vosaltres aquest treball no hagués sigut possible, gràcies.

Índex

1 Introducció

1.1 Motivació

Des dels inicis dels meus estudis universitaris el 2015, potser influït pel fet d'estudiar el doble grau Matemàtiques-Física, sempre he tingut més afinitat amb aquelles assignatures del grau de Matemàtiques més inclinades cap a la vessant aplicada. És per aquest motiu que assignatures com Probabilitats, Estadística o Equacions Diferencials han sigut les que més he gaudit al llarg del grau.

En un primer esbòs mental per pensar el tema del present treball i després d'haver realitzat el Treball Final de Grau de Física sobre Teoria Quàntica de Jocs, vaig voler encarar el meu Treball Final de Grau de Matemàtiques sobre la base matemàtica de la Teoria de Jocs. Llegint i documentant-me, els Jocs Diferencials em van cridar molt l'atenció, ja que, com diu el nom, són jocs regits per equacions diferencials; és a dir, per una de les branques de les matemàtiques que més havia gaudit al llarg dels meus estudis. Finalment, investigant sobre aquest tipus de jocs, vaig descobrir que es resolien per mitjà de la teoria del control òptim. Específicament, en els casos més interessants, aquests jocs no són purament deterministes, sinó que vénen definits per variables aleatòries (processos estocàstics) i, per tant, es resolen amb el control òptim estocàstic, teoria que dóna nom al títol del treball.

Aquesta teoria era completament desconeguda per a mi i, per si sola, ja era el suficientment extensa i complexa com per esdevenir el tema del present treball. Va ser just aquí quan va néixer la meva motivació per realitzar el Treball Final de Grau de Matemàtiques sobre el control òptim estocàstic.

1.2 Control òptim

Molts dels processos físics que es desenvolupen en el món tecnològic que ens envolta són, en general, controlables; és a dir, es poden realitzar (o controlar) d'una manera o d'una altra segons quina sigui la voluntat expressa que hi ha al darrere. En aquest context, és lògic qüestionar-se quina ha de ser la millor manera de dur a terme un procés determinat per tal d'obtenir o assolir un cert objectiu o, millor dit, quina ha de ser la millor manera de controlar aquest procés. Per tant, dit amb paraules que ens familiaritzaran més amb aquest treball, és lògic preguntar-se quin ha de ser el control òptim d'un procès.

Quan es parla de control òptim d'un procés, hom fa referència a esbrinar quina és la manera òptima d'assolir l'objectiu desitjat realitzant aquest procés. Aquest objectiu, per exemple, pot anar relacionat amb una minimització del temps de producció o amb una maximització de la quantitat produïda, entre d'altres opcions. Així doncs, seguint amb aquests exemples, la pregunta lògica que pot sorgir darrere d'un procés d'aquest tipus és la de descobrir com es pot optimitzar (és a dir, quin control s'ha d'utilitzar per a optimitzar) el temps o la quantitat de producció d'un procés.

Matemàticament parlant, aquesta explicació ens pot recordar el càlcul variacional, ja que estem demanant trobar una funció (un control d'un procés) que n'optimitzi una altra (per exemple, el temps o la quantitat de producció). Si bé aquesta idea matemàtica del càlcul de variacions no s'allunya gaire de la realitat, en aquest treball s'intentarà estudiar i desenvolupar la teoria del control òptim, una generalització del càlcul variacional.

1.3 Objectiu del treball

Aprofundint una mica en aquesta nova teoria, hi ha dues maneres clàssiques d'abordar un problema de control òptim. Per una banda, es troba el principi del màxim de Pontryagin i, per altra banda, el principi de programació dinàmica de Bellman. Així doncs, just després d'enunciar això, és també molt lògic qüestionar-se quina és la relació que s'estableix entre els dos mètodes que pretenen resoldre el mateix problema, i justament és aquesta la idea que constitueix el nucli d'aquest treball.

Tal i com veurem desenvolupat en els següents capítols, és molt important tenir present els dos camins congruents que conduïran aquest treball, i que coincideixen amb els dos mètodes clàssics de resolució del problema de control òptim. Dins de cadascun d'aquests dos camins, hi ha dos vessants que fan referència a la distinció determinista-estocàstic que es duu a terme en el treball.

D'aquesta manera, prenent el primer camí, quan s'estudiï el principi del màxim de Pontryagin, apareixerà una equació adjunta. Aquesta equació, en el cas determinista, es tractarà d'una EDO (equació diferencial ordinària) i, en el cas estocàstic, es tractarà d'una EDE (equació diferencial estocàstica). El sistema format per aquesta equació adjunta, l'equació original d'estat del procés i una condició de maximalitat (tal i com veurem en capítols posteriors) és el que anomenarem sistema Hamiltonià estès. No obstant, agafant l'altre camí, quan s'estudiï el principi de programació dinàmica de Bellman, apareixerà una EDP (equació diferencial en derivades parcials o, més comunament, equació en derivades parcials), que serà de primer ordre en el cas determinista i de segon ordre en el cas estocàstic. Aquesta equació és el que coneixerem com equació Hamilton-Jacobi-Bellman (HJB).

Aquesta petita explicació permet reformular la pregunta que ens hem fet abans i que és la idea que constitueix el nucli del present treball, de manera que qüestionar-se la relació entre els dos mètodes és totalment equivalent a preguntar-se quina és la relació entre el sistema Hamiltonià estès i les equacions HJB o, de manera més general, quina és la relació que hi ha entre una EDO/EDE i una EDP. Efectivament, intentarem donar respostes a totes aquestes preguntes al llarg del treball.

1.4 Estructura del treball

El meu desig com autor és que el present treball pugui ser entès per qualsevol estudiant que estigui finalitzant el Grau de Matemàtiques, inclús en el cas que no hagi cursat cap assignatura optativa relacionada amb el càlcul estocàstic (com era el meu cas), així que es pretén presentar el treball seguint una estructura totalment ordenada i progressiva en quant a nivell de conceptes, de manera que s'explicaran totes aquelles eines necessàries a la bestreta.

Per poder assolir l'objectiu del treball i poder respondre les preguntes que s'han citat anteriorment, s'ha estructurat la resta del treball de la següent manera:

En el **Capítol 2** es realitza una introducció al càlcul estocàstic, part de la matemàtica molt utilitzada en aquest treball i que, possiblement, un estudiant lector no l'hagi vist mai durant el grau. La idea és que aquest capítol serveixi tant per presentar i definir certs conceptes que es mencionaran al llarg del treball com per assentar les bases necessàries per poder seguir-hi el fil conductor. Tanmateix, com es tracta d'un capítol purament introductori i no és el nucli del treball, es prescindirà de la majoria de les demostracions, bé perquè augmentaria l'extensió i la dificultat del capítol o bé perquè el nivell d'aquest treball no permet realitzar-les.

Seguidament, en el **Capítol 3**, es fa una introducció del marc d'estudi general de la formulació dels dos tipus de problemes de control òptim a què ens enfrontarem al llarg del treball: el problema de control òptim determinista i el problema de control òptim estocàstic. La idea de no passar directament al cas estocàstic i, en el seu defecte, començar introduint la versió determinista rau en el fet que, per si sol, el cas determinista ja té resultats interessants i perquè una primera lectura d'aquest cas pot ajudar a entendre millor el cas estocàstic, gràcies a les relacions i similituds que es poden establir entre ambdós casos.

En el **Capítol 4** i en el **Capítol 5** s'explica, respectivament, cadascun dels dos mètodes clàssics per resoldre un problema de control òptim: el principi del màxim de Pontryagin i el principi de la programació dinàmica de Bellman. L'estructura d'ambdós capítols es basa en una primera part on s'aborda el cas determinista i en una segona part dedicada al cas estocàstic. Un fet important a comentar és que els dos mètodes estan explicats des del mateix punt de vista i partint des del mateix lloc, fet que dóna peu a desenvolupar una bona comparació entre els dos mètodes; és a dir, fet que dóna peu al següent capítol.

En el **Capítol 6** es comparen els dos mètodes estudiats, tant a nivell estocàstic com a nivell determinista, i s'arriba a veure que, sota unes bones condicions donades, el principi del màxim de Pontryagin es pot derivar directament del principi de programació dinàmica de Bellman, sobretot en el cas determinista.

Finalment, i abans d'introduir les referències de la bibliografia, es presenten les **Conclusions** del treball, on es resumeix tot el que s'ha obtingut.

1.5 Referències utilitzades

A l'apartat **Referències** es troba tota la bibliografia emprada per a la realització del present treball. No obstant, a continuació, detallarem quines referències hem utilitzat a cada capítol.

En primer lloc, per escollir l'estructura del treball s'ha seguit [6]. Això ha permès desenvolupar el treball tal com s'ha explicat a la subsecció anterior. Nogensmenys, la dificultat d'aquest llibre era prou elevada i, per aquesta raó, cadascun dels següents apartats s'ha hagut d'anar complementant amb altres referències.

En la introducció matemàtica del càlcul estocàstic, realitzada en el **Capítol 2**, s'ha consultat, principalment, els dos primers capítols de [3].

En el **Capítol 3**, per descriure els dos problemes que ens perseguiran al llarg del treball (determinista i estocàstic), ens hem basat en [6], concretament, en el segon capítol.

Tot seguit, en el primer mètode de resolució d'un problema de control òptim (**Capítol 4**), s'ha de distingir entre les referències emprades en el cas determinista i les consultades en el cas estocàstic. En el primer cas, el llibre de referència ha sigut [4], i concretament, els dos primers capítols, juntament amb el tercer capítol de [6] i el segon capítol de [5]. En canvi, en el segon cas, hem agafat informació tant del tercer capítol de [6] com del quart capítol de [1].

Referent al segon mètode de resolució (**Capítol 5**), s'ha optat, en el cas determinista, pel segon i quart capítol de [2]. Pel que fa al cas estocàstic, s'ha seguit tant el quart capítol de [6] com el tercer capítol de [1].

Finalment, en el **Capítol 6**, per dur a terme la comparativa entre els dos mètodes hem usat totes les referències emprades en cadascun dels dos mètodes per separat, però cal destacar les pàgines 99-103 de [2] en el cas determinista i el cinquè capítol de [6] en el cas estocàstic.

2 Introducció al càlcul estocàstic

És necessari introduir uns conceptes bàsics de processos estocàstics, que ens ajudaran a entendre molt millor tota la part estocàstica; és a dir, les segones subseccions dels següents tres capítols. Per fer-ho, és gairebé obligatori parlar de probabilitats, així que començarem definint unes nocions bàsiques.

Definició 2.1. *Un espai de probabilitat és una terna* $(\Omega, \mathcal{A}, P)$*, on*

(1) Ω *és l'espai mostral: és el conjunt de possibles resultats.*

(2) $\mathcal{A}$ *és una família de parts de* Ω*; és a dir,* $\mathcal{A} \subseteq \mathcal{P}(\Omega)$[1]*, i presenta una estructura de* σ*-àlgebra; és a dir:*

(i) $\Omega \in \mathcal{A}$.

(ii) $A \in \mathcal{A} \implies A^C := \Omega \setminus A \in \mathcal{A}$.

(iii) $\{A_n\}_{n\geq 1} \subseteq \mathcal{A} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ *(la mateixa condició es pot escriure equivalentment amb la intersecció numerable).*

La dupla $(\Omega, \mathcal{A})$ *s'anomena espai mesurable.*

(3) P *és una aplicació* $P : \mathcal{A} \to [0, 1]$*, anomenada probabilitat, tal que:*

(i) $P(\Omega) = 1$.

(ii) P *és* σ*-additiva; és a dir, donats* $\{A_n\}_{n\geq 1} \subseteq \mathcal{A}$ *disjunts dos a dos, es compleix que* $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$.

Definició 2.2. *Donat un espai mesurable* $(\Omega, \mathcal{A})$*, una variable aleatòria* X *és una aplicació mesurable* $X : \Omega \to \mathbb{R}^m$*, on* m *és la dimensió de l'espai on pren valors la variable aleatòria.*

Definició 2.3. *Sigui* $\mathcal{T}$ *un conjunt no buit d'índexs i* $(\Omega, \mathcal{A}, P)$ *un espai de probabilitat. S'anomena procés estocàstic a una família de variables aleatòries* $\{X(t) : t \in \mathcal{T}\}$ *que van de* $(\Omega, \mathcal{A}, P)$ *a* $\mathbb{R}^m$.

Observació 1. Un procés estocàstic també es pot escriure com $\{X(t, \omega) : t \in \mathcal{T}\}$, amb $\omega \in \Omega$, per fer incís en el fet que també depèn d'una segona variable $\omega \in \Omega$. De fet, per qualsevol $\omega \in \Omega$, l'aplicació $t \mapsto X(t, \omega)$, que representa els possibles resultats del procés estocàstic per un $\omega \in \Omega$ determinat, s'anomena trajectòria.

Observació 2. D'ara endavant, prendrem el conjunt $\mathcal{T}$ o bé com $\mathcal{T} = [0, T]$, amb $T > 0$, o bé com $\mathcal{T} = [0, \infty)$. A més, usarem indistingiblament les notacions següents per designar un procés estocàstic: $\{X(t) : t \in \mathcal{T}\}, X(\cdot), X$.

A continuació, introduïm el concepte de filtració, que ens serà molt útil per desenvolupar els espais de probabilitat filtrats.

[1] Donat un conjunt qualsevol C, la notació $\mathcal{P}(C)$ indica el conjunt de les parts de C; és a dir, el conjunt de tots els subconjunts de C.

Definició 2.4. *Donat un espai mesurable* $(\Omega, \mathcal{A})$*, direm que* $\{\mathcal{A}_t\}_{t\geq 0}$ *és una* *filtració de* $\mathcal{A}$ *si és una família creixent*[2] *de sub-*σ*-àlgebres* $\mathcal{A}_t \subseteq \mathcal{A}$*, amb* $t \in [0, T]$*.*

Definició 2.5. *Un espai de probabilitat filtrat és una quaterna* $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$*, on la terna* $(\Omega, \mathcal{A}, P)$ *és un espai de probabilitat i* $\{\mathcal{A}_t\}_{t\geq 0}$ *és una filtració de* $\mathcal{A}$*.*

Definició 2.6. *Sigui* $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ *un espai de probabilitat filtrat i sigui* $X(\cdot)$ *un procés estocàstic. Es diu que* $X(\cdot)$ *és un procés estocàstic* $\{\mathcal{A}_t\}_{t\geq 0}$*-adaptat si* $\forall t \in [0, T]$*, l'aplicació* $\omega \mapsto X(t, \omega)$ *és* $\mathcal{A}_t$*-mesurable.*

Un cop definits els espais de probabilitat filtrats i els processos estocàstics adaptats a una filtració de la σ-àlgebra de l'espai, és hora d'introduir el concepte de moviment Brownià.

Definició 2.7. *Sigui* $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ *un espai de probabilitat filtrat i* $X(\cdot)$ *un procés estocàstic* $\{\mathcal{A}_t\}_{t\geq 0}$*-adaptat. Es diu que* $X(\cdot)$ *és un* $\{\mathcal{A}_t\}_{t\geq 0}$*-moviment Brownià estàndard sobre* $[0, \infty)$ *de dimensió* m *(recordem que la variable aleatòria* $X(t)$ *pren valors en* $\mathbb{R}^m$*) si:*

(i) $P(X(0) = 0) = 1$.

(ii) L'aplicació $t \mapsto X(t)$ *és quasi segurament (q.s.)*[3] *contínua.*

(iii) $\forall\, 0 \leq s < t, X(t) - X(s)$ *és independent de* $\mathcal{A}_s$ *i segueix una distribució normal multivariant amb mitjana* 0 *i covariància* $(t - s)\mathbb{I}_m$*; és a dir, es satisfà que* $X(t) - X(s) \sim \mathcal{N}(0, (t - s)\mathbb{I}_m)$*, on* $\mathbb{I}_m$ *és la matriu identitat d'ordre* m*.*

Observació 3. Exceptuant els casos en què pugui haver confusió, simplement parlarem d'un moviment Brownià m-dimensional en lloc d'un $\{\mathcal{A}_t\}_{t\geq 0}$-moviment Brownià estàndard sobre $[0, \infty)$ de dimensió m. A més a més, generalment, designarem el moviment Brownià per $W(\cdot)$.

Tot seguit, després de presentar els processos estocàstics en un interval $[0, T]$ per un cert $T > 0$ i definir el moviment Brownià, determinarem integrals del tipus

$$\int_0^T f(t)dW(t), \tag{2.1}$$

on $f(\cdot)$ és un procés estocàstic definit en $[0, T]$ per un cert $T > 0$ i $W(\cdot)$ és un moviment Brownià.

Observació 4. Observem que, per tal que la integral anterior estigui ben definida, cal que es satisfaci una certa relació entre les dimensions del procés estocàstic $f(\cdot)$ i les dimensions del moviment Brownià $W(\cdot)$. D'aquesta manera, si volem, per exemple, que la integral sigui de dimensió n, aleshores cal que el procés estocàstic $f(\cdot)$ sigui de la forma $f : [0, T] \times \Omega \to \mathbb{R}^{n\times m}$ i cal que el moviment Brownià $W(\cdot)$ sigui m-dimensional; és a dir, $W : [0, T] \times \Omega \to \mathbb{R}^m$, per a qualsevol valor de $m \geq 1$.

[2] Parlem de família creixent per referir-nos al fet que donades dues sub-σ-àlgebres $\mathcal{A}_{t_1}, \mathcal{A}_{t_2} \subseteq \mathcal{A}$, es satisfà que $\mathcal{A}_{t_1} \subseteq \mathcal{A}_{t_2},\ \forall\, 0 \leq t_1 \leq t_2 \leq T$.

[3] Es diu que un esdeveniment es satisfà quasi segurament (abreujat com q.s.) si el seu complementari està contingut en un conjunt de mesura (o de probabilitat) nul·la. Aquesta noció de "quasi segurament" està fortament lligada a la probabilitat P d'un espai de probabilitat $(\Omega, \mathcal{A}, P)$.

En un primer cop d'ull, podríem pensar, incorrectament, que aquesta integral (2.1) es pot arribar a entendre en el sentit d'una integral de Lebesgue, prenent sempre ω com un paràmetre; i això és el que es podria fer si, $\forall\omega \in \Omega$, l'aplicació $t \mapsto W(t,\omega)$ tingués variació acotada. No obstant, com això no és cert, hem de trobar una nova manera de definir aquesta integral. Això és el que farem per mitjà de la integral d'Itô.

Definició 2.8. *Sigui* $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ *un espai de probabilitat filtrat i* $W(\cdot)$ *un moviment brownià unidimensional definit sobre aquest espai. Considerem* $T > 0$ *i sigui* $f(\cdot)$ *un procés estocàstic* $\{\mathcal{A}_t\}_{t\geq 0}$*-adaptat qualsevol, definit també sobre el mateix espai, de manera que*

$$E\bigg\{\int_0^T f(t,\omega)^2 dt\bigg\} < \infty, \tag{2.2}$$

on E *denota l'esperança (o valor esperat) respecte de la probabilitat* P*. A més a més, sigui* $\mathcal{I}_C$ *la funció indicador d'un cert conjunt* C*, definida com*

$$\mathcal{I}_C(x) = \begin{cases} 1, & \text{si } x \in C, \\ 0, & \text{si } x \notin C. \end{cases} \tag{2.3}$$

Finalment, suposem també que el procés $f(\cdot,\omega)$ *es pot expressar com*

$$f(t,\omega) = f_0(\omega)\mathcal{I}_{\{t=0\}}(t) + \sum_{i\geq 0} f_i(\omega)\mathcal{I}_{(t_i,t_{i+1}]}(t), \quad \forall t \in [0,T], \tag{2.4}$$

on $0 = t_0$ *i on,* $\forall i \geq 0$*, es compleix que* $t_i < t_{i+1}$*, que* $t_i \leq T$ *i que* $f_i(\omega)$ *és* $\mathcal{A}_{t_i}$*-mesurable amb* $\sup\limits_i \sup\limits_{\omega\in\Omega} |f_i(\omega)| < \infty$.

Aleshores, podem definir la integral de Itô $\forall t \in [0,T]$ *com*

$$\int_0^t f(s)dW(s) := \hat{I}(f)(t,\omega) := \sum_{i\geq 0} f_i(\omega)[W(\min(t,t_{i+1}),\omega) - W(\min(t,t_i),\omega)]. \tag{2.5}$$

Observació 5. De manera més general podem definir la integral d'Itô $\forall\, 0 \leq s \leq t \leq T$ com

$$\int_s^t f(r)dW(r) := \hat{I}(f)(t,\omega) - \hat{I}(f)(s,\omega). \tag{2.6}$$

Una vegada definida la integral d'Itô, que ens acompanyarà molt al llarg d'aquest treball, mencionarem les dues propietats més importants d'aquestes integrals.

Proposició 2.9. *Donades dues integrals d'Itô amb el mateix moviment Brownià unidimensional* $W(\cdot)$ *i amb dos processos estocàstics* $f(\cdot)$ *i* $g(\cdot)$ *(tot adient segons la* Definició 2.8 *i l'*Observació 4*), es satisfà que:*

(i) $E\bigg\{\displaystyle\int_s^t f(r)dW(r)\bigg\} = 0, \quad \forall\, 0 \leq s \leq t \leq T.$

(ii) $E\bigg\{\bigg[\displaystyle\int_s^t f(r)dW(r)\bigg]\bigg[\int_s^t g(r)dW(r)\bigg]\bigg\} = E\bigg\{\int_s^t f(r)g(r)dr\bigg\}, \quad \forall\, 0 \leq s \leq t \leq T.$

Demostració. Una demostració completa es pot trobar a [3], a les pàgines 49-51. □

A continuació, tractarem la regla de la cadena, anomenada també Lema d'Itô quan es tracta del cas estocàstic. Tanmateix, en primer lloc, recordarem la regla de la cadena pel cas determinista. Abans de seguir, però, cal introduir una notació que farem servir sovint en les següents pàgines.

Observació 6. Donada una funció $f : \mathbb{R}^n \to \mathbb{R}^m$, per a certs $n, m \geq 1$, que depèn de les variables $x_1, x_2, \ldots, x_n$, definim

$$f_{x_i} := \frac{\partial f}{\partial x_i}, \quad \forall i \in \{1, 2, \ldots, n\}. \tag{2.7}$$

És important no confondre aquesta nova notació amb un subíndex, però pel context i la lletra emprada es pot deduir. No obstant, en cas de dubte ho aclarirem.

Després d'aquesta petita observació per tal d'introduir una nova notació, prosseguim amb la regla de la cadena pel cas determinista. Suposem que tenim una funció $x : \mathbb{R} \to \mathbb{R}^n$ tal que $x(t)$ es pot expressar com

$$x(t) = x(0) + \int_0^t b(s)ds, \quad \forall t \in [0, T] \subseteq \mathbb{R}, \tag{2.8}$$

on $b : [0, T] \to \mathbb{R}^n$ és una funció integrable. Aleshores, per a qualsevol aplicació contínuament diferenciable $F : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ tal que $F \in \mathcal{C}^1([0, T] \times \mathbb{R}^n)$ es satisfà que

$$F(t, x(t)) = F(0, x(0)) + \int_0^t \Big\{ F_t\big(s, x(s)\big) + \big\langle F_x\big(s, x(s)\big), b(s) \big\rangle \Big\} ds, \quad \forall t \in [0, T]. \tag{2.9}$$

Després d'aquest petit recordatori de la regla de la cadena pel cas determinista, passarem a formular-la pel cas estocàstic. Abans, però, definirem un procés d'Itô.

Definició 2.10. *Sigui $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ un espai de probabilitat filtrat i sigui $X(\cdot)$ un procés estocàstic $\{\mathcal{A}_t\}_{t\geq 0}$-adaptat format per una família de variables aleatòries que van de $(\Omega, \mathcal{A}, P)$ a $\mathbb{R}^n$. Direm que $X(\cdot)$ és un procés d'Itô (de dimensió n) si el podem expressar com*

$$X(t) = X(0) + \int_0^t b(s)ds + \int_0^t \sigma(s)dW(s), \quad \forall t \in [0, T], \tag{2.10}$$

on $b(\cdot)$ és un procés estocàstic de la forma $b : [0, T] \times \Omega \to \mathbb{R}^n$ tal que $E\left\{ \int_0^T |b(t, \omega)| dt \right\} < \infty$, i on la segona integral és una integral d'Itô i, per tant, el procés estocàstic $\sigma(\cdot)$, que és de la forma $\sigma : [0, T] \times \Omega \to \mathbb{R}^{n\times m}$, i el moviment brownià m-dimensional $W(\cdot)$ satisfan tot el necessari segons la Definició 2.8.

Observació 7. En al definició anterior (Definició 2.10), hem definit un procés d'Itô de la manera més general. No obstant, nosaltres ens restringirem al cas en què b i σ són variables deterministes i no pas processos aleatoris. D'aquesta manera, tindrem funcions de la forma $b : [0, T] \to \mathbb{R}^n$ i $\sigma : [0, T] \to \mathbb{R}^{n\times m}$, respectivament.

Observació 8. Tal com hem definit el procés d'Itô en la definició anterior (Definició 2.10), observem que es tracta d'un procés continu; és a dir, que té trajectòries contínues. Això es deu al fet que la integral d'una funció integrable (encara que no sigui necessàriament contínua) és contínua. Per tant, observant (2.10), com la suma de funcions contínues és una funció contínua, aleshores podem dir que el procés d'Itô $X(\cdot)$ és un procés continu.

Ara ja podem introduir la regla de la cadena pel cas estocàstic, també anomenada Lema d'Itô.

Lema 2.11. ***(Lema d'Itô).*** *Sigui $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ un espai de probabilitat filtrat i $X(\cdot)$ un procés d'Itô. Sigui $F(t,x) : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ una aplicació de classe $\mathcal{C}^1$ respecte t i de classe $\mathcal{C}^2$ respecte x, de manera que les parcials F_t, F_x, F_{xx} són contínues. Aleshores*

$$\begin{aligned} F(t, X(t)) =& F(0, X(0)) + \int_0^t \Big\{ F_s\big(s, X(s)\big) + \big\langle F_x\big(s, X(s)\big), b(s) \big\rangle \\ &+ \frac{1}{2}\mathrm{tr}\big[\sigma(s)^\top F_{xx}\big(s, X(s)\big)\sigma(s)\big] \Big\} ds + \int_0^t \big\langle F_x\big(s, X(s)\big), \sigma(s)dW(s) \big\rangle, \end{aligned} \tag{2.11}$$

on, donats dos vectors v_1, v_2, la nomenclatura $\langle v_1, v_2\rangle$ denota el producte escalar euclidià convencional entre aquests dos vectors i on, donada una matriu A, la nomenclatura $A^\top$ denota la matriu transposada.

<u>*Demostració.*</u> La demostració general d'aquest teorema s'escapa del nivell d'aquest treball, degut a l'extensió i complexitat de la mateixa. Aquesta demostració completa es pot trobar a [3], a les pàgines 66-73. □

Observació 9. Tal com hem explicat a l'observació anterior (Observació 8), com el procés d'Itô $X(\cdot)$ té trajectòries contínues, aleshores podem dir que $X(s,\omega)$ està fitat en l'interval $[0,t]$ per un cert $t \geq 0$ fixat i per un cert $\omega \in \Omega$ també fixat. Per aquest motiu, la primera integral existeix, ja que cadascun dels tres sumands estan fitats. Per altra banda, la segona integral es defineix tal com s'ha exposat en la Definició 2.8.

Corol·lari 2.12. *Sigui $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ un espai de probabilitat filtrat i siguin Z i $\hat{Z}$ processos continus que prenen valors en $\mathbb{R}^n$ tals que*

$$\begin{cases} Z(t) = Z(0) + \displaystyle\int_0^t b(s)ds + \int_0^t \sigma(s)dW(s), & \forall t \in [0,T], \\ \hat{Z}(t) = \hat{Z}(0) + \displaystyle\int_0^t \hat{b}(s)ds + \int_0^t \hat{\sigma}(s)dW(s), & \forall t \in [0,T], \end{cases} \tag{2.12}$$

on $b : [0,T] \to \mathbb{R}^n, \hat{b} : [0,T] \to \mathbb{R}^n, \sigma : [0,T] \to \mathbb{R}^n, \hat{\sigma} : [0,T] \to \mathbb{R}^n$ són processos mesurables $\{\mathcal{A}_t\}_{t\geq 0}$-adaptats que prenen valors en $\mathbb{R}^n$, i on $W(t)$ és un moviment Brownià unidimensional. Aleshores,

$$\begin{aligned} \big\langle Z(t), \hat{Z}(t) \big\rangle =& \big\langle Z(0), \hat{Z}(0) \big\rangle + \int_0^t \Big\{ \big\langle Z(s), \hat{b}(s) \big\rangle + \big\langle b(s), \hat{Z}(s) \big\rangle + \big\langle \sigma(s), \hat{\sigma}(s) \big\rangle \Big\} ds \\ &+ \int_0^t \Big\{ \big\langle \sigma(s), \hat{Z}(s) \big\rangle + \big\langle Z(s), \hat{\sigma}(s) \big\rangle \Big\} dW(s). \end{aligned} \tag{2.13}$$

<u>*Demostració.*</u> La demostració del corol·lari és immediata després d'aplicar el Lema 2.11 amb $F(x,y) = \langle x, y\rangle$, per un $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, identificant x amb $x = Z(t)$ i y amb

$y = \hat{Z}(t)$. També s'ha de fer servir que, donades dues funcions diferenciables $a(t)$ i $b(t)$, la derivada del producte escalar entre ambdues funcions satisfà que

$$\frac{d}{dt}\Big[\langle a(t), b(t)\rangle\Big] = \left\langle a(t), \frac{d}{dt}b(t)\right\rangle + \left\langle \frac{d}{dt}a(t), b(t)\right\rangle. \tag{2.14}$$

$\square$

3 Problemes de control òptim

De la mateixa manera que hem realitzat una introducció del càlcul estocàstic, ara ja és hora d'entrar en matèria i començarem presentant el marc d'estudi general de la formulació dels problemes de control òptim a què ens enfrontarem. En les següents subseccions presentarem, respectivament, l'entorn de treball d'un problema de control òptim determinista i estocàstic.

3.1 Cas determinista

Tal com hem explicat a la secció **1.2 Control òptim**, molts processos físics es poden controlar, i la idea d'un problema de control òpim és trobar, valgui la redundància, quin és el control òptim d'un procés per tal d'assolir un determinat objectiu (d'ara en endavant, l'objectiu consistirà en la minimització[4] d'una quantitat). Per tant, molt esquemàticament, en un problema de control òptim hi podem distingir tres elements principals: el procés en qüestió (que, com es pot controlar, dependrà del control), el propi control i la funció a minimitzar, que rep el nom de funcional de cost.

Els processos deterministes que estudiarem vénen caracteritzats per una variable (determinista) n-dimensional que depèn del temps, de manera que és possible definir l'estat del procés a cada instant de temps $t \in \mathbb{R}$. Aquesta variable, representada per $x : \mathbb{R} \to \mathbb{R}^n$, s'anomena variable d'estat del procés. A més a més, és usual limitar el temps dels possibles estats d'aquests processos a un interval tancat. Per altra banda, i tal com hem mencionat, el procés també depèn del control, una funció que viu en un espai mètric separable qualsevol.

Així doncs, sigui $T \in (0, \infty)$ un valor donat, i sigui U un espai mètric separable donat amb una funció distància d associada. Sigui $x : [0, T] \to \mathbb{R}^n$ la variable d'estat que caracteritza un procés, sigui $x_0 \in \mathbb{R}^n$ un valor donat i sigui $u : [0, T] \to U$ el control del procés. Suposem que la dinàmica d'aquest procés (i la de tots els processos que estudiarem) ve regida per la següent equació diferencial[5]:

$$\begin{cases} \dot{x}(t) = b(t, x(t), u(t)), & \forall t \in [0, T], \\ x(0) = x_0, \end{cases} \tag{3.1}$$

on $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ és una funció donada.

Aquesta darrera equació (3.1) s'anomena sistema de control i els elements que hi apareixen són: la variable d'estat $x(\cdot) : [0, T] \to \mathbb{R}^n$, la condició inicial $x_0 \in \mathbb{R}^n$ i el control (o funció de control) del sistema $u(\cdot) : [0, T] \to U$. La variable d'estat $x(\cdot)$ és també la solució del sistema (3.1) sota aquest control concret $u(\cdot)$ i, per això, també rep el nom de trajectòria d'estat corresponent al control $u(\cdot)$.

Observació 10. Al llarg del treball, en qualsevol equació diferencial, prendrem el temps inicial com $t_0 = 0$. Es tracta d'una elecció que podem fer sense pèrdua de generalitat, ja que sempre podríem fer una translació de l'eix temporal que preservés totes les propietats.

[4]En tot el treball perseguirem l'objectiu de minimitzar una funció, però si desitgéssim maximitzar-la, es faria de manera completament equivalent pensant en la mateixa funció canviada de signe. I tota la teoria que desenvoluparem seguiria sent igualment vàlida.

[5]Donada una aplicació diferencial k, que depèn d'una variable temporal, la notació $\dot{k}$ representa la derivada total respecte la variable temps; és a dir, $\dot{k} := \frac{dk}{dt}$.

El següent pas és suposar també que, per cada x_0 donat i per cada control $u(\cdot)$, existeix una única solució $x(\cdot) \coloneqq x(\cdot; u(\cdot))$ del sistema anterior (3.1).

Finalment, després de parlar sobre dos dels elements principals d'un problema de control òptim, presentem el funcional de cost; és a dir, la funció que volem minimitzar per assolir un objectiu. Efectivament, el funcional de cost depèn del control que emprem, ja que, precisament, el que volem és trobar el control òptim que minimitza aquesta funció. Per tant, el funcional de cost no és res més que una mesura quantitativa del paper que ha jugat el control durant el procés. El funcional de cost dels processos que estudiarem ve donat per

$$J(u(\cdot)) = \int_0^T f(t, x(t), u(t))dt + h(x(T)), \tag{3.2}$$

on les funcions $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ i $h : \mathbb{R}^n \to \mathbb{R}$ són funcions donades.

Només ens falta definir el següent conjunt per poder postular completament el nostre problema de control òptim pel cas determinista.

Sigui $\mathcal{U}[0, T] \coloneqq \{u : [0, T] \to U | u(\cdot) \text{ és mesurable}\}$. El problema fonamental de control òptim en el cas determinista consisteix en trobar un control òptim $\overline{u}(\cdot) \in \mathcal{U}[0, T]$ tal que

$$J(\overline{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0,T]} J(u(\cdot)). \tag{3.3}$$

Tal com hem dit, tots els controls $\overline{u}(\cdot) \in \mathcal{U}[0, T]$ que satisfan l'equació (3.3) són controls òptims i, conseqüentment, les trajectòries corresponents $\overline{x}(\cdot) \coloneqq x(\cdot; \overline{u}(\cdot))$ s'anomenen trajectòries òptimes.

3.2 Cas estocàstic

Seguint amb la filosofia del que s'ha presentat en la secció anterior **3.1 Cas determinista**, el cas estocàstic es pot explicar de manera anàloga al primer, però amb la principal diferència que la variable d'estat del procés ja no és determinista, sinó que es tracta d'una variable aleatòria i, per tant, porta associada una incertesa. Comparant-ho amb el cas determinista, en un problema de control òptim estocàstic hi trobem els mateixos tres elements principals: el procés en qüestió (que, com es pot controlar, dependrà del control), el propi control i la funció a minimitzar, que, igual que abans, rep el nom de funcional de cost.

Els processos estocàstics que estudiarem vénen caracteritzats per una variable aleatòria n-dimensional que depèn del temps, l'anomenada variable d'estat del procés.

Així doncs, seguint el mateix esquema explicatiu que abans, sigui $T \in (0, \infty)$ un valor donat i sigui U un espai mètric separable donat amb una funció distància d associada. Com a novetat, sigui $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ un espai de probabilitat filtrat i sigui $W(\cdot)$ un moviment Brownià de dimensió m. Seguint com abans, però vigilant que la variable d'estat ara és aleatòria, sigui $x(\cdot) : [0, T] \times \Omega \to \mathbb{R}^n$ la variable d'estat aleatòria que caracteritza un procés, sigui $x_0 \in \mathbb{R}^n$ el valor inicial donat i sigui $u(\cdot) : [0, T] \times \Omega \to U$ el control (o funció de control aleatòria) d'aquest procés. Observem que tant $x(\cdot)$ com $u(\cdot)$ són processos $\{\mathcal{A}_t\}_{t\geq 0}$-adaptats. Finalment, suposem que la dinàmica d'aquest procés (i la de tots els processos estocàstics que estudiarem) ve regida per la següent equació diferencial estocàstica:

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), & \forall t \in [0, T], \\ x(0) = x_0, \end{cases} \tag{3.4}$$

on $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$ són funcions donades.

Observació 11. En el cas més general, en un sistema de control (procés d'Itô) com el de l'equació (3.4), les funcions b i σ podrien ser aleatòries; és a dir, podríem estar parlant de funcions $b : [0, T] \times \mathbb{R}^n \times U \times \Omega \to \mathbb{R}^n$ i $\sigma : [0, T] \times \mathbb{R}^n \times U \times \Omega \to \mathbb{R}^{n\times m}$. No obstant, tal com ja vam dir en l'Observació 7, sempre que treballem en el cas estocàstic, ens limitarem al cas en què b i σ són funcions deterministes. A més a més, sempre considerarem que la condició inicial x_0 és una constant.

Una gran diferència que val la pena comentar és que, abans, el control $u(\cdot)$ era una funció que depenia únicament del temps $t \in [0, T]$, doncs si s'arribava a conèixer la funció control, hom era capaç d'esbrinar quin era el valor d'aquest control a cada instant de temps. Ara, en canvi, només es pot saber la informació del control fins a un moment determinat; és a dir, donat un moment concret, no es pot predir quina serà la funció control per a valors temporals posteriors a aquest moment, degut a la incertesa del sistema, ja que ara $u(\cdot)$ és un procés adaptat que depèn de $\omega \in \Omega$. Per tant, com a conseqüència d'aquest fet, per a un valor de temps t donat, hom no pot escollir quin és el control $u(t)$ fins que s'assoleix aquest temps t.

Abans de continuar explicant el problema de control òptim estocàstic, cal fer una observació per aclarir una notació que apareixerà molt en el desenvolupament d'aquest cas.

Observació 12. D'ara en endavant, les funcions b i σ definides anteriorment les entendrem de la següent manera:

$$\begin{cases} b(t,x,u) = (b^1(t,x,u), \dots, b^n(t,x,u))^\top, \\ \sigma(t,x,u) = (\sigma^1(t,x,u), \dots, \sigma^m(t,x,u)), \\ \sigma^j(t,x,u) = (\sigma^{1j}(t,x,u), \dots, \sigma^{nj}(t,x,u))^\top, \quad \forall\, 1 \le j \le m; \end{cases} \tag{3.5}$$

és a dir, $b^i(t,x,u)$ és la coordenada i-èssima del vector columna $b(t,x,u)$, $\forall\, 1 \le i \le n$, $\sigma^j(t,x,u)$ és la columna j-èssima de la matriu $\sigma(t,x,u)$, $\forall\, 1 \le j \le m$, i $\sigma^{ij}(t,x,u)$ és l'element ij de la matriu $\sigma(t,x,u)$.

Finalment, falta definir, pel cas estocàstic, el tercer element principal d'un problema de control òptim: el funcional de cost. Com ara treballem amb variables aleatòries, per definir-lo cal que fem ús del valor esperat (respecte de la probabilitat P). D'aquesta manera, definirem el funcional de cost com

$$J(u(\cdot)) = E\bigg\{ \int_0^T f(t, x(t), u(t))dt + h(x(T)) \bigg\}, \tag{3.6}$$

on les funcions $f : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ i $h : \mathbb{R}^n \to \mathbb{R}$ són funcions donades.

Igual que en la secció anterior, només ens falta definir el mateix conjunt que abans, però ara pel cas estocàstic, com a pas previ abans de postular completament el nostre problema de control òptim estocàstic.

Sigui $\mathcal{U}[0,T] := \{u : [0,T] \times \Omega \to U | u(\cdot) \text{ és mesurable i és } \{\mathcal{A}_t\}_{t\ge 0}\text{-adaptat}\}$. El problema fonamental de control òptim en el cas estocàstic consisteix en trobar un control òptim $\overline{u}(\cdot) \in \mathcal{U}[0,T]$ tal que

$$J(\overline{u}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}[0,T]} J(u(\cdot)). \tag{3.7}$$

Ja per acabar, comentar que tota la nomenclatura que hem introduït pel cas determinista aplica perfectament en el cas estocàstic, només amb les diferències que ja hem assenyalat.

4 Principi del màxim de Pontryagin

Un cop ja hem explicat a la secció anterior el marc matemàtic sobre el que treballarem, ara ja podem anar directament a estudiar aquest primer mètode de resolució del problema de control òptim. Així doncs, amb la intenció de no dificultar la lectura del treball, no tornarem a definir les funcions, sinó que les presentarem directament.

4.1 Cas determinista

Seguint el formalisme de la secció **3.1 Cas determinista**, considerem el següent sistema de control

$$\begin{cases} \dot{x}(t) = b(t, x(t), u(t)), \quad \forall t \in [0, T] \\ x(0) = x_0, \end{cases} \tag{4.1}$$

amb el respectiu funcional de cost

$$J(u(\cdot)) = \int_0^T f(t, x(t), u(t))dt + h(x(T)). \tag{4.2}$$

Recordem que el nostre objectiu és resoldre el problema fonamental de control òptim determinista, que ve donat per l'equació (3.3). No obstant, abans d'atacar-lo, farem unes suposicions que facilitaran la feina (els supòsits els indicarem amb una lletra majúscula D, per fer referència que es tracta del cas determinista). Així doncs, dins d'aquest primer mètode per a resoldre el problema en qüestió, suposem que:

(D1) Les funcions $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ i $h : \mathbb{R}^n \to \mathbb{R}$ són funcions mesurables, i existeix una constant $L > 0$ i un mòdul de continuïtat[6] $\overline{\omega} : [0, \infty) \to [0, \infty)$ tal que per una funció $\phi(t, x, u) \in \{b(t, x, u), f(t, x, u), h(x)\}$ es satisfà que:

$$\begin{cases} |\phi(t, x, u) - \phi(t, \hat{x}, \hat{u})| \leq L|x - \hat{x}| + \overline{\omega}(d(u, \hat{u})), \quad \forall (t, x, \hat{x}, u, \hat{u}) \in \mathcal{T}, \\ |\phi(t, 0, u)| \leq L, \quad \forall (t, u) \in [0, T] \times U, \end{cases} \tag{4.3}$$

on d és la funció distància associada a l'espai mètric separable U i on $\mathcal{T} = [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times U \times U$.

(D2) Les funcions b, f, h compleixen que $b, f, h \in \mathcal{C}^1$ respecte x, i existeix un mòdul de continuïtat $\overline{\omega} : [0, \infty) \to [0, \infty)$ tal que per una funció $\phi(t, x, u) \in \{b(t, x, u), f(t, x, u), h(x)\}$ es satisfà que:

$$|\phi_x(t, x, u) - \phi_x(t, \hat{x}, \hat{u})| \leq \overline{\omega}(|x - \hat{x}| + d(u, \hat{u})), \quad \forall (t, x, \hat{x}, u, \hat{u}) \in \mathcal{T}. \tag{4.4}$$

Sota aquestes supòsits, ja podem abordar el problema descrit per l'equació (3.3); és a dir, el problema fonamental de control òptim determinista. En primer lloc, observem que,

[6] Un mòdul de continuïtat és una funció $\overline{\omega} : [0, \infty) \to [0, \infty)$ que s'utilitza per a mesurar quantitativament la continuïtat uniforme de funcions, de manera que, donat un interval I qualsevol, una funció $f : I \to \mathbb{R}$ admet $\overline{\omega}$ com a mòdul de continuïtat si, i només si, $|f(x) - f(y)| \leq \overline{\omega}(|x - y|)$, $\forall (x, y) \in I$. A més, un mòdul de continuïtat té la propietat que es fa infinitesimal a mesura que ens apropem a zero; és a dir, $\overline{\omega}(\epsilon) \longrightarrow 0$ quan $\epsilon \longrightarrow 0$. Finalment, com a comentari afegit, es pot demostrar que una funció és uniformement contínua si, i només si, admet un mòdul de continuïtat.

només suposant (D1), per cada x_0 donat i per cada control $u(\cdot) \in \mathcal{U}[0,T]$, existeix una única solució $x(\cdot) := x(\cdot; u(\cdot))$ del sistema anterior (4.1), gràcies al teorema d'existència i unicitat de solucions d'una equació diferencial [veure pàgines 52-54 de [6] per a més detalls]. A més a més, gràcies a aquesta unicitat de la solució $x(\cdot)$, el funcional de cost (4.2) està ben definit.

Tot seguit, amb la idea de resoldre el problema descrit per l'equació (3.3), postularem el principi del màxim de Pontryagin, teorema que dóna nom a aquesta secció i que proporcionarà condicions necessàries de primer ordre per trobar controls òptims i, conseqüentment, trajectòries òptimes.

Teorema 4.1. *Suposem que es satisfà (D1) i (D2). Sigui $\overline{u}(\cdot)$ un control òptim i sigui $\overline{x}(\cdot)$ la corresponent trajectòria òptima. Aleshores, existeix una funció $p(\cdot) : [0,T] \to \mathbb{R}^n$ tal que:*

$$\begin{cases} \dot{p}(t) = -b_x(t, \overline{x}(t), \overline{u}(t))^\top p(t) + f_x(t, \overline{x}(t), \overline{u}(t)), & \forall t \in [0,T], \\ p(T) = -h_x(\overline{x}(T)), \end{cases} \tag{4.5}$$

i que:

$$H(t, \overline{x}(t), \overline{u}(t), p(t)) = \max_{u \in U} H(t, \overline{x}(t), u, p(t)), \quad \forall t \in [0,T], \tag{4.6}$$

on la funció $H : [0,T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \to \mathbb{R}$ es defineix com:

$$H(t, x, u, p) := \langle p, b(t, x, u) \rangle - f(t, x, u). \tag{4.7}$$

La funció $p(\cdot)$ s'anomena funció adjunta i l'equació (4.5) és l'equació adjunta. En ambdues definicions hauríem d'afegir "corresponent al control òptim $\overline{u}(\cdot)$ i a la trajectòria òptima $\overline{x}(\cdot)$", però ho obviem per no carregar massa les frases. Per altra banda, la funció H s'anomena Hamiltonià.

Fent ús d'aquesta nova funció H, podem reescriure l'equació d'estat (4.1), la corresponent equació adjunta (4.5) i la condició de maximalitat (4.6) com

$$\begin{cases} \dot{x}(t) = H_p(t, x(t), u(t), p(t)), & \forall t \in [0,T], \\ \dot{p}(t) = -H_x(t, x(t), u(t), p(t)), & \forall t \in [0,T], \\ x(0) = x_0, \quad p(T) = -h_x(x(T)), \\ H(t, x(t), u(t), p(t)) = \max_{u \in U} H(t, x(t), u, p(t)), & \forall t \in [0,T], \end{cases} \tag{4.8}$$

El sistema anterior (4.8) és l'anomenat sistema Hamiltonià estès.

A continuació, demostrarem el Teorema 4.1. Per fer-ho, haurem de tenir present, per una banda, el desenvolupament en Taylor de la trajectòria d'estat i del funcional de cost (respecte d'una pertorbació en la variable de control), i per altra banda, la dualitat, per mitjà de la funció H, entre l'equació diferencial d'estat (4.1) i l'equació adjunta (4.5).

Així doncs, comencem tractant la pertorbació en el control. Sigui $\overline{u}(\cdot)$ un control òptim donat i $\overline{x}(\cdot)$ la corresponent trajectòria òptima. Prenem $\epsilon > 0$ i sigui $E_\epsilon \subseteq [0,T]$ un conjunt mesurable amb mesura de Lebesgue $|E_\epsilon| = \epsilon$. Sigui $u(\cdot) \in \mathcal{U}[0,T]$ un control qualsevol i definim

$$u^\epsilon(t) = \begin{cases} \overline{u}(t), & \text{si } t \in [0,T] \setminus E_\epsilon \\ u(t), & \text{si } t \in E_\epsilon. \end{cases} \tag{4.9}$$

Aquesta nova funció $u^\epsilon(\cdot)$ representa una petita variació (o pertorbació) del control $\overline{u}(\cdot)$ i, òbviament, satisfà que $u^\epsilon(t) \in \mathcal{U}[0,T]$.

No obstant, abans de demostrar el Teorema 4.1, és necessari fer ús d'un lema previ que ens serà de molta utilitat.

Lema 4.2. *Suposem que es satisfà (D1) i (D2). Definim la funció $x^\epsilon(\cdot)$ com $x^\epsilon(\cdot) := x(\cdot; u^\epsilon(\cdot))$; és a dir, com la solució del sistema (4.1) sota el control $u^\epsilon(\cdot)$. A més, sigui $y^\epsilon(\cdot)$ la solució de l'equació diferencial següent*[7]*:*

$$\begin{cases} \dot{y}^\epsilon(t) = b_x(t,\overline{x}(t),\overline{u}(t))y^\epsilon(t) + \{b(t,\overline{x}(t),u(t)) - b(t,\overline{x}(t),\overline{u}(t))\}\mathcal{I}_{E_\epsilon}(t), \quad \forall t \in [0,T], \\ y^\epsilon(0) = 0. \end{cases} \tag{4.10}$$

Aleshores, es satisfà:

$$\begin{cases} \max_{t\in[0,T]} |x^\epsilon(t) - \overline{x}(t)| = O(\epsilon), \\ \max_{t\in[0,T]} |y^\epsilon(t)| = O(\epsilon), \end{cases} \tag{4.11}$$

$$\max_{t\in[0,T]} |x^\epsilon(t) - \overline{x}(t) - y^\epsilon(t)| = o(\epsilon), \tag{4.12}$$

i

$$\begin{aligned} J(u^\epsilon(\cdot)) - J(\overline{u}(\cdot)) = &\langle h_x(\overline{x}(T)), y^\epsilon(T)\rangle + \int_0^T \Big\{ \big\langle f_x(t,\overline{x}(t),\overline{u}(t)), y^\epsilon(t)\big\rangle \\ &+ \big\{ f(t,\overline{x}(t),u(t)) - f(t,\overline{x}(t),\overline{u}(t))\big\}\mathcal{I}_{E_\epsilon}(t) \Big\} dt + o(\epsilon). \end{aligned} \tag{4.13}$$

Per demostrar aquest lema es necessita fer ús d'un lema molt conegut en la teoria de les equacions diferencials, el lema de Gronwall. Aquest lema només el postularem, ja que la seva demostració no té relació amb el fil conductor d'aquest treball.

Lema 4.3. *Sigui $z : I \to \mathbb{R}$ una funció contínua i no negativa en un interval qualsevol I, i sigui $t_0 \in I$. Suposem que, $\forall t \in I$, es té que*

$$z(t) \leq A + \left| \int_{t_0}^t Lz(s)ds \right|, \tag{4.14}$$

on A i L són constants no negatives. Aleshores, $\forall t \in I$, es satisfà que

$$z(t) \leq Ae^{L|t-t_0|}. \tag{4.15}$$

Ara ja podem seguir amb la demostració del lema que havíem postulat anteriorment, el Lema 4.2.

<u>*Demostració del Lema 4.2.*</u> Com sabem que es satisfà (D1), podem fixar-nos en la primera desigualtat de (4.3) pel cas en què $\phi(t,x,u) = b(t,x,u)$ quan $x = x^\epsilon(t)$, $u = u^\epsilon(t)$, $\hat{x} = \overline{x}(t)$ i $\hat{u} = \overline{u}(t)$. És a dir,

$$|b(t,x^\epsilon(t),u^\epsilon(t)) - b(t,\overline{x}(t),\overline{u}(t))| \leq L|x^\epsilon(t) - \overline{x}(t)| + \overline{\omega}(d(u^\epsilon(t),\overline{u}(t))). \tag{4.16}$$

[7]En l'equació diferencial (4.10) i en demostracions posteriors, apareix la funció indicador, que ja l'hem definida amb anterioritat dins de la Definició 2.8, concretament en (2.3).

Recordant la relació (4.1), si integrem a banda i banda la desigualtat anterior (4.16), obtenim que

$$|x^\epsilon(t) - \overline{x}(t)| \leq \int_0^t L|x^\epsilon(s) - \overline{x}(s)|ds + \int_0^t \overline{\omega}(d(u^\epsilon(s), \overline{u}(s)))ds, \quad \forall t \in [0, T]. \tag{4.17}$$

Sigui $z^\epsilon(t) := x^\epsilon(t) - \overline{x}(t)$. Tenint en ment que un mòdul de continuïtat es fa infinitesimal quan ens apropem a zero (doncs el podem acotar per ϵ) i usant la definició de $z^\epsilon(t)$, podem reescriure l'equació (4.17) com:

$$|z^\epsilon(t)| \leq \int_0^t L|z^\epsilon(s)|ds + K\epsilon, \quad \forall t \in [0, T], \tag{4.18}$$

on $K > 0$ és una constant genèrica.

Tot seguit, si apliquem la desigualtat de Gronwall (Lema 4.3), obtindrem que

$$|z^\epsilon(t)| \leq K\epsilon e^{Lt}, \ \ \forall t \in [0, T]. \tag{4.19}$$

Per tant, com a mínim, sabem que

$$\max_{t\in[0,T]} |z^\epsilon(t)| \leq K\epsilon e^{LT} = \tilde{K}\epsilon, \tag{4.20}$$

on $\tilde{K} = Ke^{LT} > 0$ és una nova constant. Així doncs, es conclou que

$$\max_{t\in[0,T]} |z^\epsilon(t)| = O(\epsilon), \tag{4.21}$$

i queda demostrada la primera igualtat de l'equació (4.11). La segona igualtat de l'equació (4.11) es demostra de manera similar.

Tot seguit, definim $w^\epsilon(t) := z^\epsilon(t) - y^\epsilon(t) = x^\epsilon(t) - \overline{x}(t) - y^\epsilon(t)$. Aleshores $\dot{w}^\epsilon(t) = \dot{x}^\epsilon(t) - \dot{\overline{x}}(t) - \dot{y}^\epsilon(t)$. Usant les equacions (4.1) i (4.10), escriurem que

$$\begin{aligned} \dot{w}^\epsilon(t) =& b(t, x^\epsilon(t), u^\epsilon(t)) - b(t, \overline{x}(t), \overline{u}(t)) - b_x(t, \overline{x}(t), \overline{u}(t))y^\epsilon(t) \\ &- \{b(t, \overline{x}(t), u(t)) - b(t, \overline{x}(t), \overline{u}(t))\}\mathcal{I}_{E_\epsilon}(t). \end{aligned} \tag{4.22}$$

Per altra banda, observem que, de manera molt abreujada, podem escriure que

$$\begin{aligned} b(t, x^\epsilon(t), u^\epsilon(t)) - & b(t, \overline{x}(t), \overline{u}(t)) - \{b(t, \overline{x}(t), u(t)) - b(t, \overline{x}(t), \overline{u}(t))\}\mathcal{I}_{E_\epsilon}(t) \\ &= \int_0^1 b_x(t, \overline{x}(t) + \theta z^\epsilon(t), u^\epsilon(t))d\theta \cdot z^\epsilon(t). \end{aligned} \tag{4.23}$$

Per tant, combinant les relacions anteriors (4.22) i (4.23), tenim que

$$\begin{aligned} \dot{w}^\epsilon(t) =& \int_0^1 b_x(t, \overline{x}(t) + \theta z^\epsilon(t), u^\epsilon(t))d\theta \cdot z^\epsilon(t) - b_x(t, \overline{x}(t), \overline{u}(t))y^\epsilon(t) \\ =& \left[\int_0^1 \big\{b_x(t, \overline{x}(t) + \theta z^\epsilon(t), u^\epsilon(t)) - b_x(t, \overline{x}(t), u^\epsilon(t))\big\}d\theta \right] z^\epsilon(t) \\ &+ b_x(t, \overline{x}(t), \overline{u}(t))w^\epsilon(t) + \{b_x(t, \overline{x}(t), u(t)) - b_x(t, \overline{x}(t), \overline{u}(t))\}z^\epsilon(t)\mathcal{I}_{E_\epsilon}(t), \end{aligned} \tag{4.24}$$

on en la segona igualtat s'ha emprat que $y^\epsilon(t) = z^\epsilon(t) - w^\epsilon(t)$ i s'han reordenat termes.

Arribats a aquest punt, com sabem que es satisfà (D2), podem fixar-nos en l'equació (4.4) pel cas $\phi(t,x,u) = b(t,x,u)$ (de manera similar al que hem fet a les equacions (4.16), (4.17) i (4.18)). A més a més, usant els resultats ja demostrats de (4.11), podem escriure que

$$|w^\epsilon(t)| \leq \int_0^t L|w^\epsilon(s)|ds + K\epsilon \int_0^T \overline{\omega}(|z^\epsilon(s)|)ds + K\epsilon^2, \quad \forall t \in [0,T]. \tag{4.25}$$

Igual que abans, apliquem el lema de Gronwall, però amb la diferència que ara hi ha una contribució quadràtica d'ϵ. Per tant, en dividir per ϵ, segueix quedant un terme proporcional a ϵ que tendeix a zero. Així doncs, hem demostrat, tal com volíem, que

$$\max_{t\in[0,T]} |w^\epsilon(t)| = o(\epsilon). \tag{4.26}$$

De manera similar, es pot demostrar l'equació (4.13). □

Ara, gràcies al lema anterior, ja estem capacitats per demostrar el Teorema 4.1.

Demostració del Teorema 4.1. Sigui $p(\cdot)$ la solució del sistema (4.5). Aplicant la regla de la cadena a $\langle p(t), y^\epsilon(t)\rangle$, obtenim la següent relació de dualitat (és la relació de dualitat entre $y^\epsilon(t)$ de l'equació variacional (4.10) i $p(\cdot)$ de l'equació adjunta (4.5)):

$$\begin{aligned} -\langle h_x(\overline{x}(T)), y^\epsilon(T)\rangle = \langle p(T), y^\epsilon(T)\rangle - \langle p(0), y^\epsilon(0)\rangle = \int_0^T \Big\{ \langle f_x(t,\overline{x}(t),\overline{u}(t)), y^\epsilon(t)\rangle \\ + \langle p(t), b(t,\overline{x}(t),u(t)) - b(t,\overline{x}(t),\overline{u}(t))\rangle \mathcal{I}_{E_\epsilon}(t) \Big\} dt. \end{aligned} \tag{4.27}$$

A continuació, fixem un temps $\bar{t} \in [0,T)$ i un control $u \in U$. Sigui $u(t) \equiv u$ i prenem $E_\epsilon = [\bar{t}, \bar{t}+\epsilon]$ amb $\epsilon > 0$ suficientment petit per tal que $E_\epsilon \subseteq [0,T]$. Finalment, combinant l'equació anterior (4.27) amb l'equació (4.13) i tenint present que $\overline{u}(\cdot)$ és un control òptim, obtenim que

$$\begin{aligned} 0 \leq J(u^\epsilon(\cdot)) - J(\overline{u}(\cdot)) &= \langle h_x(\overline{x}(T)), y^\epsilon(T)\rangle + \int_0^T \Big\{ \big\langle f_x(t,\overline{x}(t),\overline{u}(t)), y^\epsilon(t)\big\rangle \\ &+ \big\{ f(t,\overline{x}(t),u(t)) - f(t,\overline{x}(t),\overline{u}(t))\big\} \mathcal{I}_{E_\epsilon}(t) \Big\} dt + o(\epsilon) \\ &= -\int_{\bar{t}}^{\bar{t}+\epsilon} \Big\{ H(t,\overline{x}(t),u(t),p(t)) - H(t,\overline{x}(t),\overline{u}(t),p(t)) \Big\} dt + o(\epsilon). \end{aligned} \tag{4.28}$$

Partint d'aquesta darrera equació (4.28), si observem que

$$0 \leq -\int_{\bar{t}}^{\bar{t}+\epsilon} \Big\{ H(t,\overline{x}(t),u(t),p(t)) - H(t,\overline{x}(t),\overline{u}(t),p(t)) \Big\} dt + o(\epsilon), \tag{4.29}$$

aleshores podem deduir, juntament amb la condició de separabilitat de l'espai mètric U, que

$$H(t,\overline{x}(t),\overline{u}(t),p(t)) \geq H(t,\overline{x}(t),u(t),p(t)), \quad \forall u \in U; \tag{4.30}$$

és a dir, podem deduir la condició de maximalitat (4.6), tal com volíem demostrar. □

Així doncs, per resumir, acabem de demostrar el principi del màxim de Pontryagin, que ens dóna una condició necessària d'optimalitat. Ho hem fet a partir d'un desenvolupament en sèrie de Taylor i per mitjà de la relació de dualitat entre l'equació variacional i l'equació adjunta. Concretament, com $\mathcal{U}[0,T]$ no presenta, en general, una estructura lineal, el desenvolupament en sèrie de Taylor s'ha hagut de fer d'una manera especial (equació (4.13)), de manera que hem obtingut la condició necessària d'optimalitat en termes de la solució de l'equació variacional $y^\epsilon(\cdot)$ (que, a la vegada, depèn de l'elecció de $u(\cdot) \in \mathcal{U}[0,T]$). Aquesta dependència amb $y^\epsilon(\cdot)$ fa que aquesta condició necessària implícita (4.13) no sigui fàcil de manipular. És per aquest motiu que, gràcies a la relació de dualitat donada per l'equació (4.27), ens podem desfer de $y^\epsilon(\cdot)$ i presentar el resultat tal com desitgem (equació (4.30)).

Fins ara el que hem fet ha sigut donar condicions necessàries per un control òptim. Nogensmenys, encara no sabem si, donat un control que satisfà aquestes condicions necessàries, aquest control és, en realitat, òptim o no. Amb l'objectiu d'estudiar això; és a dir, de trobar les condicions suficients d'optimalitat, suposarem que:

(D3) La regió de control U és convexa amb l'interior no buit, i les aplicacions b i f són localment Lipschitz respecte u.

Amb aquest supòsit en joc, podem postular el següent teorema, que proporciona condicions suficients d'optimalitat.

Teorema 4.4. *Suposem que es satisfà (D1), (D2) i (D3). Sigui $\overline{u}(\cdot)$ un control qualsevol i sigui $\overline{x}(\cdot)$ la corresponent trajectòria. Sigui $p(\cdot)$ la corresponent variable adjunta. Suposem que $H(t,\cdot,\cdot,p(t))$ és còncava*[8] *$\forall t \in [0,T]$ i que $h(\cdot)$ és convexa. Aleshores $\overline{u}(\cdot)$ i $\overline{x}(\cdot)$ són òptims si*

$$H(t,\overline{x}(t),\overline{u}(t),p(t)) = \max_{u\in U} H(t,\overline{x}(t),u,p(t)), \quad \forall t \in [0,T]. \tag{4.31}$$

<u>*Demostració.*</u> Tal com ens diu el peu de pàgina, per la concavitat de $H(t,\cdot,\cdot,p(t))$, tenim que, $\forall u(\cdot) \in \mathcal{U}[0,T]$,

$$\int_0^T \Big\{ H(t,x(t),u(t),p(t)) - H(t,\overline{x}(t),\overline{u}(t),p(t)) \Big\} dt \le \int_0^T \langle H_x(t,\overline{x}(t),\overline{u}(t),p(t)), x(t)-\overline{x}(t) \rangle dt, \tag{4.32}$$

on $x(\cdot)$ és la trajectòria corresponent al control $u(\cdot)$.

Sigui $z(t) = x(t) - \overline{x}(t)$ tal que

$$\begin{cases} \dot{z}(t) = b_x(t,\overline{x}(t),\overline{u}(t))z(t) + \alpha(t), & \forall t \in [0,T], \\ z(0) = 0, \end{cases} \tag{4.33}$$

on

$$\alpha(t) = -b_x(t,\overline{x}(t),\overline{u}(t))z(t) + b(t,x(t),u(t)) - b(t,\overline{x}(t),\overline{u}(t)). \tag{4.34}$$

Per la relació de dualitat entre el sistema variacional (4.33) i el sistema adjunt (4.5),

[8] Es diu que una funció f d'una variable és còncava en un interval donat $\mathcal{I}$ si, per tot punt d'aquest interval, la recta tangent a f que passa per aquest punt està sempre per sobre de la gràfica. Si, a més a més, f és diferenciable, aleshores es satisfà que $f(y) - f(x) \le f'(x)(y-x)$ $\forall x,y \in \mathcal{I}$. Una funció convexa és el contrari d'una funció còncava.

tenim que

$$\langle h_x(\overline{x}(T)), z(T)\rangle = -\langle p(T), z(T)\rangle + \langle p(0), z(0)\rangle = -\int_0^T \big\{\langle f_x(t,\overline{x}(t),\overline{u}(t)), z(t)\rangle + \langle p(t), \alpha(t)\rangle\big\}dt$$
$$= \int_0^T \langle H_x(t,\overline{x}(t),\overline{u}(t),p(t)), z(t)\rangle dt - \int_0^T \langle p(t), b(t,x(t),u(t)) - b(t,\overline{x}(t),\overline{u}(t))\rangle dt. \tag{4.35}$$

Unint l'equació anterior (4.35) amb l'equació (4.32) de la concavitat de $H(t,\cdot,\cdot,p(t))$, tenim que

$$\langle h_x(\overline{x}(T)), z(T)\rangle = \int_0^T \langle H_x(t,\overline{x}(t),\overline{u}(t),p(t)), z(t)\rangle dt - \int_0^T \langle p(t), b(t,x(t),u(t)) - b(t,\overline{x}(t),\overline{u}(t))\rangle dt$$
$$\geq \int_0^T \Big\{H(t,x(t),u(t),p(t)) - H(t,\overline{x}(t),\overline{u}(t),p(t))\Big\}dt - \int_0^T \langle p(t), b(t,x(t),u(t)) - b(t,\overline{x}(t),\overline{u}(t))\rangle dt$$
$$= -\int_0^T \{f(t,x(t),u(t)) - f(t,\overline{x}(t),\overline{u}(t))\}dt. \tag{4.36}$$

Per altra banda, la convexitat de la funció $h(\cdot)$, ens diu que

$$h(x(T)) - h(\overline{x}(T)) \geq \langle h_x(\overline{x}(T)), x(T) - \overline{x}(T)\rangle = \langle h_x(\overline{x}(T)), z(T)\rangle. \tag{4.37}$$

Així doncs, ja per acabar, combinant l'equació (4.36) amb l'equació (4.37), obtenim que

$$-\int_0^T \{f(t,x(t),u(t)) - f(t,\overline{x}(t),\overline{u}(t))\}dt \leq \langle h_x(\overline{x}(T)), z(T)\rangle \leq h(x(T)) - h(\overline{x}(T)). \tag{4.38}$$

Reagrupant els termes dels extrems, obtenim que

$$J(\overline{u}(\cdot)) = \int_0^T f(t,\overline{x}(t),\overline{u}(t))dt + h(\overline{x}(T)) \leq J(u(\cdot)) = \int_0^T f(t,x(t),u(t))dt + h(x(T)); \tag{4.39}$$

és a dir, hem obtingut que, $\forall u(\cdot) \in U$,

$$J(\overline{u}(\cdot)) \leq J(u(\cdot)). \tag{4.40}$$

Per tant, concloem que $\overline{u}(\cdot)$ és òptim, tal com volíem demostrar. $\square$

4.2 Cas estocàstic

Seguint el formalisme de la secció **3.2 Cas estocàstic**, considerem el següent sistema de control

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), \quad \forall t \in [0, T] \\ x(0) = x_0, \end{cases} \tag{4.41}$$

amb el respectiu funcional de cost

$$J(u(\cdot)) = E\bigg\{ \int_0^T f(t, x(t), u(t))dt + h(x(T)) \bigg\}. \tag{4.42}$$

Recordem que el nostre objectiu és resoldre el problema fonamental de control òptim estocàstic, que ve donat per l'equació (3.7). No obstant, abans d'atacar-lo, farem unes suposicions que facilitaran la feina (els supòsits els indicarem amb una lletra majúscua E per fer referència que es tracta del cas estocàstic). Per tant, suposem que:

(E1) Les funcions $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$, $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ i $h : \mathbb{R}^n \to \mathbb{R}$ són funcions mesurables, i existeix una constant $L > 0$ i un mòdul de continuïtat $\varpi : [0, \infty) \to [0, \infty)$ tal que per una funció $\phi(t, x, u) \in \{b(t, x, u), \sigma(t, x, u), f(t, x, u), h(x)\}$ es satisfà que:

$$\begin{cases} |\phi(t, x, u) - \phi(t, \hat{x}, \hat{u})| \leq L|x - \hat{x}| + \varpi(d(u, \hat{u})), \ \forall (t, x, \hat{x}, u, \hat{u}) \in \mathcal{T}, \\ |\phi(t, 0, u)| \leq L, \ \forall (t, u) \in [0, T] \times U, \end{cases} \tag{4.43}$$

on d és la funció distància associada a l'espai mètric separable U i on $\mathcal{T} = [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times U \times U$.

(E2) Les funcions b, σ, f, h compleixen que $b, \sigma, f, h \in \mathcal{C}^2$ respecte x. A més, existeix una constant $L > 0$ i un mòdul de continuïtat $\varpi : [0, \infty) \to [0, \infty)$ tal que per una funció $\phi(t, x, u) \in \{b(t, x, u), \sigma(t, x, u), f(t, x, u), h(x)\}$ es satisfà que:

$$\begin{cases} |\phi_x(t, x, u) - \phi_x(t, \hat{x}, \hat{u})| \leq L|x - \hat{x}| + \varpi(d(u, \hat{u})), \ \forall (t, x, \hat{x}, u, \hat{u}) \in \mathcal{T}, \\ |\phi_{xx}(t, x, u) - \phi_{xx}(t, \hat{x}, \hat{u})| \leq \varpi(|x - \hat{x}| + d(u, \hat{u})), \ \forall (t, x, \hat{x}, u, \hat{u}) \in \mathcal{T}. \end{cases} \tag{4.44}$$

Sota aquestes supòsits, ja podem abordar el problema descrit per l'equació (3.7); és a dir, el problema fonamental de control òptim estocàstic. En primer lloc, observem que, només suposant (E1), per cada x_0 donat i per cada control $u(\cdot) \in \mathcal{U}[0, T]$, existeix una única solució $x(\cdot) \coloneqq x(\cdot; u(\cdot))$ del sistema anterior (4.41), gràcies al teorema d'existència i unicitat de solucions d'una equació diferencial estocàstica [veure Teorema 2.3 de la pàgina 173 i Teorema 3.1 de la pàgina 178 de [3] per a més detalls sobre l'existència i la unicitat, respectivament]. A més a més, gràcies a aquesta unicitat de la solució $x(\cdot)$, el funcional de cost (4.42) està ben definit.

Hem vist que, en el cas determinista, la variable adjunta $p(\cdot)$ jugava un rol molt important en la formulació del principi del màxim de Pontryagin. Aquesta variable adjunta es regeix per l'equació adjunta (4.5), que es tracta d'una equació diferencial ordinària amb una condició final[9].

[9]El fet que l'equació diferencial ordinària presenti una condició final no és gens rellevant en el cas determinista, ja que, invertint el temps, podem convertir la condició final en condició inicial. No obstant, en el cas estocàstic, això no és possible, ja que violaríem el fet de no poder anticipar-se a una solució donada degut a la incertesa del sistema.

En el cas estocàstic també hi ha una variable adjunta amb la seva equació adjunta, tot i que és lleugerament diferent. Així doncs, la funció adjunta estocàstica $p(\cdot)$ ve regida pel següent problema de valors finals amb la següent equació diferencial estocàstica:

$$\begin{cases} dp(t) = -\Big\{ b_x(t, \overline{x}(t), \overline{u}(t))^\top p(t) + \sum_{j=1}^{m} \sigma_x^j(t, \overline{x}(t), \overline{u}(t))^\top q_j(t) \\ \qquad\qquad - f_x(t, \overline{x}(t), \overline{u}(t)) \Big\} dt + q(t) dW(t), \quad \forall t \in [0, T], \\ p(T) = -h_x(\overline{x}(T)). \end{cases} \tag{4.45}$$

Aquí les incògnites són els dos processos $\{\mathcal{A}_t\}_{t \geq 0}$-adaptats $p(\cdot)$ i $q(\cdot)$. Es tracta d'una equació diferencial estocàstica que es resol "cap enrere", ja que, en lloc de presentar una condició inicial, el que es dóna és una condició final. No obstant, però, es demanarà que la solució $(p(\cdot), q(\cdot))$ sigui $\{\mathcal{A}_t\}_{t \geq 0}$-adaptada.

Observem que, suposant (E1) i (E2), per a cada control òptim $\overline{u}(\cdot)$ donat i cada trajectòria òptima corresponent $\overline{x}(\cdot)$, que siguin solució del sistema (4.41), el sistema (4.45) admet una única solució adaptada $(p(\cdot), q(\cdot))$ [veure Teorema 2.2 de la pàgina 349 de [6] per a més detalls].

A diferència del cas determinista, en el cas estocàstic cal introduir una nova variable (a part de la funció adjunta) per reflectir la incertesa del sistema. Aquesta nova variable $P(\cdot)$ s'anomena funció adjunta addicional i el sistema que la regeix és l'equació adjunta adicional:

$$\begin{cases} dP(t) = -\Big\{ b_x(t, \overline{x}(t), \overline{u}(t))^\top P(t) + P(t) b_x(t, \overline{x}(t), \overline{u}(t)) \\ \qquad + \sum_{j=1}^{m} \sigma_x^j(t, \overline{x}(t), \overline{u}(t))^\top P(t) \sigma_x^j(t, \overline{x}(t), \overline{u}(t)) \\ \qquad + \sum_{j=1}^{m} \Big\{ \sigma_x^j(t, \overline{x}(t), \overline{u}(t))^\top Q_j(t) + Q_j(t) \sigma_x^j(t, \overline{x}(t), \overline{u}(t)) \Big\} \\ \qquad + H_{xx}(t, \overline{x}(t), \overline{u}(t), p(t), q(t)) \Big\} dt + \sum_{j=1}^{m} Q_j(t) dW^j(t), \quad \forall t \in [0, T], \\ P(T) = -h_{xx}(\overline{x}(T)), \end{cases} \tag{4.46}$$

on $p(\cdot)$ i $q(\cdot)$ són solució del sistema anterior (4.45) i on la funció $H : [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n \times m} \to \mathbb{R}$ es defineix, en aquest cas estocàstic, com

$$H(t, x, u, p, q) = \langle p, b(t, x, u) \rangle + \operatorname{tr}\Big(q^\top \sigma(t, x, u) \Big) - f(t, x, u). \tag{4.47}$$

De nou, en el sistema (4.46) les incògnites són els dos processos $\{\mathcal{A}_t\}_{t \geq 0}$-adaptats $P(\cdot)$ i $Q(\cdot)$.

Quant a nomenclatura, el sistema (4.45) s'anomena equació adjunta de primer ordre i la variable $p(\cdot)$ és la funció adjunta de primer ordre. Paral·lelament, el sistema (4.46) és l'equació adjunta de segon ordre (o equació adjunta addicional) i la variable $P(\cdot)$ és la funció adjunta de segon ordre (o funció adjunta addicional). Per altra banda, la funció

H s'anomena Hamiltonià estocàstic.

Per seguir desenvolupant aquesta teoria, és necessari definir un Hamiltonià generalitzat. Abans, però, necessitem donar nom al conjunt de les matrius simètriques de dimensió n; és a dir, definirem el conjunt $\mathcal{S}^n$ com $\mathcal{S}^n = \{A \in \mathbb{R}^{n\times n} | A^\top = A\}$.

Ara ja sí, definim el Hamiltonià generalitzat $G : [0,T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathcal{S}^n \to \mathbb{R}$ com

$$G(t,x,u,p,P) \coloneqq \frac{1}{2}\mathrm{tr}\left(P\sigma(t,x,u)\sigma(t,x,u)^\top\right) + \langle p, b(t,x,u)\rangle - f(t,x,u). \tag{4.48}$$

Fent un abús de notació[10], podem comparar el Hamiltonià estocàstic $H(t,x,u,p,q)$ definit a (4.47) amb el Hamiltonià determinista $H(t,x,u,p)$ definit a (4.7). Observem que, clarament, el Hamiltonià estocàstic coincideix amb el determinista quan $\sigma(t,x,u) = 0$. Per altra banda, podem observar la definició (4.48) del Hamiltonià generalitzat $G(t,x,u,p,P)$ i comparar-la amb les definicions (4.47) i (4.7) que acabem de mencionar, de manera que

$$\begin{cases} G(t,x,u,p,P) = H(t,x,u,p) + \frac{1}{2}\mathrm{tr}\left(\sigma(t,x,u)^\top P\sigma(t,x,u)\right), \\ H(t,x,u,p,q) = H(t,x,u,p) + \mathrm{tr}\left(q^\top \sigma(t,x,u)\right). \end{cases} \tag{4.49}$$

Finalment, abans de postular el principi del màxim de Pontryagin en el cas estocàstic, cal definir una nova funció $\mathcal{H}$ de la forma $\mathcal{H} : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ tal que:

$$\begin{aligned} \mathcal{H}(t,x,u) &\coloneqq H(t,x,u,p(t),q(t)) - \frac{1}{2}\mathrm{tr}\left(\sigma(t,\overline{x}(t),\overline{u}(t))^\top P(t)\sigma(t,\overline{x}(t),\overline{u}(t))\right) \\ &\quad + \frac{1}{2}\mathrm{tr}\left([\sigma(t,x,u) - \sigma(t,\overline{x}(t),\overline{u}(t))]^\top P(t)[(\sigma(t,x,u) - \sigma(t,\overline{x}(t),\overline{u}(t))]\right) \\ &\coloneqq G(t,x,u,p(t),P(t)) + \mathrm{tr}\left(\sigma(t,x,u)^\top [q(t) - P(t)\sigma(t,\overline{x}(t),\overline{u}(t))]\right). \end{aligned} \tag{4.50}$$

Després d'introduir tots aquests nous conceptes ja podem postular el principi del màxim de Pontryagin en el cas estocàstic, teorema que dóna nom a aquesta secció i que proporcionarà condicions necessàries per trobar controls òptims i, conseqüentment, trajectòries òptimes.

Teorema 4.5. *Suposem que es satisfà (E1) i (E2). Sigui $\overline{u}(\cdot)$ un control òptim i sigui $\overline{x}(\cdot)$ la corresponent trajectòria òptima d'un problema fonamental de control òptim estocàstic. Aleshores existeixen processos $p(\cdot) \in \mathbb{R}^n$ i $q(\cdot) \in \mathbb{R}^{n\times m}$ que satisfan l'equació adjunta de primer ordre (4.45) i processos $P(\cdot) \in \mathcal{S}^n$ i $Q(\cdot) \in \mathcal{S}^{n\times m}$ que satisfan l'equació adjunta de segon ordre (4.46) tals que, $\forall u \in U$, es satisfà que*

$$\begin{aligned} &H(t,\overline{x}(t),\overline{u}(t),p(t),q(t)) - H(t,\overline{x}(t),u,p(t),q(t)) \\ &\geq \frac{1}{2}\mathrm{tr}\left([\sigma(t,\overline{x}(t),\overline{u}(t)) - \sigma(t,\overline{x}(t),u)]^\top P(t)[\sigma(t,\overline{x}(t),\overline{u}(t)) - \sigma(t,\overline{x}(t),u)]\right), \quad \forall t \in [0,T], \end{aligned} \tag{4.51}$$

[10] Aquest abús de notació consisteix en designar amb la mateixa variable el Hamiltonià estocàstic i el Hamiltonià determinista. No obstant, aquest paràgraf és l'únic lloc del treball on coexisteixen ambdues definicions juntes i, per aquest motiu, no hem considerat necessari haver d'introduir una nova variable. En qualsevol cas, podem distingir l'un de l'altre perquè el Hamiltonià estocàstic depèn de cinc variables (t,x,u,p,q) i el determinista només en depèn de quatre (t,x,u,p).

o, equivalentment, es satisfà que

$$\mathcal{H}(t, \overline{x}(t), \overline{u}(t)) = \max_{u \in U} \mathcal{H}(t, \overline{x}(t), u), \quad \forall t \in [0, T]. \tag{4.52}$$

<u>*Demostració.*</u> La demostració del teorema del màxim de Pontryagin pel cas estocàstic s'escapa del nivell d'aquest treball degut a la seva extensió, però, sobretot, degut a la seva complexitat. No obstant, es pot trobar una demostració completa del mateix a [6], a les pàgines 123-137. □

Així doncs, la condició de maximalitat en el cas estocàstic és l'equació (4.52). Hi ha dos casos especials que val la pena comentar.

El primer cas és quan el coeficient de difusió no depèn de la variable de control; és a dir,

$$\sigma(t, x, u) = \sigma(t, x), \quad \forall (t, x, u) \in [0, T] \times \mathbb{R}^n \times U. \tag{4.53}$$

Aleshores, en aquest cas, la condició de maximalitat (4.52) es pot reescriure com

$$H(t, \overline{x}(t), \overline{u}(t), p(t), q(t)) = \max_{u \in U} H(t, \overline{x}(t), u, p(t), q(t)), \quad \forall t \in [0, T], \tag{4.54}$$

condició que és exactament equivalent a la condició de maximalitat (4.6) del cas determinista. En aquest cas, doncs, no es necessita l'equació (4.46) per $P(\cdot)$ ni $Q(\cdot)$ i, per tant, no cal que les funcions b, σ, f i h siguin de classe $\mathcal{C}^2$ respecte x, sinó que poden ser, simplement, de classe $\mathcal{C}^1$.

El segon cas és quan la regió de control U és convexa i totes les funcions són de classe $\mathcal{C}^1$ respecte u. Aleshores, en aquest cas, la condició (4.51) implica

$$\langle H_u(t, \overline{x}(t), \overline{u}(t), p(t), q(t)), u - \overline{u}(t) \rangle \leq 0, \quad \forall (t, u) \in [0, T] \times U. \tag{4.55}$$

Aquesta equació (4.55) és la forma local del principi de maximalitat (en contraposició a la forma global donada per les equacions (4.51) o (4.52)). Observem que aquesta forma local no involucra la funció adjunta de segon ordre $P(\cdot)$.

Un cop analitzats aquests dos casos particulars, prosseguim amb el desenvolupament del cas estocàstic en aquest primer mètode de resolució d'un problema de control òptim. De manera completament anàloga al cas determinista, fent ús de la definició (4.47) del Hamiltonià estocàstic H, podem reescriure l'equació d'estat (4.41) amb la corresponent equació adjunta de primer ordre (4.45) com

$$\begin{cases} dx(t) = H_p(t, x(t), u(t), p(t), q(t))dt + H_q(t, x(t), u(t), p(t), q(t))dW(t), \quad \forall t \in [0, T], \\ dp(t) = -H_x(t, x(t), u(t), p(t), q(t))dt + q(t)dW(t), \quad \forall t \in [0, T], \\ x(0) = x_0, \quad p(T) = -h_x(x(T)). \end{cases} \tag{4.56}$$

Comparant-ho amb el cas anterior, on ja havíem definit el sistema Hamiltonià estès en (4.8), ara anomenarem sistema Hamiltonià estocàstic estès a la combinació de les equacions (4.56),(4.46) i (4.51) (o (4.52)).

Seguint el mateix plantejament que en el cas determinista, fins ara el que hem fet ha sigut donar condicions necessàries per un control òptim. Nogensmenys, per trobar unes condicions de suficiència cal que afegim un altre supòsit. Així doncs, suposem que

(E3) La regió de control U és convexa amb l'interior no buit, i les aplicacions b, σ i f són localment Lipschitz respecte u, i les parcials respecte x són contínues en (x, u).

Amb aquest supòsit en joc, podem postular el següent teorema, que proporciona condicions suficients d'optimalitat.

Teorema 4.6. *Suposem que es satisfà (E1), (E2) i (E3). Sigui $\overline{u}(\cdot)$ un control qualsevol i sigui $\overline{x}(\cdot)$ la corresponent trajectòria. Siguin $p(\cdot)$ i $P(\cdot)$ les corresponents variables adjuntes de primer i segon ordre, respectivament, i siguin $q(\cdot)$ i $Q(\cdot)$ els processos $\{\mathcal{A}_t\}_{t\geq 0}$-adaptats corresponents. Suposem que $H(t, \cdot, \cdot, p(t), q(t))$ és còncava $\forall t \in [0, T]$ i que $h(\cdot)$ és convexa. Aleshores $\overline{u}(\cdot)$ i $\overline{x}(\cdot)$ són òptims si*

$$\mathcal{H}(t, \overline{x}(t), \overline{u}(t)) = \max_{u \in U} \mathcal{H}(t, \overline{x}(t), u), \quad \forall t \in [0, T]. \tag{4.57}$$

<u>*Demostració.*</u> Per la concavitat de $H(t, \cdot, \cdot, p(t), q(t))$, tenim que, $\forall u(\cdot) \in \mathcal{U}[0, T]$,

$$\begin{aligned} &\int_0^T \Big\{ H(t, x(t), u(t), p(t), q(t)) - H(t, \overline{x}(t), \overline{u}(t), p(t), q(t)) \Big\} dt \\ &\leq \int_0^T \langle H_x(t, \overline{x}(t), \overline{u}(t), p(t), q(t)), x(t) - \overline{x}(t) \rangle dt, \end{aligned} \tag{4.58}$$

on $x(\cdot)$ és la trajectòria corresponent al control $u(\cdot)$.

Sigui $z(t) = x(t) - \overline{x}(t)$ tal que

$$\begin{cases} dz(t) = \Big\{ b_x(t, \overline{x}(t), \overline{u}(t)) z(t) + \alpha(t) \Big\} dt \\ \qquad\qquad + \displaystyle\sum_{j=1}^m \Big\{ \sigma_x^j(t, \overline{x}(t), \overline{u}(t)) z(t) + \beta^j(t) \Big\} dW^j(t), \quad \forall t \in [0, T], \\ z(0) = 0, \end{cases} \tag{4.59}$$

on

$$\begin{cases} \alpha(t) = -b_x(t, \overline{x}(t), \overline{u}(t)) z(t) + b(t, x(t), u(t)) - b(t, \overline{x}(t), \overline{u}(t)), \\ \beta^j(t) = -\sigma_x^j(t, \overline{x}(t), \overline{u}(t)) z(t) + \sigma^j(t, x(t), u(t)) - \sigma^j(t, \overline{x}(t), \overline{u}(t)), \quad \forall\, 1 \leq j \leq m. \end{cases} \tag{4.60}$$

Per la relació de dualitat entre el sistema variacional (4.59) i el sistema adjunt (4.45), tenim que

$$\begin{aligned} &E\Big\{ \langle h_x(\overline{x}(T)), z(T) \rangle \Big\} = -E\Big\{ \langle p(T), z(T) \rangle \Big\} + E\Big\{ \langle p(0), z(0) \rangle \Big\} \\ &= -E\Big\{ \int_0^T \{ \langle f_x(t, \overline{x}(t), \overline{u}(t)), z(t) \rangle + \langle p(t), \alpha(t) \rangle + \sum_{j=1}^m \langle q_j(t), \beta^j(t) \rangle \} dt \Big\} \\ &= E\Big\{ \int_0^T \langle H_x(t, \overline{x}(t), \overline{u}(t), p(t), q(t)), z(t) \rangle dt \Big\} - E\Big\{ \int_0^T \{ \langle p(t), b(t, x(t), u(t)) - b(t, \overline{x}(t), \overline{u}(t)) \rangle \\ &+ \sum_{j=1}^m \langle q_j(t), \sigma^j(t, \overline{x}(t), u(t)) - \sigma^j(t, \overline{x}(t), \overline{u}(t)) \rangle \} dt \Big\}. \end{aligned} \tag{4.61}$$

Unint aquesta darrera equació (4.61) amb l'equació (4.58) de la concavitat de $H(t,\cdot,\cdot,p(t),q(t))$, tenim que

$$\begin{aligned}
&E\Big\{\langle h_x(\overline{x}(T)),z(T)\rangle\Big\} = E\Big\{\int_0^T \langle H_x(t,\overline{x}(t),\overline{u}(t),p(t),q(t)),z(t)\rangle dt\Big\}\\
&-E\Big\{\int_0^T \big\{\langle p(t),b(t,x(t),u(t))-b(t,\overline{x}(t),\overline{u}(t))\rangle+\sum_{j=1}^m\langle q_j(t),\sigma^j(t,\overline{x}(t),u(t))-\sigma^j(t,\overline{x}(t),\overline{u}(t))\rangle\big\}dt\Big\}\\
&\geq E\Big\{\int_0^T \big\{H(t,x(t),u(t),p(t),q(t))-H(t,\overline{x}(t),\overline{u}(t),p(t),q(t))\big\}dt\Big\}\\
&-E\Big\{\int_0^T \big\{\langle p(t),b(t,x(t),u(t))-b(t,\overline{x}(t),\overline{u}(t))\rangle+\sum_{j=1}^m\langle q_j(t),\sigma^j(t,\overline{x}(t),u(t))-\sigma^j(t,\overline{x}(t),\overline{u}(t))\rangle\big\}dt\Big\}\\
&=-E\Big\{\int_0^T \{f(t,x(t),u(t))-f(t,\overline{x}(t),\overline{u}(t))\}dt\Big\}.
\end{aligned}\tag{4.62}$$

Per altra banda, la convexitat de la funció $h(\cdot)$, ens diu que

$$E\Big\{h(x(T))-h(\overline{x}(T))\Big\}\geq E\Big\{\langle h_x(\overline{x}(T)),x(T)-\overline{x}(T)\rangle\Big\}=E\Big\{\langle h_x(\overline{x}(T)),z(T)\rangle\Big\}.\tag{4.63}$$

Així doncs, ja per acabar, combinant l'equació (4.62) amb l'equació (4.63), obtenim que

$$\begin{aligned}
-E\Big\{\int_0^T \{f(t,x(t),u(t))-f(t,\overline{x}(t),\overline{u}(t))\}dt\Big\} &\leq E\Big\{\langle h_x(\overline{x}(T)),z(T)\rangle\Big\}\\
&\leq E\Big\{h(x(T))-h(\overline{x}(T))\Big\}.
\end{aligned}\tag{4.64}$$

Reagrupant els termes dels extrems i recordant que l'esperança és un operador lineal, obtenim que

$$\begin{aligned}
J(\overline{u}(\cdot)) &= E\Big\{\int_0^T f(t,\overline{x}(t),\overline{u}(t))dt+h(\overline{x}(T))\Big\}\\
&\leq J(u(\cdot)) = E\Big\{\int_0^T f(t,x(t),u(t))dt+h(x(T))\Big\};
\end{aligned}\tag{4.65}$$

és a dir, hem obtingut que, $\forall u(\cdot)\in U$,

$$J(\overline{u}(\cdot))\leq J(u(\cdot)).\tag{4.66}$$

Per tant, concloem que $\overline{u}(\cdot)$ és òptim, tal com volíem demostrar. □

5 Principi de programació dinàmica Bellman

Tot seguit, presentarem el segon mètode de resolució d'un problema de control òptim, el principi de la programació dinàmica de Bellman.

5.1 Cas determinista

Seguint el formalisme de la secció **3.1 Cas determinista**, considerem el següent sistema de control

$$\begin{cases} \dot{x}(t) = b(t, x(t), u(t)), \quad \forall t \in [0, T] \\ x(0) = x_0, \end{cases} \tag{5.1}$$

amb el respectiu funcional de cost

$$J(u(\cdot)) = \int_0^T f(t, x(t), u(t))dt + h(x(T)). \tag{5.2}$$

Observem que, tant en aquest cas com en tota la secció **4.1 Cas determinista**, el temps inicial ($t = 0$) i l'estat inicial ($x(0) = x_0$) han sigut valors fixats a priori. Ara, en canvi, la idea bàsica del principi de la programació dinàmica de Bellman és considerar una família de problemes de control òptim amb diferents condicions inicials, amb l'objectiu d'establir relacions entre elles i, en acabat, resoldre-les totes.

Així doncs, sigui $(s, y) \in [0, T) \times \mathbb{R}^n$ una condició inicial (de temps inicial i d'estat inicial) qualsevol. Ara considerem un sistema de control anàleg al (5.1):

$$\begin{cases} \dot{x}(t) = b(t, x(t), u(t)), \quad \forall t \in [s, T] \\ x(s) = y, \end{cases} \tag{5.3}$$

amb el funcional de cost associat

$$J(s, y; u(\cdot)) = \int_s^T f(t, x(t), u(t))dt + h(x(T)). \tag{5.4}$$

El nostre problema fonamental de control òptim determinista és el que havíem definit a l'equació (3.3). No obstant, ara en podem definir un altre de més general, que inclogui el primer.

Sigui ara $\mathcal{U}[s, T] := \{u : [s, T] \to U | u(\cdot) \text{ és mesurable}\}$. Per valors donats $(s, y) \in [0, T) \times \mathbb{R}^n$, el problema general de control òptim en el cas determinista consisteix en trobar, sota els lligams de l'equació (5.3), un control òptim $\overline{u}(\cdot) \in \mathcal{U}[s, T]$ tal que

$$J(s, y; \overline{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[s,T]} J(s, y; u(\cdot)). \tag{5.5}$$

Observem que aquesta formulació general del problema es tracta, realment, d'una família de problemes de control òptim parametritzada per $(s, y) \in [0, T) \times \mathbb{R}^n$, de manera que el problema fonamental descrit per l'equació (3.3) n'és un cas concret quan $s = 0$ i $y = x_0$. Així doncs, variant les condicions inicials (s, y) en $[0, T) \times \mathbb{R}^n$, obtenim informació "dinàmica"de la família de problemes de control òptim descrita per l'equació (5.5).

Fent-ho així, ara el nostre objectiu ha canviat una mica i consisteix en resoldre el problema general de control òptim determinista, que ve donat per l'equació (5.5); en contraposició al problema fonamental de control òptim determinista, que ve donat per l'equació (3.3).

Igual que en capítols anteriors, abans d'atacar el problema, farem unes suposicions que facilitaran la feina. Per tant, dins d'aquest segon mètode de resolució d'un problema de control òptim, suposem que:

(D1)$'$ Les aplicacions $b : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $f : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ i $h : \mathbb{R}^n \to \mathbb{R}$ són uniformement contínues, i existeix una constant $L > 0$ tal que per una funció $\phi(t,x,u) \in \{b(t,x,u), f(t,x,u), h(x)\}$ es satisfà que:

$$\begin{cases} |\phi(t,x,u) - \phi(t,\hat{x},u)| \leq L|x-\hat{x}|, \quad \forall (t,x,\hat{x},u) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times U, \\ |\phi(t,0,u)| \leq L, \quad \forall (t,u) \in [0,T] \times U. \end{cases} \tag{5.6}$$

Sota aquest supòsit, ja podem abordar el problema descrit per l'equació (5.5); és a dir, el problema general de control òptim determinista. En primer lloc, observem que, sota aquesta suposició, per cada $(s,y) \in [0,T) \times \mathbb{R}^n$ i per cada control $u(\cdot) \in \mathcal{U}[s,T]$, existeix una única solució $x(\cdot) := x(\cdot; s, y, u(\cdot))$ del sistema anterior (5.3), gràcies al teorema d'existència i unicitat de solucions d'una equació diferencial [veure pàgines 52-54 de [6] per a més detalls]. A més a més, gràcies a aquesta unicitat de la solució $x(\cdot)$, el funcional de cost (5.4) està ben definit. En segon lloc, definim la següent funció:

$$\begin{cases} V(s,y) = \inf_{u(\cdot) \in \mathcal{U}[s,T]} J(s,y;u(\cdot)), \quad \forall (s,y) \in [0,T) \times \mathbb{R}^n, \\ V(T,y) = h(y), \quad \forall y \in \mathbb{R}^n. \end{cases} \tag{5.7}$$

Aquesta funció $V(\cdot,\cdot)$ s'anomena funció valor del problema fonamental de control òptim determinista.

Tot seguit, amb la idea de resoldre el problema descrit per l'equació (5.5), podem postular el principi d'optimalitat de Bellman o principi de la programació dinàmica de Bellman, teorema que dóna nom a aquesta secció i que ens ajudarà a trobar controls òptims i, conseqüentment, trajectòries òptimes.

Teorema 5.1. *Suposem que es satisfà (D1)$'$. Aleshores, per qualsevol $(s,y) \in [0,T) \times \mathbb{R}^n$,*

$$V(s,y) = \inf_{u(\cdot) \in \mathcal{U}[s,T]} \left\{ \int_s^{\hat{s}} f(t, x(t;s,y,u(\cdot)), u(t))dt + V(\hat{s}, x(\hat{s}; s, y, u(\cdot))) \right\}, \quad \forall\, 0 \leq s \leq \hat{s} \leq T. \tag{5.8}$$

<u>*Demostració.*</u> Sigui $\overline{V}(s,y)$ la part dreta de la igualtat de l'equació (5.8); és a dir,

$$\overline{V}(s,y) = \inf_{u(\cdot) \in \mathcal{U}[s,T]} \left\{ \int_s^{\hat{s}} f(t, x(t;s,y,u(\cdot)), u(t))dt + V(\hat{s}, x(\hat{s}; s, y, u(\cdot))) \right\}. \tag{5.9}$$

Per la definició (5.7) de la funció $V(\cdot,\cdot)$, podem dir que

$$\begin{aligned} V(s,y) &= \inf_{u(\cdot) \in \mathcal{U}[s,T]} J(s,y;u(\cdot)) \leq J(s,y;u(\cdot)) = \int_s^T f(t,x(t),u(t))dt + h(x(T)) \\ &= \int_s^{\hat{s}} f(t,x(t),u(t))dt + \int_{\hat{s}}^T f(t,x(t),u(t))dt + h(x(T)) \\ &= \int_s^{\hat{s}} f(t,x(t),u(t))dt + J(\hat{s}, x(\hat{s}); u(\cdot)), \quad \forall u(\cdot) \in \mathcal{U}[s,T]. \end{aligned} \tag{5.10}$$

Per tant, tenim una desigualtat que és vàlida $\forall u(\cdot) \in \mathcal{U}[s,T]$. Prenent l'ínfim sobre $u(\cdot) \in \mathcal{U}[s,T]$ a l'última igualtat de l'equació (5.10) i recordant la definició (5.9) de $\overline{V}(s,y)$, obtenim que

$$V(s,y) \leq \overline{V}(s,y). \tag{5.11}$$

En canvi, si prenem un $\epsilon > 0$ qualsevol, podem dir que existeix un control $u_\epsilon(\cdot) \in \mathcal{U}[s,T]$, amb una corresponent trajectòria $x_\epsilon(\cdot) = x(\cdot; s, y, u_\epsilon(\cdot))$, de manera que

$$V(s,y) + \epsilon \geq J(s,y;u_\epsilon(\cdot)). \tag{5.12}$$

A més a més, emprant la definició de $J(s,y;u(\cdot))$ de l'equació (5.4) i la definició de $V(s,y)$ de l'equació (5.7), podem dir que

$$J(s,y;u_\epsilon(\cdot)) = J(\hat{s}, x_\epsilon(\hat{s}); u_\epsilon(\cdot)) + \int_s^{\hat{s}} f(t, x_\epsilon(t), u_\epsilon(t))dt \geq \int_s^{\hat{s}} f(t, x_\epsilon(t), u_\epsilon(t))dt + V(\hat{s}, x_\epsilon(\hat{s})). \tag{5.13}$$

Finalment, usant la definició de $\overline{V}(s,y)$ de l'equació (5.9), podem dir que

$$\int_s^{\hat{s}} f(t, x_\epsilon(t), u_\epsilon(t))dt + V(\hat{s}, x_\epsilon(\hat{s})) \geq \overline{V}(s,y). \tag{5.14}$$

Així doncs, unint les equacions (5.12), (5.13) i (5.14), obtenim que, per un valor $\epsilon > 0$ qualsevol, es satisfà la següent relació:

$$V(s,y) + \epsilon \geq \overline{V}(s,y). \tag{5.15}$$

Fent ϵ tan petit com es vulgui, s'obté que

$$V(s,y) \geq \overline{V}(s,y). \tag{5.16}$$

Per tant, observant les equacions (5.11) i (5.16), s'obté que $V(s,y) = \overline{V}(s,y)$; és a dir, s'obté, tal com volíem demostrar, la igualtat (5.8). $\square$

A continuació, farem una observació molt interessant sobre aquest teorema. En aquesta observació, usarem, principalment, les definicions (5.4) i (5.7) i el concepte de control i trajectòria òptims.

Observació 13. Suposem que, donat un problema de control òptim determinista, tenim un control òptim $\overline{u}(\cdot)$ i la corresponent trajectòria òptima $\overline{x}(\cdot)$. Sigui $\hat{s} \in (s,T)$ un valor qualsevol. Aleshores,

$$\begin{aligned} V(s,y) &= \inf_{u(\cdot)\in\mathcal{U}[s,T]} J(s,y;u(\cdot)) = J(s,y;\overline{u}(\cdot)) = \int_s^{\hat{s}} f(t,\overline{x}(t),\overline{u}(t))dt + J(\hat{s},\overline{x}(\hat{s});\overline{u}(\cdot)) \\ &\geq \int_s^{\hat{s}} f(t,\overline{x}(t),\overline{u}(t))dt + V(\hat{s},\overline{x}(\hat{s})) \geq V(s,y), \end{aligned} \tag{5.17}$$

on la darrera desigualtat es deu a l'equació (5.8). Per tant, tota l'equació (5.17) ha de consistir en igualtats encadenades i, en particular, sobté que

$$V(\hat{s},\overline{x}(\hat{s})) = J(\hat{s},\overline{x}(\hat{s});\overline{u}(\cdot)) = \int_{\hat{s}}^T f(t,\overline{x}(t),\overline{u}(t))dt + h(\overline{x}(T)), \;\; \forall \hat{s} \in (s,T). \tag{5.18}$$

Això significa que si hi ha optimalitat global, aleshores hi ha optimalitat local, ja que hem vist que si tenim un control $\overline{u}(\cdot)$ òptim en $[s,T]$ (amb condició inicial (s,y)), aleshores tenim que $\overline{u}|_{[\hat{s},T]}(\cdot)$ és un control òptim en $[\hat{s},T]$ (amb condició inicial $(\hat{s},\overline{x}(\hat{s}))$).

El Teorema 5.1 és el principi de la programació dinàmica de Bellman, i l'equació (5.8) d'aquest teorema es coneix amb el nom d'equació de la programació dinàmica, donant nom al títol d'aquesta secció i, per tant, a la segona manera clàssica de resoldre un problema de control òptim. Aquest teorema proporciona, per mitjà de la funció valor, una relació entre la família de problemes de control òptim donats per l'equació (5.5). No obstant això, es tracta d'una equació molt complicada, perquè les operacions que hi ha a la part dreta de la igualtat ($\overline{V}(s,y)$) són difícils. Per aquesta raó, el nostre objectiu d'ara és trobar una equació per $V(s,y)$ que sigui més senzilla.

Per dur a terme aquest nou objectiu, sigui $\mathcal{C}^1([0,T]\times\mathbb{R}^n) := \{v : [0,T]\times\mathbb{R}^n \to \mathbb{R} | v$ és una funció contínua i amb derivada contínua$\}$.

Proposició 5.2. *Suposem que es satisfà i (D1)$'$. Suposem que la funció valor V (definida a l'equació (5.7)) satisfà que $V\in\mathcal{C}^1([0,T]\times\mathbb{R}^n)$. Aleshores V és una solució del següent problema de valors finals d'una equació en derivades parcials:*

$$\begin{cases} -v_t + \sup\limits_{u\in U} H(t,x,u,-v_x) = 0, & (t,x)\in[0,T)\times\mathbb{R}^n, \\ v|_{t=T} = h(x), & x\in\mathbb{R}^n, \end{cases} \tag{5.19}$$

on la funció $H : [0,T]\times\mathbb{R}^n\times U\times\mathbb{R}^n \to \mathbb{R}$ es defineix com:

$$H(t,x,u,p) := \langle p, b(t,x,u)\rangle - f(t,x,u). \tag{5.20}$$

L'equació (5.19) és la famosa equació Hamilton-Jacobi-Bellman (HJB) associada al problema fonamental de control òptim determinista. La funció H definida a l'equació (5.20) s'anomena Hamiltonià i és la mateixa que la definida en el capítol anterior en (4.7).

Demostració. En primer lloc, fixem-nos que la segona relació de l'equació (5.19) és immediata si utilitzem la segona relació de la definició (5.7). Així doncs, només hem de demostrar la primera relació de l'equació (5.19). Fixem $u\in U$ i sigui $x(\cdot)$ la trajectòria d'estat corresponent al control $u(t)=u$.

A continuació, farem tendir $\hat{s}$ a s (sempre conservant que $\hat{s}-s>0$, i ho representarem per $\hat{s}\downarrow s$). Gràcies a (5.8), es pot dir que, $\forall u\in U$,

$$V(s,y) = \inf_{u(\cdot)\in\mathcal{U}[s,T]} \left\{ \int_s^{\hat{s}} f(t,x(t),u)dt + V(\hat{s},x(\hat{s})) \right\} \le \int_s^{\hat{s}} f(t,x(t),u)dt + V(\hat{s},x(\hat{s})). \tag{5.21}$$

Observant la última inequació i recordant que $\hat{s}-s>0$, podem escriure que, $\forall u\in U$,

$$0 \ge -\frac{V(\hat{s},x(\hat{s}))-V(s,y)}{\hat{s}-s} - \frac{1}{\hat{s}-s}\int_s^{\hat{s}} f(t,x(t),u)dt. \tag{5.22}$$

Observem que quan $\hat{s}\downarrow s$, aleshores $x(\hat{s}) \longrightarrow x(s)=y$; és a dir, que quan $\hat{s}\downarrow s$, la primera fracció de la part dreta de la inequació (5.22) representa la derivada total temporal de la funció $V(s,y)$. Per tant, $\forall u\in U$ i quan $\hat{s}\downarrow s$, podem dir que

$$-\frac{V(\hat{s},x(\hat{s}))-V(s,y)}{\hat{s}-s} - \frac{1}{\hat{s}-s}\int_s^{\hat{s}} f(t,x(t),u)dt \longrightarrow -\frac{d}{dt}V(s,y) - f(t,x(t),u). \tag{5.23}$$

Usant la definició de derivada total, $\forall u \in U$ i quan $\hat{s} \downarrow s$, tenim que

$$0 \geq -\frac{V(\hat{s}, x(\hat{s})) - V(s, y)}{\hat{s} - s} - \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f(t, x(t), u) dt \longrightarrow -V_t(s, y) - \langle V_x(s, y), b(s, y, u) \rangle - f(s, y, u). \tag{5.24}$$

Fixem-nos que si emprem la definició (5.20) del Hamiltonià, aleshores, $\forall u \in U$ i quan $\hat{s} \downarrow s$, es satisfà que

$$0 \geq -\frac{V(\hat{s}, x(\hat{s})) - V(s, y)}{\hat{s} - s} - \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f(t, x(t), u) dt \longrightarrow -V_t(s, y) + H(s, y, u, -V_x(s, y)). \tag{5.25}$$

Per acabar, com el resultat anterior és vàlid $\forall u \in U$, en particular deduim que

$$0 \geq -V_t(s, y) + \sup_{u \in U} H(s, y, u, -V_x(s, y)). \tag{5.26}$$

Per altra banda, de manera similar al que vam fer a l'equació (5.12), si prenem un $\epsilon > 0$ qualsevol i un $\hat{s}$ qualsevol tal que $0 \leq s < \hat{s} \leq T$ amb $\hat{s} - s > 0$ suficientment petit, podem dir que existeix un control $u(\cdot) := u_{\epsilon, \hat{s}}(\cdot) \in \mathcal{U}[s, T]$ tal que

$$V(s, y) + \epsilon(\hat{s} - s) \geq \int_s^{\hat{s}} f(t, x(t), u(t)) dt + V(\hat{s}, x(\hat{s})). \tag{5.27}$$

Reordenant la inequació, podem escriure que

$$-\epsilon \leq -\frac{V(\hat{s}, x(\hat{s})) - V(s, y)}{\hat{s} - s} - \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f(t, x(t), u) dt. \tag{5.28}$$

Com ϵ és arbitrari, el podem fer tan petit com desitgem, i al final obtindrem que

$$0 \leq -\frac{V(\hat{s}, x(\hat{s})) - V(s, y)}{\hat{s} - s} - \frac{1}{\hat{s} - s} \int_s^{\hat{s}} f(t, x(t), u) dt; \tag{5.29}$$

és a dir, obtindrem una equació molt similar a (5.22), però amb la desigualtat girada. Així doncs, seguint els mateixos passos que els que hem fet des de l'equació (5.22) fins l'equació (5.25), ara arribarem a veure que

$$0 \leq -V_t(s, y) + H(s, y, u, -V_x(s, y)) \leq -V_t(s, y) + \sup_{u \in U} H(s, y, u, -V_x(s, y)). \tag{5.30}$$

Finalment, combinant les equacions (5.26) i (5.30), obtenim, tal com volíem demostrar, que

$$0 = -V_t(s, y) + \sup_{u \in U} H(s, y, u, -V_x(s, y)). \tag{5.31}$$

□

Tot seguit, explicarem quina és la relació entre les solucions de l'equació HJB i l'obtenció d'un control òptim (que segueix sent el nostre objectiu). Sigui $V \in \mathcal{C}^1([0, T] \times \mathbb{R}^n)$

la funció valor obtinguda com a solució del sistema (5.19) després de resoldre'l. Addicionalment, suposem que el suprem de l'equació (5.19) s'assoleix per $u = \overline{\mathbf{u}}(t,x)$; és a dir, que

$$H(t,x,\overline{\mathbf{u}}(t,x),-V_x(t,x)) = \sup_{u\in U} H(t,x,u,-V_x(t,x)), \quad \forall (t,x)\in[0,T]\times\mathbb{R}^n. \tag{5.32}$$

També, suposem que, $\forall (s,y)\in[0,T)\times\mathbb{R}^n$, existeix una solució $\overline{x}(\cdot)$ del següent sistema:

$$\begin{cases} \dot{\overline{x}}(t) = b(t,\overline{x}(t),\overline{\mathbf{u}}(t,\overline{x}(t))), \quad \forall t\in[s,T] \\ \overline{x}(s) = y. \end{cases} \tag{5.33}$$

A continuació, comprovarem que aquesta solució $\overline{x}(\cdot)$ es tracta, realment, d'una trajectòria òptima, i que això ens permet concloure el nostre problema. Per fer-ho, comencem definint el control $\overline{u}(t)$ com $\overline{u}(t) = \overline{\mathbf{u}}(t,\overline{x}(t))$. Tot seguit, apliquem la regla de la cadena fixant-nos en el sistema anterior i observant l'equació (5.18), de manera que

$$\frac{d}{dt}V(t,\overline{x}(t)) = V_t(t,\overline{x}(t)) + \langle V_x(t,\overline{x}(t)), b(t,\overline{x}(t),\overline{u}(t))\rangle = -f(t,\overline{x}(t),\overline{u}(t)), \tag{5.34}$$

expressió que si l'integrem des de s fins a T, obtenim que

$$V(T,\overline{x}(T)) - V(s,\overline{x}(s)) = -\int_s^T f(t,\overline{x}(t),\overline{u}(t))dt. \tag{5.35}$$

Finalment, reagrupant i recordant les segones relacions de les equacions (5.7) i (5.33), obtenim que

$$V(s,y) = h(\overline{x}(T)) + \int_s^T f(t,\overline{x}(t),\overline{u}(t))dt = J(s,y;\overline{u}(\cdot)); \tag{5.36}$$

és a dir, obtenim que el control $\overline{u}(t)$ és, en realitat, un control òptim, i queda comprovat, efectivament, que la corresponent trajectòria $\overline{x}(t)$ és una trajectòria òptima. Per tant, hem resolt el problema general de control òptim determinista i, en particular, hem assolit el nostre objectiu inicial, ja que també hem resolt el problema fonamental de control òptim determinista (prenent com a condicions inicials $(s,y) = (0,x_0)$ en el sistema (5.33)).

Així doncs, per resumir, els passos que s'han seguit en aquesta segona tècnica de resolució del problema de control òptim en el cas determinista són:

1. Introduir el problema general de control òptim determinista, a partir del problema fonamental de control òptim determinista.
2. Definir la funció valor $V(s,y)$.
3. Definir el principi d'optimalitat de Bellman o principi de la programació dinàmica de Bellman (Teorema 5.1).
4. Derivar les equacions HJB (Proposició 5.2).
5. Resoldre l'equació HJB (equació (5.19)) per trobar la funció valor $V(t,x)$.
6. Trobar $\overline{\mathbf{u}}(t,x)$ a partir de l'equació (5.32).

7. Resoldre el sistema (5.33) (amb condicions inicials $(s, y) = (0, x_0)$) per obtenir, primer, la trajectòria òptima $\overline{x}(\cdot)$ i, en segon lloc, el control òptim $\overline{u}(\cdot)$ a partir de $\overline{u}(\cdot) = \overline{\mathbf{u}}(\cdot, \overline{x}(\cdot))$.

I així és com es resol el problema fonamental de control optim en el cas determinista segons el principi de programació dinàmica de Bellman.

5.2 Cas estocàstic

Seguint el formalisme de la secció **3.2 Cas estocàstic**, considerem el següent sistema de control

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), & \forall t \in [0, T], \\ x(0) = x_0, \end{cases} \tag{5.37}$$

amb el respectiu funcional de cost

$$J(u(\cdot)) = E\left\{ \int_0^T f(t, x(t), u(t))dt + h(x(T)) \right\}. \tag{5.38}$$

Tal com ho hem explicat en **3.2 Cas estocàstic**, en el nostre marc de treball teníem fixats un espai de probabilitat filtrat $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, P)$ i un moviment Brownià m-dimensional $W(\cdot)$. Ara, en canvi, per l'estudi del principi de la programació dinàmica de Bellman en el cas estocàstic, cal considerar una formulació diferent: és necessari variar tant l'espai de probabilitat filtrat com el moviment brownià, i considerar-los parts del control.

Emfatitzem que aquest nova formulació és una construcció auxiliar matemàtica molt efectiva per solucionar problemes formulats amb el primer marc de treball. La principal raó per la qual aquest canvi de formulació ens serveix és perquè el nostre objectiu és minimitzar el valor esperat d'una certa variable aleatòria, que només depèn de la distribució del procés en qüestió. Per tant, si les solucions del sistema (5.37) en diferents espais de probabilitat tenen la mateixa distribució de probabilitat, aleshores hom té més llibertat per escollir quin és l'espai de probabilitat més adequat amb què treballar-hi. No obstant, aquesta formulació falla si les funcions b, σ, f o h són també aleatòries, ja que, en tal cas, es necessita un espai de probabilitat definit a priori (recordem que a l'Observació 11 ja vam descartar l'opció que aquestes funcions fossin aleatòries).

Més enllà d'aquest canvi en el marc de treball, per resoldre el principi de la programació dinàmica de Bellman en el cas estocàstic cal començar com ho hem fet en el cas determinista; és a dir, considerant una família de problemes de control òptim.

Prèviament, començarem recordant que, en la primera formulació, el nostre problema consistia en trobar un control òptim $\overline{u}(\cdot) \in \mathcal{U}[0, T]$ tal que

$$J(\overline{u}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}[0,T]} J(u(\cdot)), \tag{5.39}$$

on $\mathcal{U}[0, T] \coloneqq \{u : [0, T] \times \Omega \to U | u(\cdot) \text{ és mesurable i és } \{\mathcal{A}_t\}_{t\geq 0}\text{-adaptat}\}$.

Ara ja sí, igual que hem fet en el cas determinista, sigui $(s, y) \in [0, T) \times \mathbb{R}^n$ una condició inicial qualsevol. Considerem la família de problemes de control òptim parametritzada per $(s, y) \in [0, T) \times \mathbb{R}^n$; és a dir,

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), & \forall t \in [s, T], \\ x(s) = y, \end{cases} \tag{5.40}$$

amb el funcional de cost associat

$$J(s, y; u(\cdot)) = E\left\{ \int_s^T f(t, x(t), u(t))dt + h(x(T)) \right\}. \tag{5.41}$$

Per altra banda, quant a la nova formulació, cal conèixer en quin entorn matemàtic treballarem. Per tant, fixant $s \in [0,T)$, sigui $\mathcal{U}'[s,T]$ el conjunt dels quinterns $(\Omega, \mathcal{A}, P, W(\cdot), u(\cdot))$ que satisfan que:

(i) La terna $(\Omega, \mathcal{A}, P)$ és un espai de probabilitat.

(ii) $\{W(t)\}_{t\geq s}$ és un moviment Brownià estàndard de dimensió m definit en $(\Omega, \mathcal{A}, P)$ sobre l'interval $[s,T]$ (amb $P(W(s)=0)=1$ [recordem Definició 2.7]). A més, sigui $\mathcal{A}_t^s = \sigma\{W(r) : s \leq r \leq t\}$.

(iii) $u : [s,T] \times \Omega \to U$ és un procés $\{\mathcal{A}_t^s\}_{t\geq s}$-adaptat en $(\Omega, \mathcal{A}, P)$.

(iv) Per cada control $u(\cdot) \in U$ i per cada $y \in \mathbb{R}^n$, l'equació (5.40) admet una única solució $x(\cdot)$ sobre $(\Omega, \mathcal{A}, \{\mathcal{A}_t^s\}_{t\geq s}, P)$.

(v) Les funcions $f(\cdot, x(\cdot), u(\cdot))$ i $h(x(T))$ tenen esperança finita. Aquí, l'esperança es defineix sobre l'espai de probabilitat filtrat $(\Omega, \mathcal{A}, \{\mathcal{A}_t^s\}_{t\geq s}, P)$ (associat amb el quintern donat).

Observació 14. Sempre que l'espai de probabilitat $(\Omega, \mathcal{A}, P)$ i el moviment brownià $W(\cdot)$ estiguin clars i no presentin confusions, simplement escriurem $u(\cdot) \in \mathcal{U}'[s,T]$ en lloc de $(\Omega, \mathcal{A}, P, W(\cdot), u(\cdot)) \in \mathcal{U}'[s,T]$.

Per tant, per valors donats $(s,y) \in [0,T) \times \mathbb{R}^n$, el problema general de control òptim en el cas estocàstic consisteix en trobar un quintern $\overline{u}(\cdot) := (\overline{\Omega}, \overline{\mathcal{A}}, \overline{P}, \overline{W}(\cdot), \overline{u}(\cdot)) \in \mathcal{U}'[s,T]$ tal que

$$J(s,y;\overline{u}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}'[s,T]} J(s,y;u(\cdot)). \tag{5.42}$$

Tal com hem fet en capítols anteriors, abans d'atacar el problema, farem unes suposicions que facilitaran la feina. Així doncs, suposem que:

(E1)′ Les aplicacions $b : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$, $f : [0,T]\times\mathbb{R}^n\times U \to \mathbb{R}$ i $h : \mathbb{R}^n \to \mathbb{R}$ són uniformement contínues, i existeix una constant $L > 0$ tal que per una funció $\phi(t,x,u) \in \{b(t,x,u), \sigma(t,x,u), f(t,x,u), h(x)\}$ es satisfà que:

$$\begin{cases} |\phi(t,x,u) - \phi(t,\hat{x},u)| \leq L|x-\hat{x}|, & \forall (t,x,\hat{x},u) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times U, \\ |\phi(t,0,u)| \leq L, & \forall (t,u) \in [0,T] \times U. \end{cases} \tag{5.43}$$

Sota aquest supòsit, ja podem abordar el problema descrit per l'equació (5.42); és a dir, el problema general de control òptim estocàstic. En primer lloc, observem que, sota aquesta suposició, per cada $(s,y) \in [0,T) \times \mathbb{R}^n$ i per cada control $u(\cdot) \in \mathcal{U}'[s,T]$, existeix una única solució $x(\cdot) := x(\cdot; s, y, u(\cdot))$ del sistema anterior (5.40), gràcies al teorema d'existència i unicitat de solucions d'una equació diferencial estocàstica [veure Teorema 2.3 de la pàgina 173 i Teorema 3.1 de la pàgina 178 de [3] per a més detalls sobre l'existència i la unicitat, respectivament]. A més a més, gràcies a aquesta unicitat de la solució $x(\cdot)$, el funcional de cost (5.41) està ben definit.

Anàlogament al cas determinista, podem definir la funció valor $V(\cdot,\cdot)$ del problema fonamental de control òptim estocàstic de la següent manera:

$$\begin{cases} V(s,y) = \inf_{u(\cdot)\in\mathcal{U}'[s,T]} J(s,y;u(\cdot)), & \forall (s,y)\in[0,T)\times\mathbb{R}^n, \\ V(T,y) = h(y), & \forall y\in\mathbb{R}^n. \end{cases} \tag{5.44}$$

A continuació, postularem la versió estocàstica del principi d'optimalitat de Bellman o principi de la programació dinàmica de Bellman (teorema anàleg al Teorema 5.1 del cas determinista).

Teorema 5.3. *Suposem que es satisfà (E1)′. Aleshores, per qualsevol* $(s,y)\in[0,T)\times\mathbb{R}^n$,

$$V(s,y) = \inf_{u(\cdot)\in\mathcal{U}'[s,T]} E\left\{ \int_s^{\hat{s}} f(t,x(t;s,y,u(\cdot)),u(t))dt + V\big(\hat{s},x(\hat{s};s,y,u(\cdot))\big)\right\}, \quad \forall\, 0\le s\le\hat{s}\le T. \tag{5.45}$$

<u>*Demostració.*</u> Sigui $\overline{V}(s,y)$ la part dreta de la igualtat de l'equació (5.45); és a dir,

$$\overline{V}(s,y) = \inf_{u(\cdot)\in\mathcal{U}'[s,T]} E\left\{ \int_s^{\hat{s}} f(t,x(t;s,y,u(\cdot)),u(t))dt + V\big(\hat{s},x(\hat{s};s,y,u(\cdot))\big)\right\}. \tag{5.46}$$

Per la definició (5.44) de la funció $V(\cdot,\cdot)$, podem dir que

$$\begin{aligned} V(s,y) &= \inf_{u(\cdot)\in\mathcal{U}'[s,T]} J(s,y;u(\cdot)) \le J(s,y;u(\cdot)) = E\left\{\int_s^T f(t,x(t),u(t))dt + h(x(T))\right\} \\ &= E\left\{\int_s^{\hat{s}} f(t,x(t),u(t))dt + \int_{\hat{s}}^T f(t,x(t),u(t))dt + h(x(T))\right\} \\ &= E\left\{\int_s^{\hat{s}} f(t,x(t),u(t))dt + J(\hat{s},x(\hat{s});u(\cdot))\right\}, \quad \forall u(\cdot)\in\mathcal{U}'[s,T]. \end{aligned} \tag{5.47}$$

Per tant, tenim una desigualtat que és vàlida $\forall u(\cdot)\in\mathcal{U}'[s,T]$. Prenent l'ínfim sobre $u(\cdot)\in\mathcal{U}'[s,T]$ a l'última igualtat de l'equació (5.47) i recordant la definició (5.46) de $\overline{V}(s,y)$, obtenim que

$$V(s,y)\le\overline{V}(s,y). \tag{5.48}$$

En canvi, si prenem un $\epsilon>0$ qualsevol, podem dir que existeix un control $(\Omega,\mathcal{A},P,W(\cdot),u(\cdot))\in\mathcal{U}'[s,T]$, de manera que

$$V(s,y)+\epsilon > J(s,y;u(\cdot)). \tag{5.49}$$

A més a més, emprant la definició (5.41) de $J(s,y;u(\cdot))$, podem dir que

$$J(s,y;u(\cdot)) = E\left\{\int_s^T f(t,x(t;s,y,u(\cdot)),u(t))dt + h(x(T;s,y,u(\cdot)))\right\}. \tag{5.50}$$

Separant la integral anterior, i usant que l'esperança és un operador lineal, escrivim que

$$E\bigg\{\int_s^T f(t,x(t;s,y,u(\cdot)),u(t))dt + h(x(T;s,y,u(\cdot)))\bigg\} = E\bigg\{\int_s^{\hat{s}} f(t,x(t;s,y,u(\cdot)),u(t))dt\bigg\}$$
$$+ E\bigg\{\int_{\hat{s}}^T f(t,x(t;s,y,u(\cdot)),u(t))dt + h(x(T;s,y,u(\cdot)))\bigg\}. \tag{5.51}$$

Podem canviar la condició inicial i tornar a usar la definició (5.41) per dir que

$$\begin{aligned} &E\bigg\{\int_{\hat{s}}^T f(t,x(t;s,y,u(\cdot)),u(t))dt + h(x(T;s,y,u(\cdot)))\bigg\} \\ &= E\bigg\{\int_{\hat{s}}^T f(t,x(t;\hat{s},x(\hat{s}),u(\cdot)),u(t))dt + h(x(T;\hat{s},x(\hat{s}),u(\cdot)))\bigg\} \\ &= J(\hat{s},x(\hat{s};s,y,u(\cdot));u(\cdot)). \end{aligned} \tag{5.52}$$

Així, unint l'equació (5.52) amb les equacions (5.50) i (5.51), escrivim que

$$J(s,y;u(\cdot)) = E\bigg\{\int_s^{\hat{s}} f(t,x(t;s,y,u(\cdot)),u(t))dt\bigg\} + J(\hat{s},x(\hat{s};s,y,u(\cdot));u(\cdot)). \tag{5.53}$$

Finalment, usant la definició de $V(s,y)$ de l'equació (5.44), juntament amb l'equació anterior (5.53), podem dir que

$$J(s,y;u(\cdot)) \geq E\bigg\{\int_s^{\hat{s}} f(t,x(t;s,y,u(\cdot)),u(t))dt + V(\hat{s},x(\hat{s};s,y,u(\cdot)))\bigg\}. \tag{5.54}$$

Si observem la definició de $\overline{V}(s,y)$, veiem que

$$J(s,y;u(\cdot)) \geq E\bigg\{\int_s^{\hat{s}} f(t,x(t;s,y,u(\cdot)),u(t))dt + V(\hat{s},x(\hat{s};s,y,u(\cdot)))\bigg\} \geq \overline{V}(s,y). \tag{5.55}$$

Per tant, unint les equacions (5.49) i (5.55), obtenim que, per un valor $\epsilon > 0$ qualsevol, es satisfà la següent relació:

$$V(s,y) + \epsilon \geq \overline{V}(s,y). \tag{5.56}$$

Fent ϵ tan petit com es vulgui, s'obté que

$$V(s,y) \geq \overline{V}(s,y). \tag{5.57}$$

Així doncs, observant les equacions (5.48) i (5.57), s'obté que $V(s,y) = \overline{V}(s,y)$; és a dir, s'obté, tal com volíem demostrar, la igualtat (5.45). $\square$

De la mateixa manera que en el cas determinista, a continuació, farem una observació (anàloga a l'Observació 13) molt interessant sobre aquest teorema, on usarem, principalment, les definicions (5.41) i (5.44) i el concepte de control i trajectòria òptims.

Observació 15. Suposem que, donat un problema de control òptim estocàstic, tenim un control òptim $\overline{u}(\cdot)$ i la corresponent trajectòria òptima $\overline{x}(\cdot)$. Sigui $\hat{s} \in (s,T)$ un valor qualsevol. Aleshores,

$$\begin{aligned} V(s,y) = \inf_{u(\cdot)\in\mathcal{U}'[s,T]} J(s,y;u(\cdot)) = J(s,y;\overline{u}(\cdot)) = E\bigg\{ \int_s^{\hat{s}} f(t,\overline{x}(t),\overline{u}(t))dt + J(\hat{s},\overline{x}(\hat{s});\overline{u}(\cdot))\bigg\} \\ \geq E\bigg\{ \int_s^{\hat{s}} f(t,\overline{x}(t),\overline{u}(t))dt + V(\hat{s},\overline{x}(\hat{s}))\bigg\} \geq V(s,y), \end{aligned} \tag{5.58}$$

on la darrera desigualtat es deu a l'equació (5.45). Per tant, tota l'equació (5.58) ha de consistir en igualtats encadenades i, en particular, sobté que

$$V(\hat{s},\overline{x}(\hat{s})) = J(\hat{s},\overline{x}(\hat{s});\overline{u}(\cdot)) = E\bigg\{ \int_{\hat{s}}^{T} f(t,\overline{x}(t),\overline{u}(t))dt + h(\overline{x}(T))\bigg\}, \quad \forall \hat{s} \in (s,T). \tag{5.59}$$

El Teorema 5.3 és el principi de la programació dinàmica de Bellman pel cas estocàstic. Aquest teorema ens ha proporcionat, per mitjà de la funció valor, una relació entre la família de problemes de control òptim donats per l'equació (5.42). No obstant això, es tracta d'una equació molt complicada, perquè les operacions que hi ha a la part dreta de la igualtat ($\overline{V}(s,y)$) són difícils. Per aquesta raó, el nostre objectiu d'ara és trobar una equació més senzilla per $V(s,y)$.

Per dur a terme aquest nou objectiu, sigui $\mathcal{C}^{1,2}([0,T]\times\mathbb{R}^n) := \{v : [0,T]\times\mathbb{R}^n \to \mathbb{R} | v$ és una funció contínua i v_t, v_x i v_{xx} són contínues en $(t,x)\}$; és a dir, el conjunt de les funcions contínues amb primera derivada temporal contínua i amb primera i segona derivada espacials contínues.

Proposició 5.4. *Suposem que es satisfà (E1)$'$. Suposem que la funció valor V (definida a l'equació (5.44)) satisfà que $V \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^n)$. Aleshores V és una solució del següent problema de valors finals d'una equació en derivades parcials de segon ordre (possiblement degenerada):*

$$\begin{cases} -v_t + \sup_{u\in U} G(t,x,u,-v_x,-v_{xx}) = 0, & (t,x)\in[0,T)\times\mathbb{R}^n, \\ v|_{t=T} = h(x), \quad x\in\mathbb{R}^n, \end{cases} \tag{5.60}$$

on la funció $G : [0,T]\times\mathbb{R}^n\times U\times\mathbb{R}^n\times\mathcal{S}^n \to \mathbb{R}$ es defineix com:

$$G(t,x,u,p,P) := \frac{1}{2}\mathrm{tr}\left(P\sigma(t,x,u)\sigma(t,x,u)^\top\right) + \langle p, b(t,x,u)\rangle - f(t,x,u). \tag{5.61}$$

L'equació (5.60) és la famosa equació Hamilton-Jacobi-Bellman (HJB) associada al problema fonamental de control òptim estocàstic. La funció G definida a l'equació (5.61) és el Hamiltonià generalitzat i és la mateixa funció que la definida en el capítol anterior en (4.48).

Observació 16. Comparant l'equació (5.60) amb l'equació (5.19), veiem que el problema (5.60) és de segon ordre, permetent, així, una possible degeneració. A més, (5.60) és una clara generalització de (5.19), ja que en el cas $\sigma(t,x,u) = 0$ l'equació (5.60) es redueix a l'equació (5.19).

Demostració de la Proposició 5.4. En primer lloc, fixem-nos que la segona relació de l'equació (5.60) és immediata si utilitzem la segona relació de la definició (5.44). Així doncs, només hem de demostrar la primera relació de l'equació (5.60). Fixem $(s, y) \in [0, T) \times \mathbb{R}^n$ i $u \in U$. Sigui $x(\cdot)$ la trajectòria d'estat corresponent al control $u(\cdot) \in \mathcal{U}'[s, T]$ pel problema general de control òptim estocàstic amb $u(t) = u$.

A continuació, farem tendir $\hat{s}$ a s (sempre conservant que $\hat{s} - s > 0$, i ho representarem per $\hat{s} \downarrow s$). Per la definició de l'equació (5.45), es pot dir que, $\forall u \in U$,

$$V(s,y) = \inf_{u(\cdot)\in\mathcal{U}'[s,T]} E\Big\{ \int_s^{\hat{s}} f(t, x(t), u)dt + V(\hat{s}, x(\hat{s})) \Big\} \leq E\Big\{ \int_s^{\hat{s}} f(t, x(t), u)dt + V(\hat{s}, x(\hat{s})) \Big\}. \tag{5.62}$$

Observant l'última desigualtat, recordant el Lema d'Itô, tenint en ment que l'esperança és un operador lineal i considerant el cas en què $\hat{s} \downarrow s$, podem escriure que, $\forall u \in U$,

$$\begin{aligned} 0 &\geq -\frac{E\big\{V(\hat{s}, x(\hat{s})) - V(s,y)\big\}}{\hat{s} - s} - \frac{1}{\hat{s} - s} E\Big\{ \int_s^{\hat{s}} f(t, x(t), u)dt \Big\} \\ &= \frac{1}{\hat{s} - s} E\Big\{ \int_s^{\hat{s}} \big\{ -V_t(t, x(t)) + G\big(t, x(t), u(t), -V_x(t, x(t)), -V_{xx}(t, x(t))\big) \big\} dt \Big\} \\ &\longrightarrow -V_t(s,y) + G(s, y, u, -V_x(s,y), -V_{xx}(s,y)), \end{aligned} \tag{5.63}$$

on hem emprat la definició (5.61) del Hamiltonià generalitzat.

Per acabar, com el resultat anterior és vàlid $\forall u \in U$, en particular deduim que

$$0 \geq -V_t(s,y) + \sup_{u\in U} G(s, y, u, -V_x(s,y), -V_{xx}(s,y)). \tag{5.64}$$

Per altra banda, de manera similar al que vam fer a l'equació (5.49), si prenem un $\epsilon > 0$ qualsevol i un $\hat{s}$ qualsevol tal que $0 \leq s < \hat{s} \leq T$ amb $\hat{s} - s > 0$ suficientment petit, podem dir que existeix un control $u(\cdot) := u_{\epsilon,\hat{s}}(\cdot) \in \mathcal{U}'[s, T]$ tal que

$$V(s,y) + \epsilon(\hat{s} - s) \geq E\Big\{ \int_s^{\hat{s}} f(t, x(t), u(t))dt + V(\hat{s}, x(\hat{s})) \Big\}. \tag{5.65}$$

Reordenant la desigualtat i recordant que l'esperança és un operador lineal, podem escriure que

$$\begin{aligned} -\epsilon &\leq -\frac{E\big\{V(\hat{s}, x(\hat{s})) - V(s,y)\big\}}{\hat{s} - s} - \frac{1}{\hat{s} - s} E\Big\{ \int_s^{\hat{s}} f(t, x(t), u(t))dt \Big\} \\ &= \frac{1}{\hat{s} - s} E\Big\{ \int_s^{\hat{s}} \big\{ -V_t(t, x(t)) + G\big(t, x(t), u(t), -V_x(t, x(t)), -V_{xx}(t, x(t))\big) \big\} dt \Big\} \\ &\leq \frac{1}{\hat{s} - s} E\Big\{ \int_s^{\hat{s}} \big\{ -V_t(t, x(t)) + \sup_{u\in U} G\big(t, x(t), u, -V_x(t, x(t)), -V_{xx}(t, x(t))\big) \big\} dt \Big\} \\ &\longrightarrow -V_t(s,y) + \sup_{u\in U} G(s, y, u, -V_x(s,y), -V_{xx}(s,y)). \end{aligned} \tag{5.66}$$

Com ϵ és arbitrari, el podem fer tan petit com desitgem, i al final obtindrem que

$$0 \leq -V_t(s,y) + \sup_{u\in U} G(s, y, u, -V_x(s,y), -V_{xx}(s,y)). \tag{5.67}$$

Finalment, combinant les equacions (5.64) i (5.67), obtenim, tal com volíem demostrar, que

$$0 = -V_t(s, y) + \sup_{u \in U} G(s, y, u, -V_x(s, y), -V_{xx}(s, y)). \tag{5.68}$$

□

6 Comparació entre els dos mètodes

Un cop hem estudiat en profunditat el principi del màxim de Pontryagin i el principi de la programació dinàmica de Bellman, prosseguim a fer una comparativa entre ambdós mètodes.

6.1 Cas determinista

Començarem parlant de les diferències que es veuen a simple vista entre els dos mètodes.

En primer lloc, la primera diferència que salta a la vista és que, mentres en el principi del màxim de Pontryagin només es treballa amb el problema fonamental de control òptim, en el principi de la programació dinàmica de Bellman s'introdueix el que hem anomenat com el problema general de control òptim; és a dir, una família de problemes parametritzada per unes condicions inicials $(s, y) \in [0, T] \times \mathbb{R}^n$.

En segon lloc, és molt important el fet que la programació dinàmica de Bellman ens proporciona els passos que calen seguir per obtenir un control òptim. En canvi, el principi del màxim de Pontryagin només ens dóna les condicions que satisfan un cert control òptim; és a dir, condicions necessàries. En cap cas es pot saber si un control és òptim o no, i ni tan sols hi ha un receptari per trobar aquest control òptim. No és fins el moment en què afegim un tercer supòsit que podem començar a parlar de condicions suficients.

Ara que ja hem parlat dels supòsits, arriba la tercera diferència que s'aprecia a simple vista. En el principi del màxim de Pontryagin, sense comptar el tercer supòsit que cal afegir per poder saber si un control és òptim o no (condició de suficiència), es fan dues suposicions: (D1) i (D2). En canvi, en el principi de la programació dinàmica de Bellman només se'n fa una: (D1)$'$. No obstant, observem que (D1)$'$ és més forta que (D1), ja que exigeix la continuïtat de les funcions respecte (t, x, u), mentres que (D1) no demana res sobre t. [Veure pàgines 14 i 28.]

Un cop comentades les diferències més bàsiques, veurem quina similitud hi ha entre els dos mètodes. Per fer-ho, partirem del formalisme del principi de la programació de Bellman. Així doncs, partim del següent sistema de control:

$$\begin{cases} \dot{x}(t) = b(t, x(t), u(t)), \quad \forall t \in [s, T] \\ x(s) = y, \end{cases} \tag{6.1}$$

amb el funcional de cost associat

$$J(s, y; u(\cdot)) = \int_s^T f(t, x(t), u(t))dt + h(x(T)), \tag{6.2}$$

on $(s, y) \in [0, T) \times \mathbb{R}^n$ representen les condicions inicials de temps i d'estat.

Recordem que el nostre objectiu d'ara és resoldre el problema general de control òptim en el cas determinista, que consisteix en trobar, sota els lligams de l'equació (6.1), un control òptim $\overline{u}(\cdot) \in \mathcal{U}[s, T]$ tal que

$$J(s, y; \overline{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[s,T]} J(s, y; u(\cdot)), \tag{6.3}$$

on $\mathcal{U}[s, T] \coloneqq \{u : [s, T] \to U | u(\cdot) \text{ és mesurable}\}$.

Prosseguint de la mateixa manera que sempre, toca fer les suposicions que ens facilitaran la feina. Tal com hem fet quan hem estudiat la programació dinàmica de Bellman, suposarem i (D1)$'$ [veure pàgina 28]. No obstant, com a novetat, afegirem ara el següent supòsit:

(D2)$'$ Les funcions b, f, h compleixen que $b, f, h \in \mathcal{C}^1$ respecte x, i existeix un mòdul de continuïtat $\overline{\omega} : [0, \infty) \to [0, \infty)$ tal que per una funció $\phi(t, x, u) \in \{b(t, x, u), f(t, x, u), h(x)\}$ es satisfà que:

$$|\phi_x(t, x, u) - \phi_x(t, \hat{x}, u)| \leq \overline{\omega}(|x - \hat{x}|), \quad \forall (t, x, \hat{x}, u) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times U. \tag{6.4}$$

Aquest supòsit (D2)$'$ és menys restrictiu que el supòsit (D2) del principi del màxim de Pontryagin, ja que en el primer no es demana continuïtat respecte u.

A continuació, saltarem al formalisme del principi del màxim de Pontryagin, però sense oblidar que tenim una família de problemes parametritzada per $(s, y) \in [0, T) \times \mathbb{R}^n$. D'aquesta manera, podem reescriure el sistema de control (6.1), l'equació adjunta (amb la funció adjunta $p(\cdot)$) i la condició de maximalitat [veure pàgina 15] de la següent manera:

$$\begin{cases} \dot{x}(t) = H_p(t, x(t), u(t), p(t)), \quad \forall t \in [s, T], \\ \dot{p}(t) = -H_x(t, x(t), u(t), p(t)), \quad \forall t \in [s, T], \\ x(s) = y, \quad p(T) = -h_x(x(T)), \\ H(t, x(t), u(t), p(t)) = \max_{u \in U} H(t, x(t), u, p(t)), \quad \forall t \in [s, T], \end{cases} \tag{6.5}$$

on el Hamiltonià $H : [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \to \mathbb{R}$ es defineix com:

$$H(t, x, u, p) := \langle p, b(t, x, u) \rangle - f(t, x, u). \tag{6.6}$$

Tornant de nou al principi de la programació dinàmica de Bellman, recordem que la funció valor $V(\cdot, \cdot)$ es defineix com:

$$\begin{cases} V(s, y) = \inf_{u(\cdot) \in \mathcal{U}[s,T]} J(s, y; u(\cdot)), \quad \forall (s, y) \in [0, T) \times \mathbb{R}^n, \\ V(T, y) = h(y), \quad \forall y \in \mathbb{R}^n. \end{cases} \tag{6.7}$$

Finalment, segons la Proposició 5.2, sabem que si la funció valor $V \in \mathcal{C}^1([0, T] \times \mathbb{R}^n)$, aleshores és solució de l'equació HJB; és a dir, és solució del següent sistema en derivades parcials:

$$\begin{cases} -v_t(t, x) + \sup_{u \in U} H(t, x, u, -v_x(t, x)) = 0, \quad (t, x) \in [0, T) \times \mathbb{R}^n, \\ v|_{t=T} = h(x), \quad x \in \mathbb{R}^n. \end{cases} \tag{6.8}$$

Amb tot això, i després d'aquest petit repàs, podem postular el següent resultat que ens estableix una relació entre els dos mètodes emprats: el principi del màxim de Pontryagin i el principi de la programació dinàmica de Bellman.

Teorema 6.1. *Suposem que es satisfà (D1)′ i (D2)′. Sigui* $(s, y) \in [0, T) \times \mathbb{R}^n$ *condicions inicials fixades de temps i d'estat. Sigui* $\overline{u}(\cdot)$ *un control òptim,* $\overline{x}(\cdot)$ *la corresponent trajectòria òptima i* $p(\cdot)$ *la corresponent funció adjunta d'un problema general de control òptim determinista. Suposem, també, que la funció valor* $V \in \mathcal{C}^1([0, T] \times \mathbb{R}^n)$*. Aleshores,*

$$\begin{aligned} V_t(t, \overline{x}(t)) &= H(t, \overline{x}(t), \overline{u}(t), -V_x(t, \overline{x}(t))) \\ &= \max_{u \in U} H(t, \overline{x}(t), u, -V_x(t, \overline{x}(t))), \quad \forall t \in [s, T]. \end{aligned} \tag{6.9}$$

A més a més, si $V \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n)$ *i* V_{tx} *és contínua, aleshores*

$$-V_x(t, \overline{x}(t)) = p(t), \quad \forall t \in [s, T]. \tag{6.10}$$

Demostració. Com $\overline{u}(\cdot)$ i $\overline{x}(\cdot)$ són òptims, podem dir que

$$V(t, \overline{x}(t)) = J(t, \overline{x}(t); \overline{u}(\cdot)) = \int_t^T f(r, \overline{x}(r), \overline{u}(r))dr + h(\overline{x}(T)), \quad \forall t \in [s, T]. \tag{6.11}$$

Derivant respecte el temps a banda i banda de l'equació anterior (6.11), obtenim que

$$\frac{d}{dt} V(t, \overline{x}(t)) = V_t(t, \overline{x}(t)) + \langle V_x(t, \overline{x}(t)), b(t, \overline{x}(t), \overline{u}(t)) \rangle = -f(t, \overline{x}(t), \overline{u}(t)). \tag{6.12}$$

Usant la definició (6.6) del Hamiltonià, podem reescriure l'equació anterior (6.12) com

$$V_t(t, \overline{x}(t)) = H(t, \overline{x}(t), \overline{u}(t), -V_x(t, \overline{x}(t))), \quad \forall t \in [s, T]; \tag{6.13}$$

és a dir, queda demostrada, tal com volíem, la primera igualtat de l'equació (6.9).

Per altra banda, com la funció valor $V \in \mathcal{C}^1([0, T] \times \mathbb{R}^n)$, gràcies a la Proposició 5.2, sabem que V satisfà l'equació HJB (6.8); és a dir, queda demostrada, tal com volíem, la segona igualtat de l'equació (6.9).

Tot seguit, si combinem l'equació (6.13) amb l'equació HJB (6.8), veiem que, $\forall x \in \mathbb{R}^n$,

$$0 = H(t, \overline{x}(t), \overline{u}(t), -V_x(t, \overline{x}(t))) - V_t(t, \overline{x}(t)) \geq H(t, x, \overline{u}(t), -V_x(t, x)) - V_t(t, x). \tag{6.14}$$

A més a més, si suposem que $V \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n)$ i que V_{tx} és també contínua, podem derivar la igualtat de (6.14) i dir que

$$0 = \frac{\partial}{\partial x} \{ H(t, x, \overline{u}(t), -V_x(t, x)) - V_t(t, x) \} \Big|_{x = \overline{x}(t)}. \tag{6.15}$$

Observem que podem trobar una manera alternativa d'escriure l'equació anterior si partim de l'equació (6.12) i la derivem (parcialment) respecte x a banda i banda, de manera que obtenim que

$$\begin{aligned} V_{tx}(t, \overline{x}(t)) + \langle V_{xx}(t, \overline{x}(t)), &b(t, \overline{x}(t), \overline{u}(t)) \rangle + \langle V_x(t, \overline{x}(t)), b_x(t, \overline{x}(t), \overline{u}(t)) \rangle \\ &= -f_x(t, \overline{x}(t), \overline{u}(t)). \end{aligned} \tag{6.16}$$

Finalment, reordenant termes, podem escriure

$$\begin{aligned} &f_x(t, \overline{x}(t), \overline{u}(t)) + \langle V_x(t, \overline{x}(t)), b_x(t, \overline{x}(t), \overline{u}(t)) \rangle \\ &= -\Big(V_{tx}(t, \overline{x}(t)) + \langle V_{xx}(t, \overline{x}(t)), b(t, \overline{x}(t), \overline{u}(t)) \rangle \Big) = -\frac{d}{dt} V_x(t, \overline{x}(t)). \end{aligned} \tag{6.17}$$

Per tant,

$$
\begin{aligned}
-\frac{d}{dt}V_x(t,\overline{x}(t)) &= f_x(t,\overline{x}(t),\overline{u}(t)) + \langle V_x(t,\overline{x}(t)), b_x(t,\overline{x}(t),\overline{u}(t))\rangle \\
&= -H_x(t,\overline{x}(t),\overline{u}(t),-V_x(t,\overline{x}(t))).
\end{aligned}
\tag{6.18}
$$

A més a més, observant l'equació (6.11), és obvi que

$$-V_x(T,\overline{x}(T)) = -h_x(\overline{x}(T)). \tag{6.19}$$

Així doncs, la funció $-V_x(t,\overline{x}(t))$ satisfà tant l'equació (6.18) com l'equació (6.19); és a dir, satisfà l'equació adjunta (6.5). Per tant, com sabem que, donat un control òptim i la seva corresponent trajectòria òptima, hi ha unicitat en la solució de l'equació adjunta, aleshores arribem a la conclusió que

$$-V_x(t,\overline{x}(t)) = p(t), \quad \forall t \in [s,T], \tag{6.20}$$

tal com volíem demostrar. □

És molt interessant observar que la segona igualtat de (6.9) (combinada amb (6.10)) és exactament la condició de maximalitat que trobem en el principi del màxim de Pontryagin. Per tant, acabem de demostrar que, sota uns certs supòsits, el principi del màxim de Pontryagin es pot deduir de manera directa del principi de programació dinàmica de Bellman.

6.2 Cas estocàstic

Igual que en la secció anterior, podem mencionar les diferències que es poden apreciar a simple vista entre els dos mètodes en el cas estocàstic. Concretament, les dues primeres dissimilituds que hem tractat en el cas determinista també les tenim en aquest cas: la relació entre el problema fonamental i el problema general de control òptim, i la relació entre les condicions necessàries i les suficients.

Per altra banda, en relació als supòsits, en el principi del màxim de Pontryagin se'n fan dos: (E1) i (E2); mentres que en el principi de la programació dinàmica de Bellman només se'n fa un: (E1)$'$. Igual que en el cas determinista, observem que (E1)$'$ és més fort que (E1), ja que exigeix la continuïtat de les funcions respecte (t, x, u), mentres que (E1) no demana res sobre t. [Veure pàgines 21 i 35.]

Tot seguit, seguint el mateix esquema que abans, veurem quina similitud hi ha entre els dos mètodes. Per fer-ho, partirem del formalisme del principi de la programació dinàmica de Bellman. Així doncs, partim del següent sistema de control:

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), \quad \forall t \in [s, T], \\ x(s) = y, \end{cases} \tag{6.21}$$

amb el respectiu funcional de cost

$$J(s, y; u(\cdot)) = E\left\{ \int_s^T f(t, x(t), u(t))dt + h(x(T)) \right\}, \tag{6.22}$$

on $(s, y) \in [0, T) \times \mathbb{R}^n$ representen les condicions inicials de temps i d'estat.

Recordem que, amb la nova formulació introduïda, el nostre objectiu d'ara és resoldre el problema general de control òptim en el cas estocàstic, que consisteix en trobar un quintern $\overline{u}(\cdot) := (\overline{\Omega}, \overline{\mathcal{A}}, \overline{P}, \overline{W}(\cdot), \overline{u}(\cdot)) \in \mathcal{U}'[s, T]$ tal que

$$J(s, y; \overline{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}'[s,T]} J(s, y; u(\cdot)), \tag{6.23}$$

on $\mathcal{U}'[s, T]$ ho hem definit a la pàgina 35.

En relació a les suposicions que cal fer, només suposarem (E1)$'$ [veure pàgina 35].

A continuació, saltarem al formalisme del principi del màxim de Pontryagin, però sense oblidar que tenim una família de problemes parametritzada per $(s, y) \in [0, T) \times \mathbb{R}^n$. Per tant, tal com hem vist a les pàgines 22 i 23, sabem que existeixen processos de primer ordre $p(\cdot) \in \mathbb{R}^n$ i $q(\cdot) \in \mathbb{R}^{n \times m}$ que satisfan

$$\begin{cases} dp(t) = -\Big\{ b_x(t, \overline{x}(t), \overline{u}(t))^\top p(t) + \sum_{j=1}^m \sigma_x^j(t, \overline{x}(t), \overline{u}(t))^\top q_j(t) \\ \qquad\qquad - f_x(t, \overline{x}(t), \overline{u}(t)) \Big\} dt + q(t)dW(t), \quad \forall t \in [s, T], \\ p(T) = -h_x(\overline{x}(T)), \end{cases} \tag{6.24}$$

i processos de segon ordre $P(\cdot) \in \mathcal{S}^n$ i $Q(\cdot) \in \mathcal{S}^{n\times m}$ que satisfan

$$\begin{cases} dP(t) = -\Big\{ b_x(t,\overline{x}(t),\overline{u}(t))^\top P(t) + P(t) b_x(t,\overline{x}(t),\overline{u}(t)) \\ \qquad + \sum_{j=1}^{m} \sigma_x^j(t,\overline{x}(t),\overline{u}(t))^\top P(t) \sigma_x^j(t,\overline{x}(t),\overline{u}(t)) \\ \qquad + \sum_{j=1}^{m} \Big\{ \sigma_x^j(t,\overline{x}(t),\overline{u}(t))^\top Q_j(t) + Q_j(t)\sigma_x^j(t,\overline{x}(t),\overline{u}(t)) \Big\} \\ \qquad + H_{xx}(t,\overline{x}(t),\overline{u}(t),p(t),q(t)) \Big\} dt + \sum_{j=1}^{m} Q_j(t) dW^j(t), \quad \forall t \in [s,T], \\ P(T) = -h_{xx}(\overline{x}(T)), \end{cases} \tag{6.25}$$

on $p(\cdot)$ i $q(\cdot)$ són solució del sistema anterior (6.24) i on el Hamiltonià estocàstic $H : [0,T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n\times m} \to \mathbb{R}$ es defineix com

$$H(t,x,u,p,q) = \langle p, b(t,x,u)\rangle + \operatorname{tr}\Big(q^\top \sigma(t,x,u)\Big) - f(t,x,u). \tag{6.26}$$

Així doncs, tornant de nou al principi de la programació dinàmica de Bellman, recordem que la funció valor $V(\cdot,\cdot)$ es defineix com:

$$\begin{cases} V(s,y) = \inf_{u(\cdot)\in\mathcal{U}'[s,T]} J(s,y;u(\cdot)), \quad \forall (s,y) \in [0,T) \times \mathbb{R}^n, \\ V(T,y) = h(y), \quad \forall y \in \mathbb{R}^n. \end{cases} \tag{6.27}$$

Finalment, segons la Proposició 5.4, sabem que si la funció valor $V \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^n)$, aleshores és solució de l'equació HJB pel cas estocàstic; és a dir, és solució del següent sistema en derivades parcials de segon ordre:

$$\begin{cases} -v_t(t,x) + \sup_{u\in U} G(t,x,u,-v_x(t,x),-v_{xx}(t,x)) = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n, \\ v|_{t=T} = h(x), \quad x \in \mathbb{R}^n, \end{cases} \tag{6.28}$$

on el Hamiltonià generalitzat $G : [0,T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathcal{S}^n \to \mathbb{R}$ es defineix com:

$$G(t,x,u,p,P) \coloneqq \frac{1}{2} tr\Big(P\sigma(t,x,u)\sigma(t,x,u)^\top\Big) + \langle p, b(t,x,u)\rangle - f(t,x,u). \tag{6.29}$$

Amb tot això, i després d'aquest petit repàs, podem postular el següent resultat que ens estableix una relació entre els dos mètodes emprats pel cas estocàstic. Per tant, es tracta d'un resultat completament anàleg al Teorema 6.1, però en versió estocàstica.

Teorema 6.2. *Suposem que es satisfà (E1)$'$. Siguin $(s,y) \in [0,T)\times\mathbb{R}^n$ condicions inicials fixades de temps i d'estat. Sigui $\overline{u}(\cdot)$ un control òptim, $\overline{x}(\cdot)$ la corresponent trajectòria òptima i $p(\cdot)$ i $q(\cdot)$ les corresponents solucions adaptades de l'equació (6.24) d'un problema general de control òptim estocàstic. Suposem, també, que la funció valor $V \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^n)$. Aleshores,*

$$\begin{aligned} V_t(t,\overline{x}(t)) &= G(t,\overline{x}(t),\overline{u}(t),-V_x(t,\overline{x}(t)),-V_{xx}(t,\overline{x}(t))) \\ &= \max_{u\in U} G(t,\overline{x}(t),u,-V_x(t,\overline{x}(t)),-V_{xx}(t,\overline{x}(t))), \quad \forall t \in [s,T]. \end{aligned} \tag{6.30}$$

A més a més, si $V \in \mathcal{C}^{1,3}([0,T] \times \mathbb{R}^n)$ *i* V_{tx} *és contínua, aleshores*

$$\begin{cases} -V_x(t, \overline{x}(t)) = p(t), \quad \forall t \in [s,T], \\ -V_{xx}(t, \overline{x}(t))\sigma(t, \overline{x}(t), \overline{u}(t)) = q(t), \quad \forall t \in [s,T]. \end{cases} \tag{6.31}$$

Demostració. No donarem una demostració completa d'aquest teorema. Com a idea general, la demostració segueix una estructura molt similar a la del Teorema 6.1, però amb certs tecnicismes que l'eleven a un nivell de dificultat superior al del treball. No obstant, una demostració completa es pot trobar a [6], a les pàgines 251-253. □

Establint analogies amb el cas determinista, hauríem de relacionar la segona igualtat de l'equació (6.30) (combinada amb (6.31)) amb el principi del màxim de Pontryagin pel cas estocàstic. No obstant, observem que és totalment diferent del principi del màxim de Pontryagin que nosaltres hem donat en l'equació (4.52), ja que, a diferència d'aquí, no s'involucraven les derivades de la funció valor.

7 Conclusions

Tal com hem definit a la secció **1.3 Objectiu del treball**, la intenció d'aquest projecte era estudiar quina relació hi ha entre els dos mètodes clàssics de resolució d'un problema de control òptim: el mètode del màxim de Pontryagin i el mètode de la programació dinàmica de Bellman; o de manera equivalent, veure quina és la relació que s'estableix (en el cas determinista/estocàstic) entre la condició de maximalitat de Pontryagin (EDO/EDE) i les equacions HJB de la programació dinàmica de Bellman (EDP de primer/segon ordre).

Després de començar amb una introducció del càlcul estocàstic, en el tercer capítol hem explicat en què consisteix el problema de control òptim, tant en el cas determinista com en el cas estocàstic. Amb la idea d'arribar a fer una comparació entre els dos mètodes de resolució, hem partit d'una formulació general; és a dir, hem presentat el mateix problema dins del mateix marc de treball, per després abordar-lo per dos camins diferents.

Tal com hem vist en el quart capítol, el principi del màxim de Pontryagin només ens proporciona condicions necessàries; és a dir, les condicions que satisfan un cert control òptim. No és fins el moment en què afegim un tercer supòsit que podem començar a parlar de condicions suficients.

En canvi, tal com hem vist en el cinquè capítol, el principi de la programació dinàmica de Bellman, a diferència del mètode anterior, sí que dóna condicions suficients, i inclús un receptari per poder trobar el control òptim.

Finalment, en el darrer capítol, hem realitzat la comparació entre ambdós mètodes, i és quan hem pogut donar resposta a les preguntes plantejades a la introducció. Tal com hem vist en el Teorema 6.1, sota uns certs supòsits i en el cas determinista, veiem que el principi del màxim de Pontryagin es pot deduir directament del principi de la programació dinàmica de Bellman. Per tant, dit amb altres paraules, veiem que podem deduir la condició de maximalitat (EDO) partint de les equacions HJB (EDP de primer ordre). No obstant, en canvi, en el cas estocàstic, partint del principi de la programació dinàmica de Bellman, no es pot arribar exactament al principi del màxim de Pontryagin, ja que apareixen involucrades derivades de la funció valor que no s'haurien de manifestar. Tot i així, tal com mostra el Teorema 6.2, sí que es pot donar una condició de maximalitat semblant.

Per tant, com a resum i conclusió final, podem dir que, després d'explicar el problema de control òptim juntament amb els seus dos mètodes de resolució, hem vist que, en el cas determinista, el principi de la programació dinàmica de Bellman és més potent, ja que no només presenta condicions suficients sinó que també permet deduir, sota uns certs supòsits, l'altre mètode de resolució a partir d'ell. En canvi, en el cas estocàstic, aquesta relació no es pot establir i és útil tenir en ment els dos mètodes, segons convingui estudiar condicions necessàries o suficients.

Referències

[1] R. Carmona. *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications.* Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 2016.

[2] W. H. Fleming, R. W. Rishel. *Deterministic and Stochastic Optimal Control.* New York: Springer-Verlag, 1975.

[3] N. Ikeda, S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, 2nd Edition. Amsterdam - Tokyo: North Holland - Kodansha, 1989.

[4] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko. *The Mathematical Theory of Optimal Processes.* New York: Interscience Publishers, 1962.

[5] A. Seierstad, K. Sydsæter. *Optimal Control Theory with Economic Applications.* Amsterdam: North Holland, 1987.

[6] J. Yong, X. Y. Zhou. *Stochastic Controls. Hamiltonian Systems and HJB Equations.* New York: Springer-Verlag, 1999.