UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA



EXISTENCE RESULTS AND APPLICATIONS FOR SOME NON-LINEAR OPTIMAL CONTROL PROBLEMS

Jorge Filipe Duarte Tiago

DOUTORAMENTO EM MATEMÁTICA

ANÁLISE MATEMÁTICA

2009

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA



EXISTENCE RESULTS AND APPLICATIONS FOR SOME NON-LINEAR OPTIMAL CONTROL PROBLEMS

Jorge Filipe Duarte Tiago

DOUTORAMENTO EM MATEMÁTICA

ANÁLISE MATEMÁTICA

Tese orientada pelo Prof. Doutor Pablo Pedregal
e pelo Prof. Doutor Cristian Barbarosie

2009

To those whose music,

poetry, and love,

fed my spirit over the seasons...

This research was supported by Fundação para a Ciência e Tecnologia through doctoral degree grant SFRH-BD-22107-2005 in the context of the Programa Operacional Ciência e Inovação 2010.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR Portugal

Ciência Inovação 2010

Abstract

Classical existence results for optimal control problems governed by systems of ordinary differential equations are based on typical convexity assumptions, which are quite often, very difficult to check. We present a general approach to prove existence of solutions for optimal control problems, based on several relaxations of the problem, where the convexity arises in an unexpected way. We isolate one sufficient condition for the existence of optimal solutions, which can be validated in various contexts. We end up with a main existence result for vector problems with a particular structure, motivated by underwater-vehicles-maneuvering problems.

Alternatively, we recover the classical approach based on a purely variational reformulation, which can lead to existence results by using fine existence theorems for variational problems without convexity assumptions. In particular we prove the existence of solution for autonomous scalar optimal control problems.

Finally, we apply our existence result for vector state and control variables, to prove the local existence of solution for an optimal control problem describing the control of an underwater vehicle.

Additionally to the main work described above, we introduce some ideas for future work. We propose to implement a numerical method, based on steepest descent directions, to approximate the solutions of realistic optimal control problems. Some preliminary results for academic examples are shown.

Keywords. Existence, non-convex, non-linear, relaxation, variational reformulation, underwater vehicles

Resumo

Os resultados clássicos de existência, para problemas de controlo óptimo, governados por sistemas de equações diferenciais ordinárias, baseam-se en condições de convexidade que resultam frequentemente muito difíceis de verificar. Apresentamos uma abordagem geral para este problema, com base em várias relaxações do mesmo, na qual a convexidade surge de um modo inesperado. Isolamos uma condição suficiente, para a existência de soluções óptimas, que pode ser verificada em vários contextos. Podemos chegar a um resultado de existência para problemas vectoriais, com uma estrutura particular, proveniente de problemas de controlo e manobra de veículos subaquáticos.

Numa abordagem diferente, recuperamos técnicas clássicas baseadas na reformulação variacional, que nos permitem obter resultados de existência através da aplicação de teoremas para problemas variacionais, sem condições de convexidade. Em particular, provamos a existência de solução, no caso escalar, para problemas de controlo óptimo autónomos.

Utilizamos o nosso resultado de existência, demonstrado para problemas de controlo óptimo com variáveis vectoriais, para provar a existência de solução para um problema de manobra de veículos subaquáticos.

Para terminar, apresentamos algumas ideias para trabalho futuro. Propomos a implementação de um método numérico, baseado em direcções de descida mais rápida, para aproximar soluções de problemas de controlo óptimo vindos das aplicações. Mostramos resultados preliminares desta implementação, para alguns exemplos académicos.

Palavras chave. Existência, não convexo, não linear, relaxação, reformulação variacional, veículos subaquáticos

Resumo alargado

Esta dissertação trata da existência de solução para problemas de controlo óptimo governados por sistemas de equações diferenciais ordinárias. Mais concretamente, debruçamo-nos sobre problemas autónomos, com um custo na forma integral, cuja função integranda chamamos F, e cuja dinâmica é descrita por uma função vectorial f. Além da típica condição inicial $x(0) = x_0$ para o estado $x \in \mathbf{R}^N$, consideramos ainda que o controlo $u \in \mathbf{R}^n$ deve tomar valores num conjunto de admissibilidade compacto e convexo que designamos por K.

A existência de solução para estes problemas tem sido tratada por vários autores nas últimas décadas, com resultados bastante conhecidos. Entre eles, podemos destacar dois grupos principais. Por um lado, a categoria de resultados gerais que assentam em condições de convexidade sobre o conjunto $\mathcal{A}_x = \{(v, z) : \ v > F(x, u), \ z = f(x, u), \ u \in K\}$ para cada x fixo, na linha das ideias propostas por Filippov e Roxin. Por outro, os resultados sem condições de convexidade, mas com a dependência nas variáveis de estado separadas por uma estrutura do tipo $F(x, u) = G(x) + H(u)$, tal como foi proposto por Neustadt, Cesari, Raymond, Balder, entre outros.

A razão pela qual decidimos procurar novos resultados de existência prende-se com a vontade de compreender problemas de controlo óptimo, cujas características complexas tornam muito difícil, ou até mesmo impossível, a aplicação dos resultados anteriores. De facto, considerando problemas de controlo óptimo provenientes da modelação de manobras de veículos subaquáticos, deparamo-nos com uma estrutura particular cuja dependência com respeito quer às variáveis de estado, quer às de controlo, é não linear e não necessariamente convexa. Mais concretamente, se consideramos o modelo proposto pela Universidade Politécnica de Cartagena, em colaboração com a construtora naval Navantia, temos uma dinâmica descrita por um sistema de doze equações diferenciais ordinárias, acopladas, o qual pode ser reescrito na forma $x'(t) = Q(x)\Phi(u) + Q_0(x)$. Aqui $\Phi(u) = (u_1, u_2, u_3, (u_1)^2, (u_2)^2, (u_3)^2)$, é uma aplicação vectorial que descreve a forma como as variáveis de controlo $u \in \mathbf{R}^3$ actuam no sistema. A matriz $Q(x)$ agrupa todos os coeficientes que ponderam a actuação do controlo. O vector $Q_0(x)$ agrupa todos os termos independentes das variáveis de controlo. Ambos Q_0 e Q dependem

de maneira não linear das variáveis de estado. Este sistema pode ser controlado dentro de um conjunto de admissibilidade K, de forma a atingir um critério, normalmente uma função custo a minimizar, que penalize o controlo com maior norma L^2, ao mesmo tempo que mede a distancia do estado num tempo final T, fixo, a uma configuração desejada x_T. Esta função custo pode ser reescrita de forma similar ao sistema que descreve a dinâmica, isto é, na forma $\int_0^T c(x)\Phi(u) + c_0(x)dt$. Torna-se óbvio que ao problema resultante não podemos aplicar os típicos resultados sem condições de convexidade, devido às estruturas "cruzadas" $Q(x)\Phi(u)$ e $c(x)\Phi(u)$. Por outro lado, e apesar de teoricamente os resultados do tipo de Filippov poderem em princípio ser aplicados nesta situação, a dependência não linear no controlo, juntamente com o elevado número de equações intervenientes, torna praticamente impossível uma caracterização adequada do conjunto $\mathcal{A}_x$, pelo que novas abordagens são necessárias.

A nossa contribuição principal tenta precisamente encontrar uma relação entre os ingredientes acima descritos, c, c_0, Q, Q_0, Φ e o conjunto de admissibilidade K, de forma a garantir a existência de solução para o problema de controlo óptimo resultante. Definimos como controlo admissível uma função mensurável ao qual corresponde um único estado admissível descrito por uma função absolutamente continua.

Dada a natureza não convexa do problema, procedemos a uma relaxação típica utilizando medidas parametrizadas, numa formulação actual, tal como descrita nos trabalhos de Pedregal e de Roubicek, entre outros. Aproveitando a sua estrutura particular, o problema relaxado pode ser escrito como um problema de controlo óptimo, cuja variável de controlo corresponde para cada instante t, a um vector de momentos generalizados, associados a uma medida de probabilidade. A questão da existência resume-se então a saber se o controlo óptimo para este problema de momentos está associado, a cada instante, a uma medida de probabilidade do tipo Delta, com suporte em K. Dois conjuntos jogam assim um papel fundamental. O conjunto de momentos generalizados correspondentes a medidas de probabilidade com suporte em K,

$$\Lambda = \{\int_{\mathbf{R}^n} \Phi(\lambda)d\mu(\lambda),\ \mu \in P(K)\},$$

e o seu subconjunto constituido pelos momentos correspondentes a medidas do tipo Delta

$$L = \{\int_{\mathbf{R}^n} \Phi(\lambda)d\delta_u(\lambda),\ u \in K\} = \Phi(K).$$

Como K é um conjunto compacto, Λ e L verificam a propriedade, muito conveniente,

$$\Lambda = co(L) = \bar{co}(L).$$

De facto, L coincidirá com os pontos extremos de Λ desde que assumamos Φ convexa componente a componente.

Após uma análise detalhada destes conjuntos para alguns casos de dimensão um, podemos chegar a um resultado geral, o qual pode descrever-se formalmente do seguinte modo: assumindo que estão garantidas as condições para que o sistema de controlo $x' = Q(x)\Phi(u) + Q_0(x)$ tenha solução única, para cada u admissível, por exemplo, admitindo que Q e Q_0 verificam a condição de Lipschitz. Supondo além disso que $\Phi(K)$ está contido numa superfície convexa, de modo a que o seu invólucro convexo Λ esteja sempre "do mesmo lado" de qualquer plano tangente a L. Então, se para quaisquer x e ξ, os vectores v que transladados a um ponto de L verifiquem simultaneamente $\xi = Q(x)v$ e $c(x) \cdot v \leq 0$, estão orientados para o exterior de Λ, o problema tem solução óptima.

Para descrever esta ideia geométrica necessitamos de uma caracterização de L e Λ que sirva para qualquer dimensão. Para isso supomos que existe uma aplicação Ψ de classe $\mathcal{C}^1$, convexa componente a componente, tal que L esteja contido na superfície de nível dada por $\{\Psi = 0\}$ e Λ na região do espaço definida por $\{\Psi \leq 0\}$. Com esta notação, dizer que v está orientado para o exterior de Λ significa que v pertence ao conjunto

$$\mathcal{N}(\Phi, K) = \{v : \ \nabla\Psi(m)v = 0, \text{ or } \exists i : \ \nabla\psi_i(m)v > 0, \ m \in L\}.$$

A existência de solução fica assim sujeita á verificação da seguinte condição: Se para quaisquer $x \in \mathbf{R}^N$ o conjunto

$$\mathcal{N}(c(x), Q(x)) = \{v : \ 0 = Q(x)v, \ c(x) \cdot v \leq 0\}$$

está contido em $\mathcal{N}(\Phi, K)$, então existe solução para o problema de controlo óptimo. Apesar de nos casos de dimensão um conseguirmos condições mais exaustivas, esta relação entre os ingredientes principais do problema dá-nos uma condição algébrica explicita, que poderá ser verificada nos casos de dimensões superiores, onde a condição suficiente de Filippov seja de difícil validação.

Uma outra abordagem, destinada a provar a existência de solução para problemas de controlo óptimo, sem as típicas condições de convexidade, passa por uma reformulação variacional como foi proposta por Rockafellar, segundo a qual definimos uma função

$$\varphi(x, \xi) = \min_u \{F(x, u) : \ \xi = f(x, u), \ u \in K\}$$

e consideramos o problema do Cálculo das Variações com custo $\int_0^T \varphi(x(t), x'(t))dt$ cuja minimização terá em conta apenas o estado x. Sobre certas hipóteses este problema variacional é equivalente ao problema de controlo óptimo que lhe deu origem. Esta técnica, apesar de clássica, permite provar resultados de existência no Controlo Óptimo, com base na aplicação dos resultados de existência disponíveis para o Cálculo das Variações. Na literatura existente

sobre o assunto, os resultados obtidos são essencialmente equivalentes aos demonstrados por Filippov, Neustadt ou Cesari. No entanto, um resultado recente para problemas variacionais, devido a Ornelas, no qual a convexidade é substituida por uma condição mais fraca, permite obter resultados no caso escalar, para uma estrutura geral $F(x,u)$, $f(x,u)$ sem ter de recorrer às condições de convexidade clássicas. Este resultado ilustra o poder da reformulação variacional, e indica que no futuro, mais resultados de existência poderão ser conseguidos em função dos novos avances na teoria do Calculo das Variações.

Os resultados de existência obtidos, concretamente o resultado principal com base na relaxação, permite abordar o problema da manobra de um veículo subaquático com mais detalhe. Após uma análise cuidada da matriz Q, e dos vectores c e Q_0, podemos concluir que existe um intervalo de tempo $[0,T]$ para o qual o problema está bem posto e portanto faz sentido estudar a existência de solução óptima. A verificação da condição $\mathcal{N}(c(x),Q(x)) \subset \mathcal{N}(\Phi,K)$ torna-se, neste caso, num problema de álgebra linear, que pode ser facilmente resolvido, permitindo assim concluir sobre a existência de solução óptima.

Entre as possibilidades de trabalho futuro apresentadas, destacamos a aproximação numérica para problemas de controlo óptimo não lineares, com variáveis de estado e controlo vectoriais, como o modelo acima descrito, proveniente do controlo de veículos subaquáticos. Propomos a implementação de um método baseado nas direcções de descida mais rápida. Mostramos alguns resultados preliminares para exemplos académicos. O melhoramento e a estabilização deste método poderão constituir os próximos passos a realizar no futuro.

O segundo capítulo desta dissertação correponde ao trabalho [P. Pedregal and J. Tiago, *Existence results for optimal control problems with some special non-linear dependence on state and control, SIAM J. Control Opt., 48 n.2 (2009) 415-437*].

O terceiro capítulo, por sua vez, corresponde ao trabalho [P. Pedregal and J. Tiago, *A new existence result for autonomous non convex one-dimension optimal control problems*, J. Optimiz. Theory App., 134 (2007), pp. 241-255] cujo tema foi estudado inicialmente durante a tese de mestrado e posteriormente aprofundado durante o período de doutoramento.

O quarto capítulo corresponde ao trabalho, em processo de revisão, [F. Periago and J. Tiago, *A local existence result for an optimal control problem modeling the manoeuvering of an underwater vehicle*, Nonlinear Analysis: Real World Applications, Accepted. Available online from September 2009. DOI: doi:10.1016/j.nonrwa.2009.09.002].

Palavras chave. Existência, não convexo, não linear, relaxação, reformulação variacional, veículos subaquáticos

Acknowledgments

I would like to express me gratitude to Prof. Pablo Pedregal for his patience and dedication during all these years. With him, not only mathematics can be very motivating, as it is easy to be inspired by his positive and professional attitude.

A special thank also to Prof. Cristian Barbarosie, for supporting me from the very beginning, for his opportune advices and full dedication in helping me making this project a reality.

I wish to acknowledge the Centro de Matemática e Aplicações Fundamentais, for the financial support through the CMAF Research Grant 2005/2006, before being awarded with the FCT's doctoral grant. Moreover, I am grateful to the financial support of the group OMEVA through the Departments of Mathematics of the Universidade de Castilla-La Mancha and Universidad Politécnica de Cartagena for the expenses relative to the conferences and meetings I attended.

I also want to acknowledge Prof. Jean Pierre Raymond for receiving me at Université Paul Sabatier, allowing me to introduce new subjects within a different and interesting scientific environment.

Besides, from the collaboration with Prof. Francisco Periago resulted some work integrating this thesis. I want to thank him for his interest, and for his ability to awake the enthusiasm for the applications of mathematics.

As to the people from the Department of Mathematics in Ciudad Real, my gratitude for the warm reception and for the nice mathematical and professional talks along these years. A special thank to Adrian, for his patient help in computer issues.

I also wish to thank to those friends world wide, specially to Sonia, for making life so interesting.

I cannot end without expressing my deepest debt to my parents and my family, whose joy of life, and hardworking example, have inspired me in the most difficult moments.

This research was supported by Fundação para a Ciência e Tecnologia through doctoral degree grant SFRH-BD-22107-2005 in the context of the Programa Operacional Ciência e Inovação 2010.

Contents

1 Introduction **1**

1.1 Problem Description 1

1.2 Relevance and Motivation: application to an underwater vehicle problem 3

1.3 Context and Difficulties 6

1.4 Contributions 10

1.4.1 Via Relaxation 11

1.4.2 Via Variational Reformulation 16

1.4.3 Aplication to an Underwater Vehicle Problem 18

1.5 Future Work 21

1.5.1 Numerical Approximations 21

1.5.2 Other Existence Results 25

1.5.3 Other Applications 25

1.6 Final Comments 25

2 Main Existence Result **29**

2.1 Introduction 29

2.2 Proof of Theorem 2.3 35

2.3 The set Λ and duality 36

2.4 Polynomial Dependence. The case $N = n = 1$, $p = 2$ 38

2.5 The case $N = n = 1$, $p = 3$ 40

2.6 A geometric approach to the case $N = n = 1$, $p = 3$. 42

2.7 Examples 47

2.7.1 Example 1 47

2.7.2 Example 2 49

2.7.3 Example 3 50

2.8 The case N, $n > 1$ 52

3 Some results in scalar problems **55**

3.1 Introduction 55

3.2 Variational Reformulation 58

3.3 An Existence Theorem In Dimension One 60

3.4 Examples 66

4 Application to Underwater Vehicles Models **71**

4.1 Introduction 71

4.2 A general existence and uniqueness result for some specific optimal control problems 73

4.3 Proof of Theorem 4.1 79

4.3.1 Step 1: the matrices Q and Q_0 79

4.3.2 Step 2: local existence and uniqueness of solutions for the state law . . . 83

4.3.3 Step 3: checking condition (4.11) in Theorem 4.2 85

A Some Classical Theory **89**

A.1 Measurable Selection 89

A.2 Parametrized Measures 89

A.3 Control Theory 91

B Future Directions **97**

B.1 On the numerical approximation for optimal control problems via a steepest descent method 97

B.1.1 Introduction 97

B.1.2 The general strategy: the descent method 102

B.1.3 The iteration step: the linear case 104

B.1.4 A global implementation 108

B.1.5 Numerical experiments 112

B.1.6 Conclusions 134

Bibliography **137**

Chapter 1

Introduction

1.1 Problem Description

The main purpose of this thesis is to present some results which intend to extend the existence theory for optimal control problems governed by systems of ordinary differential equations. Such situations describe the way a certain control function u can generate an output x, a state trajectory, attaining some criteria, usually in terms of a cost functional, while verifying certain viability constraints for both state and control.

This type of problems have been approached from many different perspectives concerning the generality of the problem. Among those, proving existence results according with the regularity of the main ingredients, by one side, and the generality of the viability constraints, by another, are two of the most common perspectives. Here we focus on the nonlinearities presented in the main ingredients while depending on the control variable. For this reason, we concentrate our attention over a cost functional of the Lagrange type

$$I(x,u) = \int_0^T F(x(t),u(t))dt, \tag{1.1}$$

and a so called control system

$$x'(t) = f(x(t),u(t)), \tag{1.2}$$

where $(0,T)$ is a fixed time interval, and both F and f don't depend explicitly on the time variable, letting us to better concentrate on the nonlinearity issue. Incorporating the time dependence has been done in many results, and it can possibly be done for the situations we treat here but that would require more attention, and should be the subject of future work.

In fact, our main contribution, in a vectorial frame, refers to the special case where

$$F(x,u) = c(x)\cdot\Phi(u) + c_0(x) \tag{1.3}$$

and

$$f(x,u) = Q(x) \cdot \Phi(u) + Q_0(x) \tag{1.4}$$

with c, c_0, Q, Q_0 and Φ possibly non convex, nonlinear, but continuous functions.

Also we consider only simple, but typical, constraints on the control, namely that

$$u(t) \in K \in \mathbf{R}^n \tag{1.5}$$

where K is a convex compact set, letting more general viability constraints like $u(t) \in U(t, x(t))$ to be treated elsewhere. Although in the available theory, the state variable x can be asked to verify many different type of contraints, like $(x(0), x(T)) \in B$ where B is a certain compact set, we restrict ourselves to the initial condition

$$x(0) = x_0 \tag{1.6}$$

which, when considered together with the control system (1.2), makes it an initial value problem, for which much useful theory is available. In particular, asking f to verify the (global) Lipschitz condition over $(0,T)$ with respect to the state variable, will allow us to assume that the optimal control problem is well posed, whenever we search for absolutely continuous states and measurable controls.

Summarizing, our proposed contribution turns over an optimal control problem of the type

$$(P_1) \qquad \text{Minimize in } u: \quad \int_0^T [c(x(t))\Phi(u(t)) + c_0(x(t))]\, dt \tag{1.7}$$

subject to

$$x'(t) = Q(x(t))\Phi(u(t)) + Q_0(x(t)) \text{ in } (0,T), \tag{1.8}$$

$$x(0) = x_0 \in \mathbf{R}^N \tag{1.9}$$

and

$$u(t) \in K \subset \mathbf{R}^n, \tag{1.10}$$

where K is a convex compact set.

The optimal solution should be a pair of functions (x,u) such that

$$u \in L^\infty((0,T), \mathbf{R}^n) \text{ and } x \in AC((0,T), \mathbf{R}^N). \tag{1.11}$$

where AC stands for to the space of absolutely continuous functions.

We also assume that the mappings

$$c : \mathbf{R}^N \to \mathbf{R}^s, \quad c_0 : \mathbf{R}^N \to \mathbf{R},$$

are continuous, the mappings

$$Q : \mathbf{R}^N \to \mathbf{R}^{N\times n}, \quad Q_0 : \mathbf{R}^N \to \mathbf{R}^N$$

are Lipschitz continuous and the mapping

$$\Phi : \mathbf{R}^n \to \mathbf{R}^s,$$

describing the nonlinear dependence on the control, will be assumed to be of class $\mathcal{C}^1$.

The extra regularity asked to Φ is due to the fact that we will use this mapping to describe some geometrical properties associated with the optimal control problem. Under these "strong" regularity assumptions, together with the compactness and the convexity of K, the problem is well posed in the sense that for every admissible control u there will be a (unique) state x solving the initial value problem associated, and such that the cost function is finite over the solution pair, that is,

$$-\infty < I(x,u) < +\infty. \tag{1.12}$$

Although our attention will turn mainly over problem (P_1), we will not loose sight of the more general problem (P)

$$\min_{(x,u)} \int_0^T F(x(t),u(t))$$

verifying the constraints

$$x'(t) = f(x(t),u(t)), \quad x(0) = x_0,$$

$$u(t) \in K \ \text{ a.e. } t \in (0,T).$$

Concerning this problem, we will propose some minor contributions and some directions for future work.

1.2 Relevance and Motivation: application to an underwater vehicle problem

Even if, within the general theory of optimal control, problem (P_1) seems to be quite special, important real world situations can be modelled mathematically as optimal control problems of this type. We have analyzed the particular situation of modelling the control and maneuvering of an underwater vehicle. Several mathematical models have been proposed in the literature (see for instance [Fs94]), but we considered the model proposed in [GOP09]. A twelve variable state

$$X = (V_{body}, X_{world})$$

where

$$V_{body} = (\mathtt{u}, \mathtt{v}, \mathtt{w}, \mathtt{p}, \mathtt{q}, \mathtt{r})$$

$$X_{world} = (\mathtt{x}, \mathtt{y}, \mathtt{z}, \phi, \theta, \psi)$$

is proposed. As it is described in Figure 1.1, the variables $(\mathtt{u}, \mathtt{v}, \mathtt{w})$ represent the linear velocities

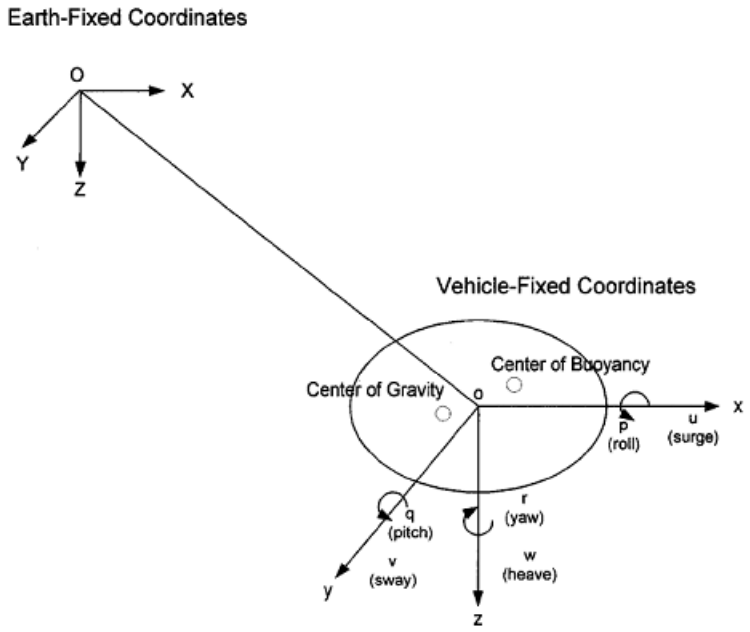


Figure 1.1: Variables in earth-fixed and vehicle coordinate system

in the surge, sway and heave directions with respect to the vehicle or body-fixed coordinate system, while $(\mathtt{p}, \mathtt{q}, \mathtt{r})$ refer to the angular velocity around the surge axe (roll movement), sway axe (pitch movement), and heave axe (yaw movement).

The variables $(\mathtt{x}, \mathtt{y}, \mathtt{z})$ give the position with respect to the earth or world-fixed coordinate system, and (ϕ, θ, ψ) the orientation accordingly with the previous axes.

The relation between both coordinate systems is given by the six ordinary differential equations

$$\dot{X}_{world} = T(\phi, \theta, \psi) V_{body}$$

where T is a transformation matrix.

The guidance of the vehicle is controlled by the three variables

$$(u_1, u_2, u_3) = (\delta_b, \delta_s, \delta_r)$$

which appear in linear and square form in the dynamics of the system

$$\frac{d}{dt} V_{world} = M^{-1} F$$

where

$$M = \begin{pmatrix} \text{mass} \\ \text{inertia and} \\ \text{added-mass} \end{pmatrix} \text{(showed to be constant)}$$

and

$$F = \begin{pmatrix} \text{velocity-related} \\ \text{inertial and} \\ \text{hydrodynamics} \\ \text{force functions} \\ \text{(drag and lift)} \end{pmatrix} + \begin{pmatrix} \text{buoyancy,} \\ \text{weight,} \\ \text{propulsion} \\ \text{and other} \\ \text{forces} \end{pmatrix}$$

Hence, a three variable control function u will be asked to be input in a twelve equation system of ordinary differential equations in order to get a twelve variable state output x. The objective will be to find how to control the system in order to lead the underwater vehicle from a given state x_0 to a final state x_T in a given time T where the control variables must remain in a compact convex cube of $\mathbf{R}^3$. A typical penalizing term should be considered to help chosing the "less expensive" control function. Therefore the most immediate cost function associated with such objective would be a Bolza cost function of the type

$$d(x(T), x_T) + \int_0^T P(u(t))dt \tag{1.13}$$

where d refers to an appropriate distance function, and P to a penalizing term. The control variables act in the control system in a linear and quadratic form, so that we can describe the control contribution in the system of twelve equations with the mapping

$$\Phi(u) = \Big(u_1, u_2, u_3, (u_1)^2, (u_2)^2, (u_3)^2\Big) \tag{1.14}$$

and rewrite the control system as

$$x'(t) = Q(x)\Phi(u) + Q_0(x) \tag{1.15}$$

If the integrand P in the cost function (1.13) is chosen to be $P(u) = \|u\|^2$, it can be rewritten (1.13) as a Lagrange cost function of the type

$$\int_0^T c(x)\Phi(u) + c_0(x)dt. \tag{1.16}$$

The functions c, c_0, Q, Q_0 and Φ will also have enough regularity to allow us to be in the exact settings of the optimal control problem (P_1).

1.3 Context and Difficulties

The answer to the question, if classical results can be applied to problem (P_1), is yes. At least theoretically. The Filippov-Roxin theory ([F62],[R62]) is sufficiently general to include this class of problems. However, the convexity requirement on the so called orientor field, the set

$$\mathcal{A}_x = \{(v,z) : \ v > F(x,u),\ z = f(x,u),\ u \in K\} \tag{1.17}$$

can be extremely difficult to check, if not impossible. This difficulty applies either in this situation, or in many others, depending on the dimensions of the problem, and the nonlinear dependence on the control. Actually, even in the scalar case, the theory cannot be applied to the problem

$$\min \int_0^1 (u(t)^2 - 1)^2 - x(t)^2 dt \tag{1.18}$$
$$x(t) = u(t) \quad x(0) = x_0 \quad u(t) \in [-1,1] \text{ a. e. } t \in (0,1)$$

for which the existence of solution can be established, even though the sets $\mathcal{A}_x$ are not convex.

In more recent results ([MP01],[Mk99]), based on ideas already presented in [GMk77], [Mk77], this condition is weakened, asking optimal trajectories to be out of an certain set associated with the lack of convexity, but the essential difficulty remains the same: characterize the set $\mathcal{A}_x$ for every x, in an explicit manner, when dealing with nonlinear vectorial problems.

On the other hand, when cross terms of the type $Q(x)\Phi(u)$, are present, we cannot apply results for non-necessarily convex problems like the ones in [N63], [O70], [Ma80] and [Ce74], neither the more recent ones [Ra90] or [B94], [RbS97], both dealing with general class of problems including integrand functions like the one in (1.18).

Concerning the former, generalized solutions, in the sense of Gamkrelidze ([Ga62]), are searched for. The special structure

$$F(x,u) = G(x) + H(u), \quad f(x,u) = g(x) + h(u),$$

together with the linearity of both G and g, are taken advantage of in order to recover solutions for the original problem by using Lyapunov's theorem (see [Ce83b]).

With respect to the result in [Ra90], in the spirit of [CC90] (more suitable for variational problems), the previous techniques are explored to what seems to be their limit. Authors use, respectively, a lower semicontinuous regularization (as in [Bu87]), and relaxation via the bipolar function ([ET74b]), as preferred alternatives to get generalized solutions. Even so, the concavity of G and g appears to be the weaker condition allowed within the framework based on the use of Lyapunov's theorem, to recover the desired solutions from the generalized ones (see [Ra94] for more comments on this).

As for [B94], and also [RbS97], where differential systems are seen as particular cases of integral equations, a different approach is proposed. A relaxation of the original problem based on parametrized measures (see [P97b], [Rb97b] or [CRV04]), also called Young measures (due to previous work in [Y37]), is used. This time, relaxation is not seen just as a way of obtaining a lower semicontinous functional, but is also used in order to profit some good properties of the relaxed functional

$$\bar{I}(\mu) = \int_0^T \int_{\mathbf{R}^m} F(x(t), \lambda) d\mu_t(\lambda) dt,$$

where x is the solution of

$$x'(t) = \int_{\mathbf{R}^m} f(x(t), \lambda) d\mu_t(\lambda), \quad x(0) = x_0,$$

and the parametrized measure $\mu = \{\mu_t\}_t$ belongs to

$$\{\mu \in L^\infty_{w*}((0,T), \mathcal{M}(\mathbf{R}^m)) : \ \mu_t \in P(K) \text{ a.e. } \ t \in (0,T)\}.$$

Here $\mathcal{M}(\mathbf{R}^m)$ stands for the set of Radon measures, while $P(K)$ refers to the set of probability measures with support on $K \in \mathbf{R}^m$. As to $L^\infty_{w*}((0,T), \mathcal{M}(\mathbf{R}^m))$, the set of weakly measurable and essentially bounded mappings $t \to \mu_t$, it can be identified with the dual space of $L^1((0,T), C_0(\mathbf{R}^m))$, so that in practice,

$$\mu = \{\mu_t\}_t \in L^\infty_{w*}((0,T), \mathcal{M}(K)) \tag{1.19}$$

means that the mapping

$$t \to \int_{\mathbf{R}^m} F(\lambda) d\mu_t(\lambda)$$

is Lebesgue measurable for all $F \in C_0(\mathbf{R}^m)$. Notice that the relaxation through Young measures was also used in [MP01] already commented above. In fact, there is a large tradition in using parametrized measures in minimization problems where lack of convexity gives rise to highly oscillating phenomena. Besides the ones mentioned above, [W62] (for control problems), [Ta79], [BK90], and [KP94] are some reference papers on this subject. The Bolza example

$$\min \int_0^1 (u(t)^2 - 1)^2 + x(t)^2 dt \tag{1.20}$$

$$x(t) = u(t) \quad x(0) = x_0 \ (x(1) = x_T) \quad u(t) \in [-1, 1] \text{ a. e. } t \in (0,1)$$

(see [IT74]) is a simple scalar problem where such oscillating behavior for minimizing sequences can be found.

In [B94], the concavity of $\bar{I}$ and the convexity and compacity of the admissible set of parametrized measures are searched for in order to allow the application of Bauer' Maximum Principle ([Ba58]), and to conclude that there is a solution to the relaxed problem which is an

extreme point of the admissible set. It turns out that such point can be characterized as delta type measures

$$\mu = \{\delta_{u(t)}\}_t$$

corresponding to optimal solutions $u(.)$ for the original problem. We remark that Balder still uses Lyapunov's theorem, but he deals with a generalized solution of the type

$$\mu = \alpha\delta_{u_1(t)} + (1-\alpha)\delta_{u_2(t)}$$

much more easy to handle than the ones in previous works, where the coefficients α were still time dependent.

Even if Balder could generalize previous results in many senses, unfortunately, in what concerns problem (P_1), the results in [B94] (or [RbS97]) cannot be applied as they still ask for the functions, in both cost and dynamics, to have a structure based on a separate dependence for the state and the control in two different terms, plus the concavity of the cost integrand function with respect to the state variable. Actually, during a formal attempt to try to extend these ideas to more general cost functions, we found very difficult to check the concavity of a functional coming from a structure different from the one assumed in Balder's work . This doesn't mean that such approach cannot provide further results.

Another class of results we should have in mind are the ones coming from the techniques based on the Rockafellar's variational reformulation introduced in [Rk75], and well-described in [Ce83b], [ET74b] or recently in [P03]. According to these, a new integrand φ, some times called density function, should be defined, incorporating the different constraints from the control problem as follows

$$\varphi(x,\xi) = \begin{cases} \inf\{F(x,u) :\ \xi = f(x,u),\ u \in K\} & \text{if } \xi \in f(x,K) \\ +\infty & \text{else.} \end{cases} \tag{1.21}$$

Sufficient conditions should be looked for in order to guarantee that the variational problem

$$(VP) \quad \min I(x) = \int_0^T \varphi(x(t),u(t))dt$$

$$x(0) = x_0, \quad x \in AC((0,T),\mathbf{R}^N)$$

is equivalent to the original optimal control problem. The existence of solution for the control problem is then a question of applying an existence result for the associated equivalent variational problem. As we can understand from the complete exposition in [Ce83b], if we stick to classical existence theory, these techniques don't allow us to go much further than the previous ones. Actually, if we want to apply the classical Tonelli's existence theorem for calculus of

variations [To14] to problem (VP), the convexity with respect to the state derivative in φ turns out to be equivalent to the convexity of the orientor field $\mathcal{A}_x$ (1.17) in Filippov-Roxin's theory. If we admit the lack of convexity, and we want to apply appropriate traditional results in calculus of variations without convexity, then, again, we will need a structure based on a separate dependence on the state variable and the control variable (see [Ce83b]). We felt this difficulty while trying, without any success, to apply the Cellina-Colombo existence result [CC90], to the variational problem (VP) coming from an optimal control problem as general as (P) where cross dependence with respect to the state and control variables is probable to be true.

A possible way out of this difficulty, could be choosing more recent existence results for calculus of variations where the separate dependence is not an issue.

One of such result, included in [Ra94], follows the spirit of some classical results for the calculus of variations (see [AT79]), where again, the integrand was assumed to be of the type

$$\varphi(x) = G(x) + H(x').$$

In this direction, the author choose to use a monotony condition to prove the existence of solution, this time for a general integrand φ, without the need of the typycal separate structure, neither any type of concavity. The typical Lyapunov's theorem is not used, and instead, solutions are seen to be out of a certain set associated with the lack of convexity with respect to the state derivative. Such monotony condition is characterized in terms of partial derivatives of the bipolar function φ^{**} given in the sense of [ET74b]. Such characterization demands some regularity on φ, which is definitely not verified, in general, by the density function

$$\varphi : \mathbf{R}^N \times \mathbf{R}^N \to \mathbf{R} \cup \{+\infty\}$$

coming from the variational reformulation. While such regularity can't be weakened, it seams difficult to apply this class of results to go further in existence theory for optimal control problems.

More recently, in [Or07], an existence result was proved for the calculus of variations in the one dimensional, scalar case, where $N = 1$, but for a general lower semicontinuous autonomous integrand φ verifying a special condition, later called zero-convexity. Such condition states that

$$\varphi^{**}(x, 0) = \varphi(x, 0), \ \forall x,$$

where, as above $\varphi**$ refers to the bipolar function of φ with respect to the state derivative,

$$\varphi^{**}(x, \xi) = \sup \{A(\xi) : A(.) \text{ is an affine function}, \ A(v) \leq \varphi(x, v), \ \forall v \in \mathbb{R}\}$$

with x fixed as a parameter. It turns out that such special convexity condition can be translated directly into the ingredients of scalar optimal control problems. Doing this is one of our proposed contributions, and it will be the subject of the forthcoming discussion in next section.

As to one-dimensional vectorial problems, a recent result ([CO07]), guarantees the existence of solution under similar regularity conditions, but seeing the zero-convexity as a property for almost convex functions. Such concept, which was introduced in [CO03], describes the behavior of the integrand φ in the regions of lack of convexity with respect to the state derivative, that is, where the bipolar function is strictly smaller than the image itself. Hence, we say that φ is almost convex if considering ξ such that

$$\varphi^{**}(x,\xi) < \varphi(x,\xi)$$

where x is fixed as parameter, there are $\lambda, \alpha \in [0,1]$ and $\beta \in [1,+\infty)$ such that

$$\xi = \alpha(\lambda\xi) + (1-\alpha)(\beta\xi)$$

and

$$\varphi^{**}(x,\xi) = \alpha\varphi(x,\lambda\xi) + (1-\alpha)\varphi(x,\beta\xi).$$

In particular, this property implies the zero-convexity in [Or07], which is seen not to be sufficient to prove the existence in the vectorial case. The main technique is based on bimonotonicity, a property introduced in [Or98] (after preliminary ideas in [FMO98]), which turns up to be present in the minimizers after certain transformation.

Checking how almost convexity can be translated to optimal control problems, could be a possible direction to follow in future work.

A briefly comment can be done about the possibility of using the known equivalence between Lagrange optimal control problems like (P) or (P_1), and other optimal control problems like Mayer and Minimum Time problems. Concerning the former, as we can see in [Ce83b], the equivalence with Lagrange problems can be established in such a way that the available results apply for both problems. As to the later, the equivalence is not as general as for the Mayer problems, and although recent existence results ([CFM06]) are available for this problem, using it for Lagrange problems does not seem to be quite appropriate, as we will illustrate later.

1.4 Contributions

As mentioned above, even if some of the classical results can be, in theory, applied to our optimal control, in practice, some of the requested conditions are extremely difficult to check,

especially for some important problems coming from real world situations. So that giving alternative existence results can be quite useful to a better understanding of such situations.

While stating our problem, we have seen that the regularity of the main ingredients, and the generality of the constraints are not an issue. What is our main purpose then?

We would like to answer the following general questions:

(I) How does the nonlinear dependence on the control variable relates with the existence of solution?

(II) In what way should the integrand F, the vector field f and the viability set K relate to each other, in order to guarantee the existence of an optimal solution?

To help answering these questions, we would like to give some alternative conditions to the convexity of $\mathcal{A}_x$ in (1.17), which can fit better in situations where vectorial problems involving highly nonlinear dependences on the control are analyzed.

A more specific question, directed to problem (P_1) can be:

(III) Given an admissible set K, and a particular nonlinear dependence described by $\Phi(u)$, what class of mappings Q, Q_0, c and c_0 will allow problem (P_1) to have an optimal solution?

Giving a checkable, sufficient condition characterizing such admissible ingredients could help in answering this question.

1.4.1 Via Relaxation

We believe we can help answering question (III) in the vectorial case when we are in presence of problem (P_1). To be more precise, let us first introduce some notation. Consider a new ingredient of the problem related to Φ. Suppose that there is a $\mathcal{C}^1$ mapping

$$\Psi : \mathbf{R}^s \to \mathbf{R}^{s-n}, \quad \Psi = (\psi_1, ..., \psi_{s-n}), \quad (s > n), \tag{1.22}$$

so that $\Phi(K) \subset \{\Psi = 0\}$. This is simply saying, in a rough way, that the embedded (parametrized) manifold $\Phi(K)$ of $\mathbf{R}^s$ is part of the manifold defined implicitly by $\Psi = 0$. For instance, if we consider

$$\Phi(u) = \Big(u_1, u_2, u_3, (u_1)^2, (u_2)^2, (u_3)^2\Big)$$

as in (1.14) then $\Psi : \mathbf{R}^6 \to \mathbf{R}^3$ can be given by

$$\Psi(v) = ((v_1)^2 - v_4, (v_2)^2 - v_5, (v_3)^2 - v_6) \in \mathbf{R}^3.$$

Also, for a pair (c, Q), put

$$\mathcal{N}(c, Q) = \{v \in \mathbf{R}^s : Qv = 0, cv \leq 0\}. \tag{1.23}$$

Similarly, set

$$\mathcal{N}(K, \Phi) = \tag{1.24}$$

$$\{v \in \mathbf{R}^s : \text{ for each } u \in K, \text{ either } \nabla\Psi(\phi(u))v = 0, \text{ or there is } i \text{ with } \nabla\psi_i(\phi(u))v > 0\}.$$

Under this setting we will see that if

$$\mathcal{N}(c(x), Q(x)) \subset \mathcal{N}(K, \Phi), \tag{1.25}$$

then problem (P_1) admits at least one solution. Under an additional condition, we can also establish that this solution is in fact, unique.

Notice that condition (1.25), gives us a sufficient relation between c, Q, Φ and K to ensure the existence of solution. It does not include c_0 nor Q_0, whenever these functions are such that the problem is well posed in the sense described in Section 1.1.

A geometrical interpretation is behind this vector sets. Suppose, additionally, that Ψ is componentwise convex. Hence, if we consider the sets

$$L = \Phi(K), \quad \Lambda = co(L), \tag{1.26}$$

since Ψ is convex, we can see that Λ can be identified with part of the set $\{\Psi \leq 0\}$. Thus, a sufficient condition for a vector placed in a point $\Phi(u)$ of $L \subset \{\Psi = 0\}$ to be oriented to the "exterior" of Λ is that it belongs to the set $\mathcal{N}(K, \Phi)$, that is, v cannot point into the interior of the "$(s-n)$-hyperoctant" defined at point $\Phi(u)$ by $\{v : \nabla\Psi(\Phi(u)) \cdot v \leq 0\}$, where the set Λ is contained. This condition must be verified for all $u \in K$. In fact we could add some assumptions to ensure that $\mathcal{N}(K, \Phi)$ will never be empty, like asking Φ to have at least one linear component, which will be the case in every situation we are dealing with. But we prefer to assume $\mathcal{N}(K, \Phi)$ to be non empty, for the sake of more generality.

Therefore, condition (1.25) means that, for all x, the directions in $\mathcal{N}(c(x), Q(x))$, those simultaneously belonging to the hyperspace $\{v : Q(x)v = 0\}$ and the half space $\{v : c(x) \cdot v \leq 0\}$, point outside Λ whenever placed on any point of L.

In Figure 1.2, we illustrate the ideas behind condition (1.25), for the simple case were

$$N = n = 1, \ s = 2, \ K = [a_1, a_2],$$

$$\Phi(u) = (u, u^2) \text{ and } \Psi(m) = (m_1)^2 - m_2.$$

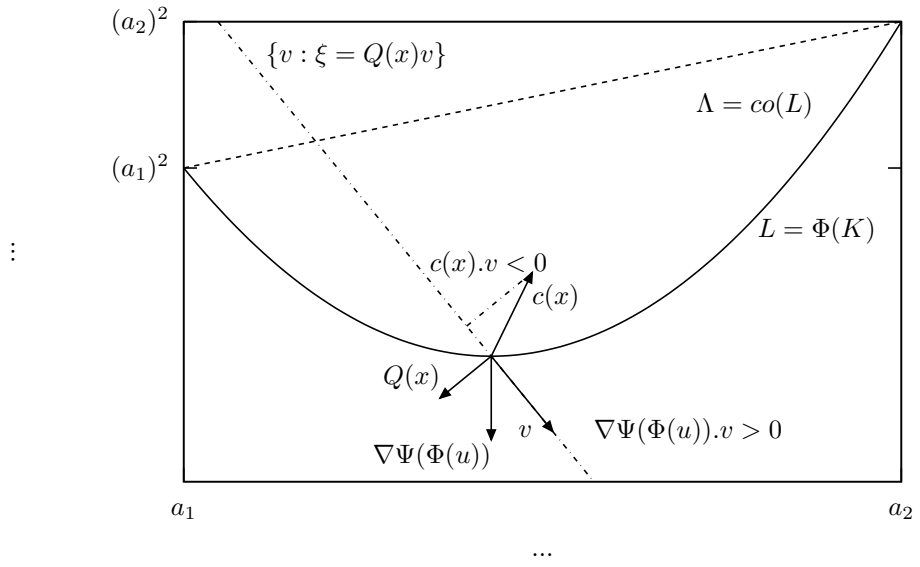


Figure 1.2: $v \in \mathcal{N}(c(x), Q(x)$ verifies $\nabla\Psi(\Phi(u)) \cdot v > 0$

There, v represents the vectors in $\mathcal{N}(c(x), Q(x))$ and verifies $\nabla\Psi(\Phi(u)) \cdot v > 0$ for the selected $\Phi(u) \in L$.

This condition is sufficient to ensure the following assumption.

Hypothesis 1.1. *For each fixed $x \in \mathbf{R}^N$, and $\xi \in Q(x)\Lambda \subset \mathbf{R}^N$, the minimum*

$$\min_{m \in \Lambda} \{c(x) \cdot m : \xi = Q(x)m\}$$

is only attained in L, where $L = \Phi(K)$, and $\Lambda = co(L)$.

Actually, this hypothesis is crucial to prove our main contribution.

Theorem 1.1. *Consider problem (P_1). Assume that the mapping Ψ as above is componentwise convex and $\mathcal{C}^1$. If for each $x \in \mathbf{R}^n$, the functions Q and c are such that condition (1.25) is verified, then the corresponding optimal control problem (P) has at least one solution.*

In fact, after proceeding to a relaxation of problem (P_1), with respect to the control variable, based on parametrized measures lying in

$$\{\mu \in L^\infty_{w*}((0,T), \mathcal{M}(\mathbf{R}^m)) : \ \mu_t \in P(K) \text{ a.e. } \ t \in (0,T)\},$$

and proving that the relaxed problem have an optimal solution, Hypothesis 1.1 will allow us to conclude that indeed this optimal solution, is a Dirac type measure, corresponding to a an optimal solution for problem (P_1).

It may be useful to understand where the sets Λ and L come from. Due to the particular structure with Φ describing the exact dependence on the control variable, we can rewrite the

relaxed problem in terms of generalized vector moments of the type

$$m = \int_K \Phi(\Lambda) d\mu(\lambda) \in R^s$$

where μ is a probability measure with support on K. These techniques are closely related to the classical moment problem ([Ak61], [ST70] or more recently [EMeP03], [Me04]). In the resulting problem, the set

$$\Lambda = \{m \in R^s : m = \int_K \Phi(\lambda) d\mu(\lambda) \ \mu \in P(K)\}$$

plays an important role. In particular the set $L \subset \Lambda$ corresponding to delta type probability measures, and therefore defined as

$$L = \{m \in \mathbf{R}^s : m = \int_K \Phi(\lambda) d\delta_u(\lambda), \ u \in K\}$$

$$= \{m \in \mathbf{R}^s : m = \Phi(u), \ u \in K\}$$

will also be very important, as we should prove that the solutions of the relaxed problem should lie in this kind of set. Convexity and closeness of Λ are important properties ([Me04]). In our framework this doesn't suppose any special difficulty because we assumed K to be convex and compact, in such a way that Λ can be given by $\Lambda = co(L)$, a convex compact set. The connection between these concepts and Hypothesis 1.1 will be done via an equivalent variational reformulation within the framework of classical linear problems [Ce83b], and hence, a well established technique.

Summarizing, the proof of Theorem 1.1 is based on the usage of well known techniques, put together in such a way that the existence of solution will depend on the verification of the technical Hypothesis 1.1, which in turn should be ensured by a more explicit condition characterizing the "good" functions c and Q. Condition (1.25) plays the role of such explicit condition. It surely will not be checkable in all cases we could wish, but applying this result to the optimal control problem coming from the underwater vehicle control model, described in Section 1.2, can give us a feeling of how such condition can be checked in many vector control problems, with nonlinear dependence on the control.

As to the assumptions made before using the relaxation, generalized moments, variational reformulation, the classical techniques leading to Hypothesis 1.1, it is sure that the fact of K being bounded has simplified a lot their usage. Even so, we believe that appropriate growth conditions, can be considered and successfully used in approaching the case where K is unbounded with similar techniques. Those have been used for instance in [MP01], [Ce83b] and [EMeP03] for dealing with relaxation through Young measures, variational reformulation and problems of moments respectively. Doing this can be the subject of future work.

Also, besides the application to the underwater vehicle, we propose a class of problems to which a direct recipe can be given as a sufficient condition to have (1.25):

(SP)

$$\text{Minimize in } u: \quad \int_0^T \left[\sum_{i=1}^n c_i(x(t))u_i(t) + \sum_{i=1}^n c_{n+i}(x(t))u_i^2(t) \right] dt \tag{1.27}$$

subject to

$$x'(t) = Q_0(x(t)) + Q_1(x(t))u(t) + Q_2(x(t))u^2(t) \text{ in } (0,T), \tag{1.28}$$

$$x(0) = x_0 \in \mathbf{R}^n, \text{ and } u(t) \in K \subset \mathbf{R}^n. \tag{1.29}$$

This problem is square in the sense that we take here $N = n$. Also, Q_1 and Q_2 are $n \times n$ matrices, with Q_1 non-singular, and such that, together with the vector Q_0, comply with appropriate technical hypotheses so that the state law is a well-posed problem. Under these hypotheses, we can set

$$Q = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix}, \quad c = \begin{pmatrix} C_1 & C_2 \end{pmatrix}, \tag{1.30}$$

where $C_1, C_2 \in \mathbf{R}^n$, and put

$$D(x) = -(Q_1)^{-1}Q_2, \quad E(x) = C_1 D + C_2, \quad U(m,x) = 2\sum_i m_i e_i \otimes e_i D - \mathrm{id}, \quad m = \Phi(u), \tag{1.31}$$

where the e_i's stand for the vectors of the canonical basis of $\mathbf{R}^n$, id is the identity matrix of size $n \times n$, and

$$\Phi(u) = (u_1, ..., u_n, (u_1)^2, ..., (u_n)^2).$$

We are then in condition to state the following result.

Theorem 1.2. *Suppose that for the ingredients (c, Q, K) of (SP), we have*

1. *the matrix U is always non-singular for $u \in K$, and $x \in \mathbf{R}^n$;*
2. *for such pairs (u, x), we always have $U^{-T}E < 0$, componentwise.*

Then, condition (1.25) is verified and the optimal control problem admits solutions.

The proof is mainly algebraic and consist in seeing that if the functions (c, Q) verify these algebraic properties with respect to K and Φ, then they also verify the algebraic conditions characterizing (1.25).

The problem with giving "recipes" is that they depend on the structure of the problem. In Theorem 1.2 the squared feature is essential to the conclusion. This indicates that there could be as many recipe results as there are different problems. While we are not sure about

this, the application to the underwater vehicle model seems to indicate that checking directly condition (1.25) could be a good approach.

As these questions relate to condition (1.25), we may think of alternative ways to ensure Hypothesis 1.1. A few possibilities are proposed in Chapter 2, where some exhaustive work is included for the cases $N = n = 1$, $\Phi(u) = (u, u^2)$ and $\Phi(u) = (u, u^2, u^3)$.

1.4.2 Via Variational Reformulation

The following comments refer to the ideas considered in the author's master degree, which where also explored during the period of this thesis. We include such results here for completeness.

We have seen in Section 1.3 that the techniques based on the variational reformulation introduced in [Rk75] don't take us much further in optimal control theory, if we stick to the classical results in the calculus of variations. However, as we also commented, some recent results [Or07], [CO07] may change this scenario. As a matter of fact, applying the results in [Or07] allows us to give some partial answers to question (II) in Section 1.4. It is just for the one-dimensional case, but for general F and f, K not necessarily bounded, and with the possibility of including a viability constraint on the state of the type $x \in L$, where L is a closed set. The fact of being just for one-dimensional problems makes this a minor contribution, but we believe it can be used in the future for higher dimensional situations since the techniques used are highly dependent on the advances on the theory of Calculus of Variations.

We recall that in the mentioned result, the convexity is reduced to a minimum, the so called zero-convexity which states that

$$\varphi^{**}(x, 0) = \varphi(x, 0), \ \forall x,$$

where φ is the integrand in the cost function of the variational problem, and $\varphi**$ represents, once more, the bipolar function of φ with respect to the state derivative.

The equivalence between the optimal control problem (P), and its variational reformulation (VP) proceeds in a standard way, and the zero-convexity can be seen to be equivalent to the condition

$$\sup_{\xi<0} \frac{\varphi(\xi) - \varphi(0)}{\xi} \leq \inf_{\xi>0} \frac{\varphi(\xi) - \varphi(0)}{\xi} \ . \tag{1.32}$$

Therefore translating this condition for the control problem, writing it in terms of F, f and K gives

$$\sup_{u\in K,\ f(x,u)<0} \frac{F(x,u) - m(x)}{f(x,u)} \leq \inf_{u\in K,\ f(x,u)>0} \frac{F(x,u) - m(y)}{f(x,u)} \tag{1.33}$$

where $m(x) = \min_{u\in K}\{F(x,u) : f(x,u) = 0\}$ for every $x \in L$.

Hence an existence theorem can be proved, assuming appropriate, typical growth and regularity conditions, but replacing either the traditional or more recent convexity assumptions based on the orientor field

$$\mathcal{A}_x = \{(v,z) : \ v > F(x,u),\ z = f(x,u),\ u \in K\}$$

with the easily checkable condition (1.33).

Theorem 1.3. *Consider problem (P) with the additional constraint $x \in L \subset \mathbf{R}$. Suppose that F is lower semi-continuous, f is continuous, K and $L \subseteq \mathbf{R}$ are closed sets. Assume that condition (1.33) is true, together with the typical growth conditions:*

(i) *F is coercive with respect to K in the sense*

$$\lim_{|u|\to\infty,\ u\in K} F(x,u) = +\infty, \qquad \textit{uniformly in } x;$$

(ii) *F is bounded from below, and*

$$\lim_{|f(x,u)|\to\infty,\ u\in K} F(x,u)/|f(x,u)| = +\infty, \ \textit{uniformly in } x.$$

Then, there is an absolutely continuous function y, and a measurable function u, such that the pair (y,u) is an optimal solution for (P).

In case of K being a bounded set, the typicall growth conditions lose their meaning and the zero-convexity can be translated to a even simpler condition

$$\min_u\{F(y,u) : \ f(y,u) = 0,\ u \in K\} = \min_u\{F(y,u) : \ u \in K\}.$$

As the zero-convexity seems not to be sufficient to prove the existence of solution for vector variational problems, the definition of almost convex function is considered (see [CO07]). The way how such definition can be interpreted in terms of optimal control problems can be the purpose of subsequent work.

Some simple examples can be given. They don't pretend to compare these results with classical ones, but give an illustration of how the sufficient conditions can be easily checked. In fact, the results commented here don't pretend to generalize others results. They are just an example of the power of the equivalent variational reformulations, applied to control problems, whenever good results for the calculus of variations are available.

1.4.3 Aplication to an Underwater Vehicle Problem

As mentioned in Section 1.2, the problem of controlling and maneuvering an underwater vehicle can be described by an optimal control problem of Bolza type, which by turn can be rewritten as a problem of the type (P_1). In fact (see [GOP09]), let us consider the functional

$$\frac{1}{2}\left\|x\left(T\right)-x^{T}\right\|^{2}+\int_{0}^{T}\left\|u\left(t\right)\right\|^{2}dt$$

which simplifies (except for some constant coefficients) the cost function, whose minimizers (x,u) should be such that, the state $x\in\mathbf{R}^{12}$ approaches x^T at time T, while $u\in\mathbf{R}^{3}$ is kept to be the less expensive possible. It can be rewritten as

$$\frac{1}{2}\left\|x\left(0\right)-x^{T}\right\|^{2}+\int_{0}^{T}\left[\frac{1}{2}\frac{d}{dt}\left\|x\left(t\right)-x^{T}\right\|^{2}+\left\|u\left(t\right)\right\|^{2}\right]dt$$

$$=\frac{1}{2}\left\|x\left(0\right)-x^{T}\right\|^{2}+\int_{0}^{T}\left[<x(t)-x^{T},x'(t)>+\left\|u\left(t\right)\right\|^{2}\right]dt.$$

Furthermore, the state is given by the control system $x'=f(\mathtt{x},u)$, described in [GOP09]. Even if in such system the dependence on both state and control variables is nonlinear, its structure allows us to rewrite it in the form

$$x'(t)=Q\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right)+Q_{0}\left(x\left(t\right)\right)$$

where

$$\Phi(u)=\left(u_1,u_2,u_3,(u_1)^2,(u_2)^2,(u_3)^2\right)$$

and

$$Q:\mathbf{R}^{12}\rightarrow\mathcal{M}^{12\times6}\quad\text{and}\quad Q_0:\mathbf{R}^{12}\rightarrow\mathbf{R}^{12}.$$

We can therefore consider a cost given by

$$\begin{aligned}&\int_{0}^{T}\left[<x\left(t\right)-x^{T},\ Q\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right)+Q_{0}\left(x\left(t\right)\right)>+\left\|u\left(t\right)\right\|^{2}\right]dt\\ =\ &\int_{0}^{T}\left[c\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right)+c_{0}\left(x\left(t\right)\right)\right]dt\end{aligned}$$

where the vector c is given by

$$\begin{cases}c_{i}\left(x\right)=\sum_{j=1}^{12}\left(x-x^{T}\right)_{j}Q_{ji}, & i=1,2,3,\\ c_{i}\left(x\right)=\sum_{j=1}^{12}\left(x-x^{T}\right)_{j}Q_{ji}+1, & i=4,5,6,\end{cases}$$

and

$$c_{0}\left(x\right)=<x-x^{T},\ Q_{0}\left(x\right)>.$$

Admissible controls u are measurable functions that should lie in a certain set $K \subset \mathbf{R}^3$, which, in our case, is given by

$$K = [-a_1, a_1] \times [-a_2, a_2] \times [-a_3, a_3],$$

with $0 < a_1, a_2, a_3 < \pi/2$. Hence we are dealing with a control problem of the type

$$\text{(UVP)} \quad \text{Minimize in } u : \int_0^T \left[c\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right) + c_0\left(x\left(t\right)\right)\right] dt$$

subject to

$$x'\left(t\right) = Q\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right) + Q_0\left(x\left(t\right)\right), \quad t \in (0,T)$$

$$x\left(0\right) = x^0 \in \Omega$$

$$x\left(t\right) \in \Omega \quad \text{and} \quad u\left(t\right) \in K,\ 0 \le t \le T.$$

The question if the initial value problem IVP

$$\mathtt{x}'(t) = Q\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right) + Q_0\left(x\left(t\right)\right)$$

$$x(0) = x_0$$

is well posed is not minor, as the mappings Q and Q_0 can be very complicated. Using the dimensionless coefficients in [GOP09] we see that those matrices have the following structure:

$$Q(x) =$$

$$= (x_7)^2 \begin{pmatrix} & 0_{6\times 6} & & & & \\ -0.0056307 & -0.0056219 & 0.0002292 & -0.0028418 & -0.0011310 & -0.0037067 \\ 0 & 0 & -0.0001291 & 0 & 0 & 0 \\ 1.527832 & 1.4903911 & -0.0617573 & -0.0001656 & -0.0000659 & -0.0002160 \\ 0 & 0 & 0.0001049 & 0 & 0 & 0 \\ -0.0162938 & -0.0162684 & 0.0006631 & 0 & 0 & 0 \\ 0 & 0 & -0.0002773 & 0 & 0 & 0 \end{pmatrix}.$$

In particular, this means that Q only depends on the surge velocity $\mathtt{u}$ (here x_7);

$$Q_0(x) = \begin{pmatrix} \begin{pmatrix} J_1(x_4, x_5, x_6) & 0_{3\times 3} \\ 0_{3\times 3} & J_2(x_4, x_5, x_6) \end{pmatrix} \begin{pmatrix} x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \end{pmatrix} \\ \bar{F}_0(x) \end{pmatrix}$$

where J_1, J_2 and $\bar{F}_0$ are nonlinear mappings (see [GOP09] for explicit expressions).

The specific structure of the dynamics, namely the composition of the matrices J_1 and J_2 and the vector $\bar{F}_0$, which include, component-wise, only polynomial terms, absolute values, terms of the type $\sqrt{x_j^2 + x_k^2}$, and products of this type of terms, allows us to conclude that there is a minimum time T for which the mapping

$$(t, x) \to Q(x)\Phi(u(t)) + \Phi(x)$$

verify the typical Lipschitz condition with respect to x, over the interval $(0, T)$, and uniformly in $u \in L^{+\infty}((0, T), K)$, such that for every measurable function u taking values in K, the corresponding (IVP) has a unique solution x.

The next issue is to see that condition (1.25) is verified. For this purpose, we need to characterize for every $x \in \mathbf{R}^{12}$ the set

$$\mathcal{N}(c(x), Q(x)) = \left\{ v \in \mathbf{R}^6 : Q(\mathbf{x})v = 0, c(x) \cdot v \leq 0 \right\},$$

to guarantee that it is contained in

$$\mathcal{N}(K, \Phi) =$$

$$\left\{ v = (v_1, \cdots, v_6) \in \mathbf{R}^6 : \text{for each } u \in K, \text{ either } \nabla\Psi(\Phi(u))v = 0 \text{ or there is } i \text{ with} \nabla\Psi_i(\Phi(u))v > 0 \right\},$$

where $\Psi : \mathbf{R}^6 \to \mathbf{R}^3$ is the component-wise convex $\mathcal{C}^1$ mapping

$$\Psi(m) = (m_1^2 - m_4, m_2^2 - m_5, m_3^2 - m_6), \quad m = (m_1, \cdots, m_6).$$

After some computation, using the particular composition of Q, we can see that the condition

$Qv = 0$ is equivalent to

$$\begin{cases} v_1 = 0 \\ v_2 = 0 \\ v_3 = 0 \\ v_6 = -\frac{1}{Q_{16}}(Q_{14}v_4 + Q_{15}v_5) = \\ -\frac{1}{Q_{36}}(Q_{34}v_4 + Q_{35}v_5). \end{cases}$$

This can be used to conclude that for every v such that $Qv = 0$ the vector $\nabla\Psi(m) \cdot v$ is in fact

$$2\ \mathtt{diag}[m_1, m_2, m_3] \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} - I_3 \begin{pmatrix} v_4 \\ v_5 \\ v_6 \end{pmatrix} = - \begin{pmatrix} v_4 \\ v_5 \\ v_6 \end{pmatrix}.$$

The conclusion follows then without many problems.

We notice that all the computations became easier due to the relatively simple structure of the matrix Q.

We are, therefore, in conditions to state the following existence result, for the optimal control problem coming from the underwater vehicle model.

Theorem 1.4. *For $T > 0$, small enough, there exists an optimal solution for (UVP).*

1.5 Future Work

1.5.1 Numerical Approximations

Improving numerical approximations for the solutions of optimal control problems has been one of the most important research lines within this field. Many different techniques have been used to approach optimal control problems of the type we consider here. Among the most common are the shooting methods for problems with initial and final constraints, techniques based on the necessary conditions given by the Pontryagin maximum principle, or the sufficient conditions described by the Hamilton-Jacobi-Bellman equation. Also descent methods, and more recently, the direct discretization of the original problem have been considered. For more details on this see for instance [Py99], [Sa00], [Sm06] and [Tr08].

Considering the optimal control problem (UVP) described above, the numerical approximation of solutions was tested successfully in [GOP09]. There, necessary optimality conditions were used, although this method seemed to be rather demanding in terms of computational cost.

Hence, we would like to implement another algorithm sufficiently stable to approximate optimal solutions for such a class of problems, and capable, in an efficient manner, to deal with the non linear dependences on the state and control. As these variables take values in $\mathbf{R}^{12}$ and $\mathbf{R}^3$ respectively, we believe we should avoid either solving Hamilton-Jacobi-Bellman equation (a partial differential equation), either the previous analysis required for the shooting method. Also, although these type of problems can be considered to have low-medium dimension, the presence of nonlinear terms, together with the possibility of needing large time discretizations, make us suspect that a full discretization will not be very stable. This could be an option for partial problems with smaller dimensions involved.

Accordingly to Pytlak's opinion ([Py99]), gradient (descent) methods can handle with state and control constraints quite well, and approximate solutions with the precision required for problems in engineering.

Let us explain better the idea behind the descent method.

Consider the optimal control problem

$$\text{Minimize} \quad \mathcal{I}(x,u) = \int_{t0}^{t1} F(t,x,u)dt$$

subject to

$$x' = f(t,x,u)$$

$$u \in K \subset \mathbf{R}^n,\ x(t0) = x_0$$

where K is the set of vectors u s.t. $k_i^0 \leq u_i \leq k_i^1$.

Let $I = (t0, t1)$. An admissible pair (x,u) verifies

$$x \in AC(I, \mathbf{R}^m) \quad \text{and} \quad u \in L^\infty(I, \mathbf{R}^n).$$

Notice that we have included here the explicit time dependence in F and f. This is because many interesting examples have such dependence, and the ideas for numerical approximation don't change much with such generality.

Typically, the steepest descent direction is computed with respect to the augmented cost functional with integrand $H = F + pf$, the so called Hamiltonian, where p represents the costate vector function ([EP76], [MPD70]).

Another way of proceeding, also standard (see [MPD70], [PyV98], [MW03]), is to consider the original cost $\mathcal{I}(x,u)$, and for an admissible pair (x,u), find a direction, say U, such that

$$\mathcal{I}(x_\varepsilon, u + \varepsilon(U-u)) \leq \mathcal{I}(x,u)$$

where x_ε is the solution of the initial value problem associated with the control $u_\varepsilon = u+\varepsilon(U-u)$ and $u_\varepsilon \in K$. To this purpose, ε should lie in $(0,1)$ and either set to be sufficient small or found by one dimension minimization of the ε-parametrized cost

$$\mathcal{I}(x_\varepsilon, u+\varepsilon(U-u)).$$

A manner of searching the direction U is finding the solution the linear optimal control problem

$$(LP) \quad \min_U \int_I \nabla F(t,x(t),u(t)) \cdot (X(t),U(t))dt$$

where X is the solution of

$$X'(t) = \nabla f(t,x(t),u(t)) \cdot (X(t),U(t)-u(t))$$

such that

$$X(0) = 0, \quad U(t) \in K, \quad t \in I.$$

Actually, this means that $(X, u-U)$ minimize the first variation of $\mathcal{I}$ around (x,u)

$$\int_I \nabla F(t,x,u) \cdot (X, U-u)dt$$

where X verifies the linearization of the control system on the neighborhood of (x,u)

$$X' = \nabla f(t,x,u) \cdot (X, U-u).$$

The important step of this idea is solving (LP), which can be done either via optimality conditions, or by a direct discretization into a linear mathematical programming problem, solved by appropriate methods for such finite dimension optimization problems. We prefer this option as available packages for linear optimization, like CPLEX, can be very efficient.

A different approach can be to see that at each iterative step, (LP) can be written as

$$\text{Minimize in } U \in K: \quad \int_I (a(t)X(t) + b(t)U(t))\,dt,$$

subject to

$$X'(t) = A(t)X(t) + B(t)(U(t)-u(t)), t \in I, \quad X(t0) = 0, \tag{1.34}$$

$$U(t) \in K, \quad t \in I,$$

proceed to the Rockafellar's variational reformulation

$$\varphi(t,X,\xi) = \begin{cases} +\infty \quad \text{if} \quad \xi \notin A(t)X + B(t)(K-u(t)) \\ \min\limits_{U\in K}\{a(t)X + b(t)U : \xi = A(t)X + B(t)(U-u(t))\} \quad \text{else} \end{cases}$$

and consider the genuine variational problem (VP)

$$\text{Minimize in } X: \quad \int_I \varphi(t, X(t), X'(t))\, dt$$

subject to $X(t0) = x_0$.

The equivalence between problems (LP) and (VP) is well established ([BP01], [P03], [PT07]). Therefore, we can use (VP) to find X and then recover U via the algebraic system of equations (B.3). This is one advantage, as the minimization will proceed only with respect to the state variable, instead of (X, U). Another one is that using an appropriate minimization method for (VP) can be potentially better than, for instance, solving optimality conditions for (LP), a system of ordinary differential equations. Notice also that even if φ is defined through a minimum, evaluating it consists in a lower dimension linear optimization problem. A major difficulty in using (VP) is that φ is not a continuous function and takes infinite values.

The question of choosing the best method to solve (VP) and compare the results between both approaches is still open, and it could be a line of future work. In fact, we would like to carry trough three different stages:

- (I) Implementation of the global method with the iterative step based on direct discretization of (LP).
- (II) When the global method is considered to be stable, try to improve it by applying an appropriate method to approximate the solutions of (VP) instead of dealing with (LP).
- (III) Implementation of the global method to the problem (UVP) related to the underwater vehicles model described in [GOP09].

At the moment, we have made some advances for concluding the first stage. We have analyzed many academical examples where we tried to approximate the optimal solution with the method described above. We used CPLEX compiled with GAMS (Generic Algebraic Modeling System) to approximate the solution of problem (LP), at the iterative step. For some of them, we compared the results with the ones obtained by solving the nonlinear problem coming from the full discretization of the original problem (P).

Although we can obtain good results for some class of problems, we consider that we should improve the implementation, before going through the next stages. This is because when we increase the state and control dimensions, while searching for a structure similar to the one of (UVP), some unexpected bad behavior occurs.

1.5.2 Other Existence Results

Concerning the relaxation methods through parametrized measures, possible future work can be:

- Go through a deeper understanding of condition (1.25), finding more examples, counter-examples, and possible alternatives and generalizations to it, as a sufficient condition to ensure Hypothesis 1.1.
- Find necessary conditions to Hypothesis 1.1.
- Generalize the existence results for the cases $c = (t, x)$, $Q(t, x)$, $\Phi(t, u)$ and $U = U(t, x)$, either bounded or unbounded. Include different types of constraint over the state variable as well as to Bolza type cost functions.
- Apply these techniques to more general situations than the one in (P_1).

Considering the classical techniques based on the variational reformulation, an immediate goal can be

- Apply the existence results for calculus of variations in [CO07] to the equivalent variational problem to the general optimal control problem (P), and see how the almost convexity with respect to the state derivative can be translated into the control problem. This should be done in the spirit of [PT07].

1.5.3 Other Applications

Concerning optimal control problems coming from underwater vehicles, several possibilities can be borne in mind, namely

- Numerical approximation of solutions for the optimal control problem coming from the modelling of submarines ([GOP09]),
- Existence results which can be applied to more complicated models including time dependent viability constraints over the state and control variables, and constraints on the derivative, eventually taking us to the field of differential inclusions $x'(t) \in Q(t)$.

1.6 Final Comments

As we have seen, with this thesis we pretend to go further in the existence theory available for optimal control problems of the type (P_1). This class of problems are still general enough

to be applied to important problems coming from the maneuvering and control of underwater vehicles and possibly other applied problems as well. We also can give some contribution for the more general problem (P) in the scalar case.

For such problems, techniques based on convexity properties of the orientor field $\mathcal{A}$, although general, can be very hard to check for vectorial problems, with nonlinear dependence on the control like (P_1). Techniques based on using Lyapunov's Theorem to obtain an optimal solution from the generalized optimal solution from the relaxed problem seems to be associated with to a separate structure of the type $F(x,u) = G(u) + H(u)$. Similar difficulties seems to arise when trying to check the concavity of the functional coming from the relaxation of the cost function using parametrized measures. The equivalent variational reformulation's technique depends directly of the available theory for the calculus of variations. For this reason, future paths are not excluded.

Concerning the search of sufficient conditions for an optimal solution of the relaxed problem, to be an optimal solution for the original problem, while using parametrized measures and some standard techniques, a technical crucial hypothesis is presented whose easier verification will be connected, in a non expected way, to the convexity of the space of moments

$$\Lambda = \{m \in R^s : m = \int_K \Phi(\lambda) d\mu(\lambda) \ \mu \in P(K)\}$$

where Φ is a vectorial mapping describing the nonlinear dependence on the control.

An existence result can be established and applied locally to the optimal control coming from the real model proposed in [GOP09].

Among future lines of research, we would like to stress improving standard methods for numerical approximation of optimal solutions and application to underwater vehicle problems.

The chapters that follow are organized as follows:

In Chapter 2, our main contribution to the existence theory with respect to problem (P_1) is introduced. The proof of Theorem 1.1, Theorem 1.2, as well as the proof of other existence results for the scalar case, are presented according to the comments in Section 1.4.1. This chapter corresponds to the work in [PT09].

As to Chapter 3, some existence results for scalar general problems of type (P) are shown, as indicated in Section 1.4.2. This corresponds to the work published in [PT07], concluded during the thesis period, but started earlier within the master thesis's programme [Ti04].

Chapter 4 deals with the application of the existence result for problem (P_1) to the problem (UVP) coming from the underwater vehicle control problem. This material is included in the work [PerT09] submitted for publication.

Some preliminary results on Young measures and optimal control problems are gathered in Appendix A.

In Appendix B, we present some ideas for future research. Among them, an approach to some numerical techniques is proposed, while a couple of simple academic examples are shown for illustration.

Chapter 2

Main Existence Result

2.1 Introduction

This chapter focuses on the analysis of optimal control problems of the general form

$$(P_1) \qquad \text{Minimize in } u: \quad \int_0^T [\sum_{i=1}^{s} c_i(x(t))\phi_i(u(t))]\, dt \tag{2.1}$$

subject to

$$x'(t) = \sum_{i=1}^{s} Q_i(x(t))\phi_i(u(t)) \text{ in } (0,T), \quad x(0) = x_0 \in \mathbf{R}^N, \tag{2.2}$$

and

$$u \in L^\infty(0,T), \quad u(t) \in K, \tag{2.3}$$

where $K \subset \mathbf{R}^m$ is compact. The state $x : (0,T) \to \mathbf{R}^N$ takes values in $\mathbf{R}^N$.

The mappings

$$c_i : \mathbf{R}^N \to \mathbf{R}, \quad \phi_i : \mathbf{R}^m \to \mathbf{R}, \quad Q_i : \mathbf{R}^N \to \mathbf{R}^N$$

as well as the restriction set $K \subset \mathbf{R}^m$ will play a fundamental role. We assume, at this initial stage, that c_i are continuous, ϕ_i are of class $\mathcal{C}^1$, and each Q_i is Lipschitz so that the state system is well-posed.

In such a general form, we cannot apply results for non-necessarily convex problems like the ones in [B94], [Ce74], [Ra90] or [RbS97]. Besides, techniques based on Bauer' Maximum Principle ([Ba58]) are quite difficult to extend to our general setting because it is hard to analyze the concavity of the cost functional when the dependence on both state and control comes in product form. Also the Rockafellar's variational reformulation introduced in [Rk75], and well-described in [Ce83b], [ET74b] or recently in [P03] or [PT07], looks as if it cannot avoid assuming a separated dependence on the state and control variables, since this is the structure of the variational problem for which the existence of solution has been so far ensured ([CC90]).

Concerning the classical Filippov-Roxin theory introduced in [F62] and [R62], it is not easy at all to know if typical convexity assumptions hold, or when they may hold, as we can see from the examples and counter-examples in [Ce83b]. When analyzing explicit examples, one realizes such difficulties coming from the need of a deep understanding of typical orientor fields. The same troubles would arise when applying refinements of this result as the ones in [Mk99] and [MP01].

Recently ([CFM06]), an existence result has been shown for minimum time problems where the typical convexity assumptions over the set valued function on the differential inclusion has been replaced by more general conditions. In fact, the intersection of this result with the ones we present here is not empty although, as we will comment, our frame extends to situations not covered by this result. Such analysis can be done by writing problem (P_1) as a minimum time problem as suggested in [Ce83b].

Our aim is to provide hypotheses on the different ingredients of the problem so that existence of solutions can be achieved through an independent road. Actually, it is not easy to claim whether our results improve on classical or more recent general results. They provide an alternative tool which can be more easily used in practice than such results when one faces an optimal control problem under the special structure we consider here. As a matter of fact, convexity will also occur in our statements but in an unexpected and non-standard way.

Before stating our main general result, a bit of notation is convenient. We will write

$$c : \mathbf{R}^N \to \mathbf{R}^s, \quad \phi : \mathbf{R}^n \to \mathbf{R}^s, \quad Q : \mathbf{R}^N \to \mathbf{R}^{Ns}, \tag{2.4}$$

with components c_i, ϕ_i, and Q_i, respectively. These are the main ingredients of the problem. Some geometrical properties of $\Phi(K)$ will play an important role but although such properties can be easily handled in lower dimensions, when n increases, we need to see $\Phi(K)$ as part of a n-submanifold embedded in $\mathbf{R}^s$. For that purpose, consider a new ingredient of the problem related to ϕ. Suppose that there is a $\mathcal{C}^1$ mapping

$$\Psi : \mathbf{R}^s \to \mathbf{R}^{s-n}, \quad \Psi = (\psi_1, ..., \psi_{s-n}), \quad (s > n), \tag{2.5}$$

so that $\Phi(K) \subset \{\Psi = 0\}$. This is simply saying, in a rough way, that the embedded (parametrized) manifold $\Phi(K)$ of $\mathbf{R}^s$ is part of the manifold defined implicitly by $\Psi = 0$. In practical terms, it suffices to check that the composition $\Psi(\phi(u)) = 0$ for $u \in K$.

For a pair (c, Q), put

$$\mathcal{N}(c, Q) = \{v \in \mathbf{R}^s : Qv = 0, cv \leq 0\}. \tag{2.6}$$

Similarly, set

$$\mathcal{N}(K,\phi) = \tag{2.7}$$

$$\{v \in \mathbf{R}^s : \text{ for each } u \in K, \text{ either } \nabla\Psi(\phi(u))v = 0 \text{ or there is } i \text{ with } \nabla\psi_i(\phi(u))v > 0\}.$$

Our main general result is the following.

Theorem 2.1. *Assume that the mapping Ψ as above is strictly convex (componentwise) and $\mathcal{C}^1$. If for each $x \in \mathbf{R}^N$, we have*

$$\mathcal{N}(c(x), Q(x)) \subset \mathcal{N}(K,\phi), \tag{2.8}$$

then the corresponding optimal control problem (P_1) admits at least one solution.

As it stands, this result looks rather abstract, and it is hard to grasp to what extent may be applied in more specific situations.

A particular, yet still under some generality, situation where this result can be implemented is the case of polynomial dependence where the ϕ_i's are polynomials of various degrees. The main structural assumption, in addition to the one coming from the set K, is concerned with the convexity of the corresponding mapping Ψ.

Suppose we take $\phi_i(u) = u_i$, for $i = 1, 2, \ldots, n$, and $\phi_{n+i}(u)$, $i = 1, 2, \ldots, s-n$, convex polynomials of whatever degree, or simply polynomials whose restriction to K is convex. In particular, K itself is supposed to be convex. Then we can take

$$\Psi_i(v) = \phi_{n+i}(\overline{v}) - v_{n+i}, \quad i = 1, 2, \ldots, s-n, \quad \overline{v} = (v_i)_{i=1,2,\ldots,n}. \tag{2.9}$$

In this case, it is clear that

$$\Psi(\phi(u)) = 0 \text{ for } u \in K,$$

by construction, and, in addition, Ψ is smooth and convex. The important constraint (2.8) can also be analyzed in more concrete terms, if we specify in a better way the structure of the problem.

As an illustration, though more general results are possible, we will concentrate on an optimal control problem of the type

(P)

$$\text{Minimize in } u: \quad \int_0^T \left[\sum_{i=1}^n c_i(x(t))u_i(t) + \sum_{i=1}^n c_{n+i}(x(t))u_i^2(t)\right] dt \tag{2.10}$$

subject to

$$x'(t) = Q_0(x(t)) + Q_1(x(t))u(t) + Q_2(x(t))u^2(t) \text{ in } (0,T), \tag{2.11}$$

$$x(0) = x_0 \in \mathbf{R}^n, \text{ and } u(t) \in K \subset \mathbf{R}^n. \tag{2.12}$$

We are taking here $N = n$. Q_1 and Q_2 are $n \times n$ matrices that, together with the vector Q_0, comply with appropriate technical hypotheses so that the state law is a well-posed problem. Set

$$Q = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix}, \quad c = \begin{pmatrix} c_1 & c_2 \end{pmatrix}, \tag{2.13}$$

where Q_1 is a non-singular $n \times n$ matrix, and $c_1 \in \mathbf{R}^n$. In addition, we put

$$D(x) = -(Q_1)^{-1}Q_2, \quad E(x) = c_1 D + c_2, \quad U(m,x) = 2\sum_i m_i e_i \otimes e_i D - \mathrm{id}, \quad m = \phi(u), \tag{2.14}$$

where the e_i's stand for the vectors of the canonical basis of $\mathbf{R}^n$, and id is the identity matrix of size $n \times n$.

Theorem 2.2. *Suppose that for the ingredients (c, Q, K) of (P), we have*

1. *the matrix U is always non-singular for $u \in K$, and $x \in \mathbf{R}^n$;*
2. *for such pairs (u, x), we always have $U^{-T}E < 0$, componentwise.*

Then the optimal control problem admits solutions.

As a more specific example of the kind of existence result that can be obtained through this approach, we state the following corollary whose proof amounts to going carefully through the arithmetic involved after Theorem 2.2.

Corollary 2.1. *Consider the optimal control problem*

$$\textit{Minimize in } u: \quad \int_0^T [c_1(x(t))(u_1(t))^2 + c_2(x(t))u_2(t)^2]\,dt$$

under

$$x_1'(t) = u_1(t) - u_2(t) + q_1(x)u_1(t)^2 + u_2(t)^2,$$
$$x_2'(t) = q_2(x)u_1(t) + u_2(t) + u_1(t)^2 + u_2(t)^2,$$

and an initial condition $x(0) = x_0$, where $u(t) \in K = [0,1]^2$,

$$q_1(x) \in (\frac{1}{3}, 1), \quad q_2(x) \in (-1, 1), \quad c_1(x), c_2(x) > 0,$$

and

$$\frac{q_1(x)+1}{2}c_2(x) < c_1(x) < \frac{2(q_1(x))^2 + q_1(x)(q_2(x)+1) - q_2(x) - 3}{4(q_1(x)-1)}c_2(x).$$

Then there is, at least, one optimal solution of the problem.

Our strategy to prove these results is not new as it is based on the well-established philosophy of relying on relaxed versions of the original problem, and then, under suitable assumptions, prove that there are solutions of the relaxed problem which are indeed solutions of the original one ([Ck75], [Ga62], [Ms40], [Ms67], [W62] and [Y37]). From this perspective, it is a very good example of the power of relaxed versions in optimization problems.

The relaxed version of the problem that we will be using is formulated in terms of Young measures associated with sequences of admissible controls. These so-called parametrized measures where introduced by L. C. Young ([Y37], [Y42] and [Y69]), and have been extensively used in Calculus of Variations and Optimal Control Theory (see for example [MP01], [P97b], [Rb96] and [Rb97b]). Because of the special structure of the dependence on u, we will be concerned with (generalized) "moments" of such probability measures. Namely, the set

$$L = \{m \in \mathbf{R}^s : m_i = \phi_i(u), 1 \leq i \leq s, u \in K\}, \tag{2.15}$$

and the space of moments

$$\Lambda = \left\{ m \in \mathbf{R}^s : m_i = \int_K \phi_i(\lambda)\, d\mu(\lambda), 1 \leq i \leq s, \mu \in P(K) \right\} \tag{2.16}$$

will play a fundamental role. Here $P(K)$ is the convex set of all probability measures supported in K. Since the mapping

$$M : \mu \in P(K) \mapsto \Lambda, \quad M(\mu) = \int_K \phi(\lambda)\, d\mu(\lambda)$$

is linear, we easily conclude that Λ is a convex set of vectors, and, in addition, that the set of its extreme points is contained in L. In fact, for some particular ϕ_i's of polynomial type, the set of the extreme points of Λ is precisely L. We examine and comment on the set Λ in Section 2.3. This is closely related to the classical moment problem ([Ak61], [ST70] or more recently [EMeP03], [Me04]).

A crucial fact in our strategy is the following.

Assumption 2.1. *For each fixed $x \in \mathbf{R}^N$, and $\xi \in Q(x)\Lambda \subset \mathbf{R}^N$, the minimum*

$$\min_{m \in \Lambda} \{c(x) \cdot m : \xi = Q(x)m\}$$

is only attained in L.

It is interesting to realize the meaning of this assumption. If we drop the linear constraint $\xi = Qm$ on the above minimum, then the minimum is always attained in a certain point in L simply because a linear function on a convex set will always take its extreme values on extreme points of such convex set. However, precisely the presence of the linear constraint $\xi = Qm$

makes the hypothesis meaningful as the extreme points of the section of Λ by such set of linear constraints may not (indeed most of the time they do not) belong to L, so that the extreme points of the linear function $c \cdot m$ over such convex section may not attain its minimum on L. Our main hypothesis establishes that this should be so, and fundamentally, that the minimum is only attained in L.

Under this assumption, and the other technical requirements indicated at the beginning, one can show a general existence theorem of optimal solutions for our problem.

Theorem 2.3. *Under Assumption 2.1 and the additional well-posedness hypotheses on (c, Q) indicated above, the initial optimal control problem (P_1) admits a solution.*

Notice that we are not assuming any convexity on the set K in this statement. The proof of this theorem can be found in Section 2.2. As remarked before, the proof is more-or-less standard, and it involves the use of an appropriate relaxed formulation of the problem in terms of moments of Young measures ([MP01], [Rb97b]).

Condition (2.8) in Theorem 2.1 is nothing but a sufficient condition to ensure Assumption 2.1 in a more explicit way. As a matter of fact, all of our efforts are directed towards finding in various ways more explicit conditions for the validity of this assumption. In this vein, the rest of the chapter focuses on exploring more fully our Assumption 2.1 either through duality, geometric arguments, or in order to prove Theorem 2.1. Ideally, one would like to provide explicit results saying that for a certain set $\mathcal{M}$, Assumption 2.1 holds if for each $x \in \mathbf{R}^N$, $(c(x), Q(x)) \in \mathcal{M}$. In fact, by looking at Assumption 2.1 from the point of view of duality, one can write a general statement whose proof is a standard exercise.

Proposition 2.1. *If for any $x \in \mathbf{R}^N$, $(c, Q) = (c(x), Q(x))$ are such that for every $\eta \in \mathbf{R}^N$ there is a unique $m(\eta) \in L$ solution of the problem*

$$\text{Minimize in } m \in L: \quad (c + \eta Q)m \tag{2.17}$$

then Assumption 2.1 holds.

We briefly comment on this in Section 2.3. One then says that $(c, Q) \in \mathcal{M}$ if this pair verifies the condition on this proposition. A full analysis of this set $\mathcal{M}$ turns out to depend dramatically on the ingredients of the problem. In particular, we will treat the cases $n = N = 1$, and the typical situation of algebraic moments of degree 2 and 3 in Sections 2.4, 2.5 and 2.6. In Section 2.7 we apply our results to a few explicit examples and compare it with the application of the classical Filippov-Roxin theory.

Situations where either $N > 1$ or $n > 1$ are much harder to deal with, specially because existence results are more demanding on the structure of the underlying problem. In particular,

we need a convexity assumption on how the non-linear dependence on controls occurs. We found that (2.8) turns out to be a general sufficient condition for the validity of Assumption 2.1, thus permitting to prove Theorem 2.1 based on Theorem 2.3. Both Theorem 2.2, and Corollary 2.1 follow then directly from Theorem 2.1 after some algebra. This can be found in Section 2.8.

Finally, we would like to point out that one particular interesting example, from the point of view of applications, that adapts to our results comes from the control of underwater vehicles (submarines). See [Br94], [Fs94], and [HL93]. This served as a clear motivation for our work. We plan to go back to this problem in the near future.

2.2 Proof of Theorem 2.3

Consider the following four formulations of the same underlying optimal control problem.

(P_1) The original optimal control problem described in (2.1)-(2.3).

(P_2) The relaxed formulation in terms of Young measures ([MP01], [P97b], [Rb96], [Rb97b]) associated with sequences of admissible controls:

$$\text{Minimize in } \mu = \{\mu_t\}_{t\in(0,T)} : \quad \tilde{I}(\mu) = \int_0^T [\int_K \sum_i c_i(x(t))\phi_i(\lambda)\, d\mu_t(\lambda)]\, dt$$

subject to

$$x'(t) = \int_K \sum_i Q_i(x(t))\phi_i(\lambda)\, d\mu_t(\lambda)$$

and

$$\text{supp}(\mu_t) \subset K, \quad x(0) = x_0 \in \mathbf{R}^N.$$

(P_3) The above relaxed formulation (P_2) rewritten by taking advantage of the moment structure of the cost density and the state equation. If we put $c = (c_1, ..., c_s) \in \mathbf{R}^s$, $Q \in M_{N\times s}$ and m such that

$$m_i = \int_K \phi_i(\lambda)\, d\mu_t(\lambda)\ \forall i \in \{1, ..., s\},$$

then we pretend to

$$\text{Minimize in } m \in \Lambda : \quad \int_0^T c(x(t)) \cdot m(t)\, dt$$

subject to

$$x'(t) = Q(x(t))m(t), \quad x(0) = x_0.$$

(P_4) Variational reformulation of formulation (P_3) (as in [Rk75], [Ce83b], [P03] and [PT07]). This amounts to defining an appropriate density by setting

$$\varphi(x, \xi) = \min_{m\in\Lambda}\{c(x) \cdot m \ : \xi = Q(x)m\}.$$

Then we would like to

$$\text{Minimize in } x(t): \quad \int_0^T \varphi(x(t), x'(t))\, dt$$

subject to $x(t)$ being Lipschitz in $(0,T)$ and $x(0) = x_0$.

We know that the three versions of the problem (P_2), (P_3), and (P_4) admit solutions because they are relaxations of the original problem (P_1). In fact, since K is compact, (P_2) is a particular case of the relaxed problems studied in [MP01] and [Rb97b]. The existence of solution for the linear optimal control problem (P_3) is part of the classical theory ([Ce83b]). Indeed, (P_3) is nothing but (P_2) rewritten in terms of moments, so that the equivalence is immediate. (P_4) is the reformulated problem introduced in [Rk75] whose equivalence to (P_3) was largely explored in [Ce83b] and [P03], [PT07].

Let $\tilde{x}$ be one such solution of (P_4). By Assumption 2.1 applied to a. e. $t \in (0,T)$, we have

$$\varphi(\tilde{x}(t), \tilde{x}'(t)) = \min_{m \in \Lambda} \{c(\tilde{x}(t)) \cdot m(t) : \ \tilde{x}'(t) = Q(\tilde{x}(t))m(t)\} = c(\tilde{x}(t)) \cdot \tilde{m}(t)$$

for a measurable $\tilde{m}(t) \in L$, a solution of (P_3) (see [P03]). The fundamental fact here (through Assumption 2.1) is that $\tilde{m}(t) \in L$ for a.e. $t \in (0,T)$, and this in turn implies that $\tilde{m}(t)$ is the vector of moments of an optimal Dirac-type Young measure $\mu = \{\mu_t\}_{t\in(0,T)} = \{\delta_{\tilde{u}(t)}\}_{t\in(0,T)}$ for an admissible $\tilde{u}$ for (P_1). This admissible control $\tilde{u}$ is optimal for (P_1). This finishes the proof.

2.3 The set Λ and duality

The moment set Λ deserves some comments before proceeding further. Consider the mapping ϕ as in (2.4) and L as in (2.15).

We can regard L as part of an embedded n-manifold in $\mathbf{R}^s$, $s > n$, and ϕ its standard or canonical parametrization. The moment set Λ defined in (2.16) is contained in the convex hull of this manifold.

The most important fact about Λ that one may need in our analysis is stated in the next proposition.

Proposition 2.2. *The set of extreme points of Λ is contained in L.*

Proof. First notice that, as it was shown in [Me04] in a context similar to ours, the compactness of K implies

$$\overline{co(L)} = co(L) = \bar{\Lambda} = \Lambda.$$

The fact of K being bounded plays an important role because otherwise Λ can be shown to be not necessarily closed ([EMeP03]).

Since $\Lambda = co(L)$ then it is known from convex analysis ([Rk70]) that

$$ext(\Lambda) \subseteq L,$$

where $ext(\Lambda)$ represents the extreme points of Λ. □

Remark 2.1. *For some ϕ's it is possible to conclude that $ext(\Lambda) = L$. This is the case, for example when ϕ contains all the linear and quadratic terms of a n-variable polynomial. However this is not essential in what follows.*

Due to this result the proof of Proposition 2.1 is standard (see [Rk70]), so that we shall only make a few remarks.

Since

$$ext(\Lambda) \subseteq co(L)$$

which is a compact set, the minimum of

$$(c + \eta Q)m$$

in Λ is always attained at least in one point of L (it can be attained also in points of $\Lambda \setminus L$). However, if this point happens to be unique, because of Proposition 2.2, it is also immediate to check that it must be the unique minimizer in Λ.

The condition (2.17) in Proposition 2.1 means that

$$\min_{m\in\Lambda}(c + \eta Q)m = \min_{m\in L}(c + \eta Q)m = (c + \eta Q)\phi(a)$$

for a single $a \in K$, which also verifies

$$\min_{m\in\Lambda}(c + \eta Q)m - \eta\xi = (c + \eta Q)\phi(a) - \eta\xi$$

for $\xi \in Q(x)\Lambda$, that is, such that Assumption 2.1 is non empty.

In particular the associated Karush-Kuhn-Tucker vector $\bar{\eta}$ verifies (see [Rk70])

$$c \cdot \phi(a) + \bar{\eta}(Q\phi(a) - \xi) = \min_{m\in\Lambda}\{c \cdot m : \ Qm = \xi\} = c \cdot \phi(a)$$

for a single $a \in K$ complying with $Q\phi(a) = \xi$. As a consequence, for all admissible $m \in \Lambda$ different from $\phi(a)$, we have

$$c \cdot m > c \cdot \phi(a).$$

2.4 Polynomial Dependence. The case $N = n = 1$, $p = 2$

Until Section 8, we concentrate in the situation where

$$\phi : \mathbf{R}^n \to \mathbf{R}^s$$

is such that $\phi_i(u) = u_i$, for $i = 1, 2, \ldots, n$, and $\phi_{n+i}(u)$, $i = 1, 2, \ldots, s-n$ are convex polynomials of some degree p, or simply polynomials whose restriction to K is convex. We will consider K itself to be convex.

Our goal is to explore different possibilities to apply directly Theorem 2.3 by ensuring Assumption 2.1. In other words, we will search for functions

$$c : \mathbf{R}^N \to \mathbf{R}^s, \quad Q : \mathbf{R}^N \to \mathbf{R}^{Ns},$$

such that for every $x \in \mathbf{R}^N$,

$$(c(x), Q(x)) \in \mathcal{M}$$

where $\mathcal{M}$ represents the set

$$\left\{ (c, Q) : \ \forall\, \xi \in Q\Lambda, \ \arg\min_{m \in \Lambda} \{ c \cdot m : \xi = Qm \} \subset L \right\} \tag{2.18}$$

During the following three sections we will focus on the scalar case $N = n = 1$ and use some ideas based on duality (Proposition 2.1) and in geometric interpretations.

In Sections 4, 5, and 6, we explore various scenarios where Assumption 2.1 can be derived, and defer explicit examples until Section 7. In particular, we consider in this section the situation where ϕ is given by $\phi(a) = (a, a^2)$. We are talking about polynomial components of degree less or equal than $p = 2$.

Let $K = [a_1, a_2]$, L, and Λ as in (2.15)-(2.16). Here, we have $s = 2$ and

$$c : \mathbf{R} \to \mathbf{R}^2, \quad Q : \mathbf{R} \to \mathbf{R}^2$$

can be identified with vectors in $\mathbf{R}^2$, or more precisely, with plane curves parametrized by x. To emphasize that function Q is not a matrix-valued but vector-valued, we will call it q.

Next we describe sufficient conditions for $(c(x), q(x)) \in \mathcal{M}$.

Lemma 2.1. *Let K, L and ϕ be as above. For every $x \in \mathbf{R}$, let $q = q(x)$ and $c = c(x)$ be vectors such that one of the following conditions is verified*

1. $q_1 + q_2(a_1 + a_2) = 0$ *and*

$$det \begin{pmatrix} c_1 & c_2 \\ q_1 & q_2 \end{pmatrix} \neq 0;$$

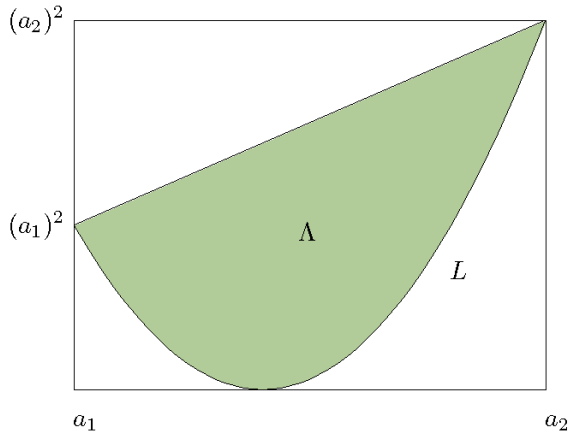


Figure 2.1: $\Lambda = co(L)$ for $p = 2$

2. $q_1 + q_2(a_1 + a_2) \neq 0$ *and*

$$(q_1 + q_2(a_1 + a_2))det \begin{pmatrix} c_1 & c_2 \\ q_1 & q_2 \end{pmatrix} < 0.$$

Then $(c, q) \in \mathcal{M}$*, and consequently Assumption 2.1 is verified.*

Proof. Suppose there is η such that the minimum of $(c + \eta q) \cdot m$ is attained in more than one point of $L = \phi(K)$. This means that the real function

$$g(t) = (c + \eta q) \cdot \phi(t) = (c_1 + \eta q_1)t + (c_2 + \eta q_2)t^2$$

has more than one minimum point over K. For that to happen, either g is constant on t, i. e.,

$$\begin{cases} c_1 + \eta q_1 = 0 \\ c_2 + \eta q_2 = 0 \end{cases} \Leftrightarrow \det \begin{pmatrix} c_1 & c_2 \\ q_1 & q_2 \end{pmatrix} = 0,$$

which contradicts our hypothesis; or else we must have

$$c_2 + \eta q_2 < 0, \quad g'\left(\frac{a_1 + a_2}{2}\right) = 0.$$

This condition can be written as

$$c_1 + (a_1 + a_2)c_2 + \eta[q_1 + (a_1 + a_2)q_2] = 0.$$

If $q_1 + q_2(a_1 + a_2) = 0$, but $c_1 + (a_1 + a_2)c_2 \neq 0$ (condition 1. in statement of lemma), then this equation can never be fulfilled. Otherwise, there is a unique value for η, by solving this equation, which should also verify the condition on the sign of $c_2 + \eta q_2$. It is elementary, after going through the algebra, that the condition on this sign cannot be true under the second condition on the statement of the lemma. □

2.5 The case $N = n = 1$, $p = 3$

We study the case where $\phi(a) = (a, a^2, a^3)$, $s = 3$, and c and q can be identified as vectors in $\mathbf{R}^3$. The understanding of the set Λ and its sections by planes in $\mathbf{R}^3$ is much more subtle however.

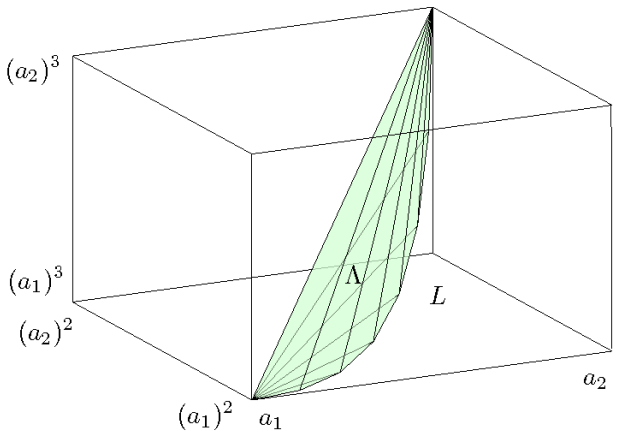


Figure 2.2: $\Lambda = co(L)$ for $p = 3$

To repeat the procedure used for $p = 2$, and apply Proposition 2.1, we would like to give sufficient conditions for the function

$$g(t) = (c + \eta q) \cdot \phi(t) = (c_1 + \eta q_1)t + (c_2 + \eta q_2)t^2 + (c_3 + \eta q_3)t^3 \tag{2.19}$$

to have a single minimum over $K = [a_1, a_2]$ for every η. As indicated, and after some reflection, a complete analysis of the situation is rather confusing and the conditions on the vectors c and q much more involved. To illustrate this, we give a sufficient condition in the following form.

Lemma 2.2. *For all $x \in \mathbf{R}$, let $c = c(x)$ and $q = q(x)$ be vectors in $\mathbf{R}^3$ such that*

$$q_2^2 - 3q_1q_3 < 0, \quad (2c_2q_2 - 3c_1q_3 - 3q_1c_3)^2 - 4(c_2^2 - 3c_1c_3)(q_2^2 - 3q_1q_3) < 0,$$

then $(c, q) \in \mathcal{M}$, and Assumption 2.1 is verified.

Proof. The proof consists in the realization that the conditions on the vectors c and q ensure that the cubic polynomial (2.19) is monotone in all of $\mathbf{R}$ (avoiding degenerate situations), and thus it can only attain the minimum in a single point of any finite interval. Notice that this condition is independent of the interval. In fact, we have to discard the possibility for the derivative of the polynomial $g(t)$ to have roots. This amounts to the negativity of the corresponding discriminant. And this, in turn, is a quadratic expression in η that ought to be always negative.

This occurs when that parabola has a negative discriminant, and the leading coefficient is also negative. These two conditions are exactly the ones in the statement of this lemma. □

A more general condition would focus on considering the local maximizer and the local minimizer of $g(t)$, M_+ and M_-, respectively, and demanding that the interval $[a_1, a_2]$ have an empty intersection with the interval determined by M_+ and M_-. But this would lead to rather complicated expressions. Even so, some times under more specific hypotheses on the form of the vectors c and q, these conditions can be exploited.

Remark 2.2. *Notice that the relation*

$$ext(\Lambda) = L$$

is not true for a general K if it has positive and negative values. However, it is true if we consider $a_1 > 0$ or $a_2 < 0$.

Lemma 2.3. *Let $K = [a_1, a_2]$ with $a_1 > 0$ and*

$$(c, q) = ((0, c_2, c_3), (0, q_2, q_3))$$

such that

$$-\frac{q_2}{q_3} < 0, \quad (c_2, c_3) \cdot (1, -\frac{q_2}{q_3}) < 0.$$

Then the assumptions of Proposition 2.1 are valid, and consequently so is Assumption 2.1.

Proof. In this situation, the maximizer M_+ referred to above is given by

$$M_+ = \frac{-(c_2 + \eta q_2) - |c_2 + \eta q_2|}{3(c_3 + \eta q_3)}$$

If $q_2 > 0$, then $q_3 > 0$, and we have

$$-\frac{c_2}{q_2} > -\frac{c_3}{q_3}.$$

Hence if $\eta \in]-\infty, -\frac{c_2}{q_2}] \setminus \{-\frac{c_3}{q_3}\}$

$$M_+ = \frac{-(c_2 + \eta q_2) + c_2 + \eta q_2}{3(c_3 + \eta q_3)} = 0.$$

If $\eta > -\frac{c_2}{q_2}$,

$$M_+(\eta) = \frac{-(c_2 + \eta q_2) - (c_2 + \eta q_2)}{3(c_2 + \eta q_2)} = \frac{-2(c_2 + \eta q_2)}{3(c_3 + \eta q_3)} < 0.$$

In any case $M_+(\eta) \leq 0$, thus $a_1 > M_+$.

Also if $\eta = -\frac{c_3}{q_3}$,

$$g(t) = (c_2 + \eta q_2)t^2 = (c_2, c_3) \cdot (1, -\frac{q_2}{q_3})t^2$$

which has a unique minimum in K since we have assumed $a_1 > 0$. We conclude that the condition (2.17) in Proposition 2.1 is verified. □

In a very similar way we can prove the following.

Lemma 2.4. *Let $K = [a_1, a_2]$ with $a_2 < 0$ and*

$$(c, q) = ((0, c_2, c_3), (0, q_2, q_3))$$

such that

$$-\frac{q_2}{q_3} > 0, \quad (c_2, c_3) \cdot (1, -\frac{q_2}{q_3}) < 0.$$

Then $(c, q) \in \mathcal{M}$ and consequently Assumption 2.1 is valid.

2.6 A geometric approach to the case $N = n = 1$, $p = 3$.

As we have seen, the use of Proposition 2.1 is simpler only when restricted to some particular classes of examples. Thus we propose a general criteria for obtaining Assumption 2.1, based on a geometric approach.

We first give a result that generalizes the strictly convexity of a ϕ-parametrized plane curve for a 3-dimensional one.

Lemma 2.5. *Let $K = [a_1, a_2]$ with $a_1 > 0$, $\phi(t) = (t, t^2, t^3)$, and L the curve parametrized by ϕ for t in K.*

1. *Given t in K, then for all $s \in K$ such that $s \neq t$ we have*

$$(\phi(s) - \phi(t)) \cdot N(t) > 0$$

   where $N(t)$ is the normal vector to ϕ at t.

2. *For every $t \in K$, $v \in \Lambda = co(L) \setminus \{\phi(t)\}$, we have*

$$(v - \phi(t)) \cdot N(t) > 0.$$

Proof. To check the first part of the statement notice that since

$$\phi'(t) = (1, 2t, 3t^2)$$

and

$$\phi''(t) = (0, 2, 6t)$$

we have that the normal vector, colinear to $\phi'(t) \times \phi''(t)$, is given by

$$N(t) = c_t(-9t^2 - 2t, 1 - 9t^4, 6t^3 + 3t)$$

where $c_t > 0$ is a normalizing constant. Setting

$$N_1 = -9t^3 - 2t, \quad N_2 = 1 - 9t^4, \quad N_3 = 6t^3 + 3t,$$

we find that the solution s of

$$(\phi(s) - \phi(t)) \cdot N(t) = 0$$

also verifies

$$N_3 s^3 + N_2 s^2 + N_1 s - N \cdot \phi(t) = 0,$$

which is equivalent to

$$(s - t)^2(N_3 s + N_2 + 2tN_3) = 0.$$

This means that

$$\hat{s} = -\frac{N_2}{N_3} - 2t = -\frac{3t^4 + 6t^2 + 1}{6t^3 + 3t}$$

is the only solution different from t, but also that it is negative for all $t > 0$, and consequently that it should be excluded. Once we assumed $K \subset \mathbf{R}^+$ and $s \neq t$ the conclusion is immediate.

By using the previous discussion, proving the second part of the statement is trivial once we notice that both m and $\phi(t)$ can be rewritten as

$$\sum_{i=1}^{4} \alpha_i \phi(s_i) \text{ and } \sum_{i=1}^{4} \alpha_i \phi(t)$$

respectively, where $s_i \in K$ and $\sum_{i=1}^{4} \alpha_i = 1$.

□

Another useful lemma.

Lemma 2.6. *If q and c are such that*

$$(\phi'(t) \times (c \times q)) \cdot (\phi(s) - \phi(t)) \tag{2.20}$$

does not change sign for $t, s \in K$, $s \neq t$, then if $v \in \Lambda$, $v \neq \phi(t)$, and $q \cdot (v - \phi(t)) = 0$ we have

$$c \cdot (v - \phi(t)) \neq 0.$$

This means that the linear function c cannot take the same value over $\phi(t)$ and any $v \neq \phi(t)$ in the plane section

$$\{v \in \Lambda : \ q \cdot v = q \cdot \phi(t)\}.$$

Proof. Notice that for $v \in \Lambda$,

$$\begin{aligned}(\phi'(t) \times (c \times q)) \cdot (v - \phi(t)) &= (\phi'(t) \times (c \times q)) \cdot (\sum_{i=1}^{4} \alpha_i \phi(s_i) - \sum_{i=1}^{4} \alpha_i \phi(t)) \\ &= \sum_{i=1}^{4} \alpha_i (\phi'(t) \times (c \times q)) \cdot (\phi(s_i) - \phi(t)) > 0(\text{ or } < 0),\end{aligned}$$

so that the condition stated is also verified for any $v \in \Lambda$.

Suppose now that $v \in \Lambda$ verifies $q \cdot (v - \phi(t)) = 0$ for given $t \in K$ with $v \neq \phi(t)$ and is such that $c \cdot (v - \phi(t)) = 0$, then

$$\begin{aligned}(\phi'(t) \times (c \times q)) \cdot (v - \phi(t)) &= [(\phi'(t) \cdot q)c - (\phi'(t) \cdot c)q] \cdot (v - \phi(t)) \\ &= (\phi'(t) \cdot q)c \cdot (v - \phi(t)) - (\phi'(t) \cdot c)q \cdot (v - \phi(t)) = 0,\end{aligned}$$

a contradiction concerning the argument above. □

We now define the set $\mathcal{M}_1$ of pairs $(c, q) \in \mathbf{R}^3 \times \mathbf{R}^3$ through the following requirements:

- the quantity in (2.20) does not change sign over the pairs $t, s \in K,\ s \neq t$;
- whenever there is a unique $a \in K = [a_1, a_2]$ such that

$$(\phi(a_1) + \phi(a_2) - 2\phi(a)) \cdot q = 0, \tag{2.21}$$

then

$$(\phi(a_1) + \phi(a_2) - 2\phi(a)) \cdot c > 0.$$

Once more we can establish the following result.

Proposition 2.3. *Let $\mathcal{M}$ be as in (2.18).*

If $a_1 > 0$, and $(c, q) \in \mathcal{M}_1$, then $(c, q) \in \mathcal{M}$ and Assumption 2.1 holds.

Proof. 1. Suppose first that there is $a \in K$ such that we have (2.21).

Let

$$v_a = \frac{\phi(a_1) + \phi(a_2)}{2}.$$

Consider $v \in \Lambda$ such that

$$[v - \phi(a)] \cdot q = 0.$$

Suppose

$$c \cdot [v - \phi(a)] < 0,$$

and consider the continuous function

$$G(v, u) = c \cdot (v - u)$$

over the bounded path connecting $(v_a, \phi(a))$ and $(v, \phi(a)$ given by

$$S = \{\alpha[(v, \phi(a)) - (v_a, \phi(a))] + (v_a, \phi(a)) : \ \alpha \in [0, 1]\}.$$

It is easy to check that every component of a vector of S is contained in the section

$$\{v \in \Lambda : q \cdot v = q \cdot \phi(a)\}.$$

Then there exists α such that

$$G(\alpha[(v, \phi(a)) - (v_a, \phi(a))] + (v_a, \phi(a))) = 0,$$

or in other words

$$c \cdot [\alpha(v - v_a) + v_a - \phi(a)] = 0,$$

which by Lemma 2.6 means that necessarily

$$\alpha(v - v_a) + v_a = \phi(a).$$

Consequently

$$\alpha(v - \phi(t)) \cdot N(t) + (1 - \alpha)(v_t - \phi(t)) \cdot N(t) = 0$$

and this is in contradiction with Lemma 2.5. Hence

$$c \cdot [v - \phi(a)] > 0 \text{ if } c \cdot [v_a - \phi(a)] > 0.$$

Let $\bar{t}$ be such that

$$q \cdot \phi(a_1) = q \cdot \phi(\bar{t})$$

and $t \neq a$, $t \geq \bar{t}$, such that

$$v_t = \alpha[\phi(a_2) - \phi(a_1)] + \phi(a_1) \in \Lambda$$

verifies

$$[v_t - \phi(t)] \cdot q = 0$$

Considering once more the continuous function $G(v, u)$ over the path connecting $(v_t, \phi(t))$ and $(v_a, \phi(a))$, as

$$\alpha[v_t - v_a] + v_a \in \{v : q \cdot v = q \cdot \phi(t)\},$$

we can, as we did above, conclude that if

$$c \cdot [v_t - \phi(t)] < 0$$

then for certain α,

$$\alpha[v_t - v_a] + v_a = \phi$$

and consequently

$$c \cdot [v_t - \phi(t)] > 0$$

for any $t \geq \bar{t}$. The same type of arguments show that

$$c \cdot [v - \phi(t)] > 0$$

for any v such that

$$q \cdot v = q \cdot \phi(t).$$

If $t < \bar{t}$, there exists $s \in K$ such that

$$q \cdot \phi(s) = q \cdot \phi(t).$$

In this situation, again the continuity of G should be applied to the path connecting

$$(v_{\bar{t}}, \phi(\bar{t})) = (\phi(a_1), \phi(\bar{t}))$$

and

$$(\phi(s), \phi(\bar{t})),$$

repeatedly until the limit case when $\phi(s) = \phi(\bar{t})$.

If there is $\bar{t} \neq a_2$ such that

$$q \cdot \phi(a_2) = q \cdot \phi(\bar{t})$$

we shall proceed in an analogous way.

2. Suppose now that there are $a, b \in K$ such that

$$(v_a - \phi(a)) \cdot q = (v_a - \phi(b)) \cdot q = 0.$$

Then it is not difficult to conclude that

$$a = a_1 \text{ and } b = a_2.$$

Hence assuming (without loss of generality) that

$$(\phi(a_1) - \phi(a_2)) \cdot c > 0$$

we can, once again, use the continuity of G to conclude

$$c \cdot [\phi(s) - \phi(t)] > 0$$

where $\phi(s)$ and $\phi(t)$, verify

$$(\phi(s) - \phi(t)) \cdot q = 0$$

and after that, for a general v such that

$$(v - \phi(b)) \cdot q = 0.$$

□

Remark 2.3. *This type of argument can be also deduced for the case $N = n = 1$, $p = 2$ where it can be seen to be equivalent to the conditions in Lemma 2.1. However when the parameters N, n and p increase their values, it becomes very hard to give geometrically-based sufficient conditions in such an exhaustive manner as we have done here. Even so, in Section 2.8 we show how to give more restrictive yet more general sufficient conditions (Theorem 2.2 and Theorem 2.1) for interesting high dimensional particular situations, where some geometrical ideas can be used as a way to verify Assumption 2.1.*

2.7 Examples

Before going further to higher dimensional situations we gather in this section some typical, academic examples for which either Lemma 2.1, Lemma 2.3, or Proposition 2.3 can be applied.

2.7.1 Example 1

Let us consider the optimal control problem

$$\text{Minimize in } u: \quad \int_0^T [c(x(t))u(t) + u^2(t)]\, dt$$

under

$$x'(t) = q(x(t))u(t) + u^2(t), \quad x(0) = x_0$$

where $|u(t)| \leq 1$.

We have the following remarkable existence result.

Lemma 2.7. *If the functions q and c are Lipschitz, and*

$$q(q - c) > 0.$$

then the optimal control problem admits solution.

The proof reduces to performing some elementary algebra to check the conditions of Lemma 2.1.

Instead of applying that lemma, as both our cost and state-equation functions have cross dependence on x and on u so that we can't apply results in [Ba58], [Ra90], one can try the

classical existence result based on the classical Filippov-Roxin theory. For that we need to check if the orientor field

$$\mathcal{A}_x = \{(\xi, v) : \ v \geq c(x)u + u^2, \ \xi = q(x)u + u^2, \ u \in K = [-1, 1]\}$$

is a convex set. Note that K is bounded so coercivity is not an issue here. Proceeding in that direction, we can see that

$$\xi = q(x)u + u^2$$

is equivalent to

$$u_1 = -\frac{q + \sqrt{q^2 + 4\xi}}{2} \text{ or } u_2 = -\frac{q - \sqrt{q^2 + 4\xi}}{2},$$

which are possible solutions when ξ is such that $\xi \geq -\frac{q^2}{4}$, and at least one of them belongs to $K = [-1, 1]$. Letting

$$F_i(x, \xi) = c(x)u_i + u_i^2, \ i = 1, 2,$$

we see that

$$F_2 \leq F_1,$$

for all ξ as above. Consequently

$$\mathcal{A}_x = \mathcal{A}_x^1 \cup \mathcal{A}_x^2 =$$

$$\{(\xi, v) : \ v \geq F_2(x, \xi), \xi \in u_2^{-1}(K) \cap [-\frac{q^2}{4}, +\infty[\} \bigcup$$

$$\{(\xi, v) : \ v \geq F_1(x, \xi), \xi \in \left(u_1^{-1}(K) \setminus u_2^{-1}(K)\right) \cap [-\frac{q^2}{4}, +\infty[\}$$

where, for $i = 1, 2$, u_i^{-1} refers to the pre-image of the solutions u_i as functions of ξ.

Because of the assumption on (c, q) it is easy to see that $\mathcal{A}_x^2 = \emptyset$, and consequently that the convexity of $\mathcal{A}_x$ reduces to the convexity of the function

$$F_2(\xi) = \frac{q - c}{2}(q - \sqrt{q^2 + 4\xi})$$

over a certain convex set

$$u_2^{-1}(K) \cap [-\frac{q^2}{4}, +\infty[.$$

This can be checked by elementary calculus.

We now turn over the possibility of applying the result in [CFM06] to this example. First, in order to write our problem as a minimum time problem, we need that $c(x)u + u^2$ never changes sign in $\mathbf{R} \times K$ ([Ce83b]). So a first restriction must be imposed. For example, consider $c(.)$ and $q(.)$ such that

$$q(x) > c(x) > 1.$$

The right member of the differential equation of the minimum time problem is given by

$$\mathcal{F}(x,K) = \{\frac{q(x)u+u^2}{c(x)u+u^2} :\ u \in K\}.$$

The result in [CFM06] doesn't ask for the convexity of the set-valued map $\mathcal{F}$, but it requires a linear boundedness in the sense that

$$\exists \alpha,\ \beta \text{ s. t. } \forall x \in \mathbf{R},\ \forall \xi \in \mathcal{F}(x,K) \text{ then}$$

$$\|\xi\| \leq \alpha\|x\| + \beta.$$

It is easy to see that this condition places a real constraint on the relative growth of pairs (c,q), even before verifying the remaining assumptions in [CFM06].

2.7.2 Example 2

Look at the problem

$$\text{Minimize in } u: \quad \int_0^T [c(x(t))u^2(t) + u^3(t)]\,dt$$

under

$$x'(t) = [q(x(t))]u^2(t) + u^3(t), \quad x(0) = x_0$$

where $u(t) \in [a_0, a_1]$, $a_0 > 0$.

Lemma 2.8. *If the functions $q(x)$ and $c(x)$ are Lipschitz,*

$$c(x) < q(x)\ \forall x,$$

and $q(x)$ is always positive, then the optimal control problem admits solutions.

This result comes directly by applying Lemma 2.3 and Theorem 2.3.

Let us see, what we would need to do if, alternatively, we decided to use the classical existence theory.

Like we have seen in the previous example we need to check the convexity of the orientor field

$$\mathcal{A}_x = \{(\xi, v) :\ v \geq c(x)u^2 + u^3,\ \xi = q(x)u^2 + u^3,\ u \in K = [a_0, a_1]\}.$$

In this case, accordingly to the discriminant

$$\Delta = 27\xi^2 - 4\xi q$$

of the equation

$$\xi = q(x)u^2 + u^3$$

we will have from one to three possible real solutions. Consider for each ξ

$$F_i = c(x)u_i^2 + u_i^3$$

such that

$$F_1 \leq F_2 \leq F_3$$

where $u_i = u_i^x(\xi),\ i = 1, 2, 3$ are the three, possible equal, real solutions. Then

$$\mathcal{A}_x = \mathcal{A}_x^1 \cup \mathcal{A}_x^2 \cup \mathcal{A}_x^3 =$$

$$\{(\xi, v) :\ v \geq F_1,\ \xi \in u_1^{-1}(K)\} \bigcup$$

$$\{(\xi, v) :\ v \geq F_2,\ \xi \in u_2^{-1}(K) \setminus u_1^{-1}(K)\} \bigcup$$

$$\{(\xi, v) :\ v \geq F_3,\ \xi \in u_3^{-1}(K) \setminus \left(u_2^{-1}(K) \cup u_1^{-1}(K)\right)\}$$

Checking the convexity of this set, or alternatively, of the function

$$\varphi_x(\xi) = \begin{cases} F_1(\xi) \text{ if } & \xi \in u_1^{-1}(K) \\ F_2(\xi) \text{ if } & \xi \in u_2^{-1}(K) \setminus u_1^{-1}(K) \\ F_3(\xi) \text{ if } & \xi \in u_3^{-1}(K) \setminus \left(u_2^{-1}(K) \cup u_1^{-1}(K)\right) \end{cases}$$

is not an easy task at all, specially when compared to the almost immediate exercise of verifying the conditions of Lemma 2.3. It is also plausible that the inherent difficulties to apply classical theory will increase until a practically impossible scenario when we let N, n and p grow.

2.7.3 Example 3

In order to give an heuristic for using the criteria given in Proposition 2.3 let us consider the previous problem, just by rewriting q as $c - \beta$ and for a specific K.

$$\text{Minimize in } u: \quad \int_0^T [c(x(t))u^2(t) + u^3(t)]\, dt$$

under

$$x'(t) = [c(x(t)) - \beta(x(t))]u^2(t) + u^3(t), \quad x(0) = x_0$$

where $u(t) \in [1, 2]$.

Lemma 2.9. *If the functions β and c are Lipschitz, and*

$$\beta < \min\{0, c\}.$$

then the optimal control problem admits solutions.

Proof. First notice that for $a \in K = [a_1, a_2]$, we can find α such that the vector

$$B = \alpha[\phi(a_2) - \phi(a_1)] + \phi(a_1)$$

verifies

$$[B - \phi(a)] \cdot q = 0.$$

Moreover, it is not difficult to see that

$$\alpha = \frac{a^3 - a_1^3 - m(a^2 - a_1^2)}{a_2^3 - a_1^3 - m(a_2^2 - a_1^2)}$$

and in the projection plane yz, (B_2, B_3) belongs to the line of slope m passing through (a^2, a^3),

$$B_3 - a^3 = m(B_2 - a^2),$$

where

$$B_2 - a^2 = \frac{(a - a_1)[a^2(a_2 + a_1) - a_2^2(a + a_1)]}{a_2^2 + a_1 a_2 + a_1^2 - m(a_2 + a_1)}$$

and $m = -\frac{q_2}{q_3}$.

In our case $K = [1, 2]$, so, because of what we have just seen, taking $a_1 = 1$ and $a_2 = 2$ we see that for $a \in K$, we can find

$$\alpha = \frac{a^3 - ma^2 + m - 1}{7 - 3m} \in [0, 1]$$

such that

$$[\alpha[\phi(a_2) - \phi(a_1)] + \phi(a_1) - \phi(a)] \cdot q = 0,$$

where

$$m = -\frac{c - \beta}{1} = \beta - c < 0.$$

Furthermore, it is easy to see that the equation $\alpha = \frac{1}{2}$ has a unique solution in K. Consequently, if we consider $q = (0, c - \beta, 1)$ and $\bar{c} = (0, c, 1)$, there exists a unique $a \in K$ such that

$$[\phi(a_2) - \phi(a_1) - 2\phi(a)] \cdot q = 0.$$

Also, because of what we have seen above

$$\begin{aligned}[\frac{1}{2}(\phi(a_2) - \phi(a_1)) - \phi(a)] \cdot \bar{c} \ &= (B_2 - a^2)c + (B_3 - a^3) = (B_2 - a^2)(c + m) \\ &= \tfrac{(a-1)(3a^2-4a-4)}{7-3m}(c + \beta - c) > 0.\end{aligned}$$

In addition, given $t, s \in K,\ s \neq t$,

$$(\phi'(t) \times (c \times q)) \cdot (\phi(s) - \phi(t)) = 0 \Leftrightarrow$$

$$\beta t(0, 3t, -2) \cdot (\phi(s) - \phi(t)) = 0 \Leftrightarrow$$

$$(s - t)[3t(s + t) - 2(s^2 + st + t^2)] = 0 \Leftrightarrow$$

$$s = -\frac{t}{2} \vee s = t$$

which is impossible since $s \in K = [1, 2]$ and $s \neq t$. The result follows then by applying Proposition 2.3.

□

2.8 The case N, $n > 1$

The previous analysis makes it very clear that checking Assumption 2.1 may be a very hard task as soon as n and/or N become greater than 1. Yet in this section we would like to show that there are chances to prove some non-trivial results.

The three main ingredients in Assumption 2.1 are:

- the vector $c \in \mathbf{R}^s$ in the cost functional;
- the matrix $Q \in \mathbf{M}^{N \times s}$ occurring in the state equation;
- the convexification Λ of the set of moments L.

For (c, Q) given, consider the set $\mathcal{N}(c, Q)$ as it was defined in (2.6). Let Ψ be as in (2.5) and such that $\nabla\Psi(m)$ is a rank $s - n$ matrix and L can be seen as the embedded (parametrized) manifold of $\mathbf{R}^s$ in the manifold defined implicitly by $\Psi = 0$ ([KB76]). This means that $\Psi(\phi(u)) = 0$ for all $u \in K$.

Consider also the set of vectors $\mathcal{N}(K, \phi)$ described in (2.7), that is, the set of "ascent" directions for Ψ at points of L.

We are now in conditions to prove Theorem 2.1.

Proof. The proof is rather straightforward. Firstly, note that due to the convexity assumption on Ψ, and the fact that $L \subset \{\Psi = 0\}$, we have $\Lambda \subset \{\Psi \leq 0\}$.

Suppose that $m_0 \in L$ and $m_1 \in \Lambda$, so that

$$\Psi(m_0) = 0,\ \Psi(m_1) \leq 0,\ cm_1 \leq cm_0, \text{ and } Qm_1 = Qm_0\ (= \xi).$$

Then it is obvious that $m = m_1 - m_0 \in \mathcal{N}(c, Q)$. Because of our assumption, $m \in \mathcal{N}(K, \phi)$. We have two possibilities:

1. $\nabla\Psi(m_0)m = 0$. Because of the convexity of each component of Ψ, we have

$$\Psi(m_1) - \Psi(m_0) - \nabla\Psi(m_0)m \geq 0.$$

But then

$$0 = \Psi(m_0) \leq \Psi(m_1) \leq 0,$$

so that $m_1 \in L$. Because of the strict convexity of each component of Ψ, this means that $m_1 = m_0$, and Assumption 2.1 holds.

2. $\nabla\psi_i(m_0)m > 0$ for some i. Once again we have

$$\psi_i(m_1) - \psi_i(m_0) - \nabla\psi_i(m_0)m \geq 0.$$

But this is impossible because $\psi_i(m_1) > 0$ cannot happen for a vector in Λ.

□

Remark 2.4. *Notice that if in the original problem (P_1) we would have considered the dynamics given by*

$$Q(x)\phi(u) + Q_0(x)$$

instead of just $Q(x)$, Assumption 2.1 and Theorem 2.1 could be written exactly in the same way.

Though Theorem 2.1 can be applied to more general cases, we will focus on a particular situation motivated by the control of underwater vehicles ([Br94]). We will briefly describe the structure of the state equation. Indeed, it is just

$$x'(t) = Q_1(x)\phi(u) + Q_0(x)$$

where the state $x \in \mathbf{R}^{12}$ incorporates the position and orientation in body and world coordinates, and the control $u \in \mathbf{R}^{10}$ accounts for guidance and propulsion. Under suitable simplifying assumptions ([Br94], [Fs94], [HL93]), the components of the control vector u only occur as either linear or pure squares, in such a way that $\phi(u) = (u, u^2) \in \mathbf{R}^{20}$, and $u^2 = (u_i^2)_i$, componentwise. Q_1 and Q_0 are matrices which may have essentially any kind of dependence on the state x.

To cover this sort of situations just described, we will concentrate on the optimal control problem (P) already stated in (2.10)-(2.12), and set D, E and U as in (2.13)-(2.14).

We can now prove Theorem 2.2.

Proof. Notice that accordingly to (2.9), as $s = 2n$, we have, for $m \in \mathbf{R}^s$,

$$\psi_i(m) = m_i^2 - m_{n+i}, \quad i = 1, 2, \ldots, n,$$

which are certainly smooth and (strictly) convex. Moreover,

$$\nabla\Psi(m) = \begin{pmatrix} 2\tilde{m} & -\mathrm{id} \end{pmatrix}$$

where

$$\tilde{m} = 2\sum_i m_i e_i \otimes e_i,$$

and e_i is the canonical basis of $\mathbf{R}^n$.

Suppose we have, for a vector $v \in \mathbf{R}^{2n}$, $v = (v_1, v_2)$, that

$$Qv = 0, \quad cv \leq 0.$$

A more explicit way of writing this is

$$Q_1 v_1 + Q_2 v_2 = 0, \quad c_1 v_1 + c_2 v_2 \leq 0.$$

So

$$v_1 = Dv_2, \quad Ev_2 \leq 0.$$

We have to check that such a vector v is not a direction of descent for every function ψ_j, or it is an ascent direction for at least one of them. Note that

$$\nabla\Psi(m)v = Uv_2, \quad Ev_2 \leq 0.$$

It is an elementary Linear Algebra exercise to check that if $U^{-T}E < 0$, then condition (2.8) is verified so that Theorem 2.1 can be applied. □

Corollary 2.1 is a specific example of the kind of existence result that can be obtained through this approach. Its proof amounts to going carefully through the arithmetic while checking that matrix U and vector E defined from such given class of $(c(.), Q(.))$ verify the assumptions of Theorem 2.2.

By using the same ideas, more general situations can be treated, for example the number of controls could be greater than the components of the state. This is in fact the situation in the model that has served as an inspiration for us. We will pursue a closer analysis of such a particular situation, even stressing the more practical issues, in a forthcoming work.

Chapter 3

Some results in scalar problems

3.1 Introduction

We place ourselves in the context of a typical optimal control problem

$$\text{(P)} \quad \text{Minimize the functional} \quad I(y,u) = \int_0^T F(y(t),u(t))dt$$

with the state $y(t)$ and the control $u(t)$ subject to the state equation

$$y'(t) = f(y(t),u(t)),\ a.e.\ t \in (0,T), \tag{3.1}$$

the initial and/or final conditions

$$y(0) = y_0,\ y(T) = y_T, \tag{3.2}$$

and the viability constraints

$$u(t) \in K,\ y(t) \in L, \tag{3.3}$$

where

$$f : \mathbf{R}^N \times \mathbf{R}^n \longrightarrow \mathbf{R}^N,\ F : \mathbf{R}^N \times \mathbf{R}^n \longrightarrow \mathbf{R},\ K \subseteq \mathbf{R}^n, \quad L \subseteq \mathbf{R}^N.$$

Appropriate typical technical hypotheses are assumed on F and f so that, for instance, the state equation admits absolutely continuous (AC for short) solutions, and the state y can be recovered (in a unique way) from the control u. To avoid more technical complications, we will assume that the set of admissibility of controls is constant throughout time as well as the admissibility set for the state. After all, our main contribution here is for autonomous systems so that we will not assume such time dependence from the start.

As we will see, this optimal control problem is equivalent to the variational problem

$$\text{(VP)} \quad \text{Minimize the functional} \quad J(y) = \int_0^T \varphi(y(t),y'(t))dt$$

with the state $y(t)$ subject to the same initial and final states $y(0) = y_0$, $y(T) = y_T$, where the integrand φ is defined by putting

$$\varphi(y, \xi) = \inf_u \{F(y, u) : \xi = f(y, u), u \in K\} \tag{3.4}$$

if $y \in L$, and $\varphi = +\infty$ if $y \notin L$ where

$$\varphi : \mathbf{R}^N \times \mathbf{R}^N \longrightarrow \mathbf{R} \cup \{+\infty\}.$$

Notice, in addition, that if the set $\{u \in K \ : \ \xi = f(y, u)\}$ turns out to be empty, then $\varphi(y, \xi) = +\infty$. If further assumptions are verified by F and f, this integrand φ is well-defined (in the sense that the infimum in (3.4) is a minimum) and so is the associated variational problem. We will be more precise about this later.

The whole point is that whenever we can show existence of optimal solutions for the variational problem determined by φ, we will also have optimal solutions for the original optimal control problem. Indeed, a general situation would be to apply the classical Tonelli's theorem [To14] in its modern version [BGH98] to (VP), and see how coercivity and convexity for φ translate into explicit properties for F and f. This is classical and corresponds to standard existence theorems under convexity ([Ce83b]).

What is more interesting is to apply finer existence theorems for variational problems to our situation, and translate them into the ingredients of the optimal control problem (P). Our main contribution here is to explore one such simple, but non-trivial, situation corresponding to autonomous, one-dimensional problems. The most general such theorem, as far as we know, can be found in [Or03], [Or07], and reads as follows. In those papers, convexity is reduced to a minimum as it is only required at the origin, when considering the usual convexificacion of a function given by the bipolar function. To be specific, for a function $f : X \to \bar{\mathbf{R}} = \mathbf{R} \cup \{+\infty\}$, the bipolar function f^{**} is defined [ET74b] by

$$f^{**}(x) = \sup\{A(x) : A(.) \text{ is an affine function, } A(v) \leq f(v), \ \forall v \in X\}.$$

When we consider y as a fixed parameter, the bipolar function of φ with respect to the variable ξ is given by

$$\varphi^{**}(y, \xi) = \sup\{A(\xi) : A(.) \text{ is an affine function, } A(v) \leq \varphi(y, v), \ \forall v \in \mathbf{R}\}.$$

As just indicated, y is taken here as a parameter.

Requiring the convexity at the origin means that we should have

$$\varphi^{**}(y, 0) = \varphi(y, 0), \ \forall y.$$

By applying such a result to our problem (VP) when $n = 1$ (although dimension m could be larger than 1), and reinterpreting appropriately the convexity at the origin, we prove our main contribution in this chapter, Theorem 3.3, below. Proceed to Section 3.3 for a rigorous precise statement and proof. Here we discuss this result, more informally.

For the existence of an optimal solution for problem (P), we must ask for f to be continuous, F lower semicontinuous, coercive with respect to K, bounded from below, and verifying the growth condition

$$\lim_{|f(y,u)|\to\infty,\ u\in K} F(y,u)/|f(y,u)| = +\infty, \text{ uniformly in } y.$$

In addition, typical convexity conditions required in classical results for existence are based on variational reformulations ([Ce83b] or [P03]), and amount to the convexity of the epigraph of φ with respect to ξ. In our result, we replace it by the condition

$$\sup_{u\in K,\ f(y,u)<0} \frac{F(y,u)-m(y)}{f(y,u)} \le \inf_{u\in K,\ f(y,u)>0} \frac{F(y,u)-m(y)}{f(y,u)}$$

where

$$m(y) = \min_{u\in K}\{F(y,u) : f(y,u) = 0\},$$

which is in fact a necessary and sufficient condition for the convexity at the origin of φ (with respect to ξ).

We will also see (Corollary 3.1 and Corollary 3.2, below) that when K is bounded, the simple condition

$$\min_u\{F(y,u) :\ f(y,u) = 0,\ u \in K\} = \min_u\{F(y,u) :\ u \in K\},$$

will be sufficient to have the existence of optimal solutions.

What is interesting about these results is that we need not care about convexity in any way, and particularly this last condition (Corollary 3.2) can be applied very easily to many situations.

We have organized all of the material in another three sections. Section 3.2 is concerned with some elementary ideas concerning the proof of the equivalence of (P) and (VP) used to specify main technical assumptions. Section 3.3 is the main part of the chapter and there we prove Theorem 3.3. Finally, Section 3.4 includes several (academic) examples where these results are applied.

It is true that the one-dimensional situation treated here is a very special case and that the higher dimensional situation is much more complicated. Even so, we would like to address it in the future.

3.2 Variational Reformulation

We start our analysis by introducing a new format for the Lagrange Optimal Control Problem (P). This approach has been recently examined in different papers, for instance, in [BP01] and [P03], where a variational reformulation was undertaken for the situation where only boundary conditions are assumed on the state. The process we will follow here is the same. The viability constraints (3.3) can also be incorporated in this approach. Additional integral constraints in the form of inequalities can be treated too, but for the sake of simplicity we will not take these into account.

We define the integrand for our reformulation by putting

$$\varphi(y,\xi) = \inf_{u\in K}\{F(y,u) : \xi = f(y,u), y \in L\}$$

which is in fact

$$\varphi(y,\xi) = \begin{cases} \inf\limits_{u\in K}\{F(y,u) : \xi = f(y,u)\}, & \text{if } y \in L \text{ and } \{u \in K \ : \xi = f(y,u)\} \neq \emptyset, \\ +\infty, & \text{else.} \end{cases}$$

Consider the variational problem (VP)

$$\text{Minimize in } y: \quad J(y) = \int_0^T \varphi(y(t), y'(t))dt$$

with

$$\varphi : \mathbf{R}^N \times \mathbf{R}^N \longrightarrow \mathbf{R} \cup \{+\infty\},$$

and $y(0) = y_0$ and $y(T) = y_T$.

We can see that, under certain mild hypotheses, φ is well-defined, in the sense that $\varphi > -\infty$ always, and the infimum is attained.

Lemma 3.1. *Assume that:*

i) K is closed and, if unbounded, F should be coercive with respect to K in the sense that if

$$U(y,\xi) = \{u \in K : \xi = f(y,u), y \in L\}$$

then

$$\lim_{u\in K,\ |u|\to\infty} F(y,u) = +\infty, \tag{3.5}$$

uniformly in y inside $U(y,\xi)$, which means that

$$\forall M > 0, \quad \forall R > 0, \quad \exists P_{M,R}, \quad \forall u, y, \xi, \text{ such that}$$

$$(|u| > P_{M,R}, \quad ||(y,\xi)|| \leq R, \quad u \in \{u \in K : \xi = f(y,u)\} \text{ then } F(y,u) > M);$$

ii) $F : \mathbf{R}^N \times \mathbf{R}^n \longrightarrow \mathbf{R}$ is lower semi-continuous, and $f : \mathbf{R}^N \times \mathbf{R}^n \longrightarrow \mathbf{R}$ is continuous.

Then $\varphi : \mathbf{R}^N \times \mathbf{R}^N \longrightarrow \mathbf{R} \cup \{+\infty\}$ given above is well-defined.

Proof. Notice that (3.5) can also be interpreted in the following fashion:

$$\forall M > 0, \quad \forall R > 0, \quad \exists P_{M,R}, \quad \forall u, y, \xi, \text{ such that}$$

$$(\quad ||(y,\xi)|| \leq R, \quad u \in \{u \in K : \xi = f(y,u)\}, \quad F(y,u) \leq M, \text{ then } |u| \leq P_{M,R}).$$

Using this fact, if we consider, by contradiction, a minimizing sequence u_n in K such that for some ξ and y, we have $\xi = f(y, u_n)$ and $F(y, u_n)$ converging to $-\infty$, then $\{u_n\}$ is a bounded sequence, and so it has a subsequence u_{n_k} converging to certain $\tilde{u}$ such that, due to the lower semi-continuity of F,

$$-\infty = \liminf_n F(y, u_{n_k}) \geq F(y, \tilde{u}) > -\infty,$$

a contradiction. Therefore $\varphi(y, \xi) > -\infty$.

In a similar manner, by using the continuity of f, it can be easily argued that the infimum defining φ is a minimum. □

We can now establish the next result which states the equivalence between problems (P) and (VP).

Theorem 3.1. *Problem (P) is equivalent to problem (VP) in the sense that the infimum of I is the same as that of J.*

Moreover, if (y, u) is an optimal solution for (P), then y is an optimal solution for (VP). Conversely, if y is optimal for (VP), then there is a measurable admissible u such that (y, u) is optimal for (P).

Proof. Step 1 - Let (y, u) be an admissible pair for (P), then $F(y(t), u(t)) < +\infty$ a.e. t in $[0, T]$. By the definition of φ

$$\int_0^T \varphi(y(t), y'(t))dt \leq \int_0^T F(y(t), u(t))dt < +\infty,$$

and so y is admissible for (VP). Moreover

$$J(y) \leq I(y, u).$$

Step 2 - If y is an admissible solution of (VP), then $y \in AC$, and

$$\varphi(y(.), y'(.)) \leq +\infty \text{ a.e. } t \text{ in } [0, T].$$

By definition of φ, $y(t) \in L$ a.e. t in $[0, T]$, and

$$\{u \in K \ : y'(t) = f(y(t), u)\} \neq \emptyset \text{ a.e. } t \text{ in } [0, T].$$

However, for a.e. t in $[0,T]$, there is $u(t) \in K$ such that

$$\varphi(y(t), y'(t)) \leq F(y(t), u(t)) < +\infty,$$

and $y'(t) = f(y(t), u(t))$. Consequently, the set-valued function

$$M(t) = \{u \in K \ : \varphi(y(t), y'(t)) = F(y(t), u), \quad y'(t) = f(y(t), u)\}$$

has non-empty values for a.e. t in $[0,T]$. In addition, $M(t)$ is closed. To this end, consider $\{u_n\} \subset M(t)$ such that $u_n \to \bar{u}$. Then

$$\varphi(y(t), y'(t)) = F(y(t), u_n),$$

$$y'(t) = f(y(t), u_n), \quad \forall n \in \mathbb{N},$$

and since F is lower semicontinuous and f is continuous,

$$F(y(t), \bar{u}) \leq \liminf_n F(y(t), u_n) = \varphi(y(t), y'(t)),$$

$$y'(t) = f(y'(t), \bar{u}), \quad \forall n \in \mathbb{N}.$$

But by definition of φ we have

$$\varphi(y(t), y'(t)) \leq F(y(t), \bar{u}) \leq \varphi(y(t), y'(t)),$$

and so $\bar{u} \in M(t)$.

In this situation we can apply to $M(t)$ a typical selection theorem ([AF90b] or Appendix A) to conclude that there is a measurable selection $u : [0,T] \to K \subseteq \mathbf{R}^n$ such that

$$\varphi(y(t), y'(t)) = F(y(t), u(t)), \ y'(t) = f(y(t), u(t)) \text{ a.e. } t \text{ in } [0,T]$$

and

$$\int_0^T \varphi(y(t), y'(t))dt = \int_0^T F(y(t), u(t))dt < +\infty.$$

Hence, (y,u) is admissible for (P) and $I(y,u) = J(y)$.

Step 3 - In particular, if y is optimal for (VP) then there is a measurable u such that (y,u) is optimal for (P) (and consequently $\inf I = \inf J$) since otherwise by Step 1, (y,u) would not be optimal. A similar argument give us the reciprocal. □

3.3 An Existence Theorem In Dimension One

In the particular case when the state y is one-dimensional (the state of the system is a single parameter), then new existence results can be given. One can use recent existence results for (VP) (see [FMO98], [Or03], [Or07]) where the full convexity of the associated integrand φ is changed just to the convexity at the origin as it has been indicated in the Introduction.

Theorem 3.2. *(See [Or07]) Suppose that*
$\varphi : \mathbf{R} \times \mathbf{R} \longrightarrow [0, +\infty]$ *verifies the following conditions:*

(i) φ *is* $\mathcal{L} \otimes \mathbf{B}$*-measurable and* $\varphi(.,.)$ *is lower semicontinuous;*

(ii) $\varphi^{**}(y, 0) = \varphi(y, 0),\ \forall y$*;*

(iii) $\varphi(y, \xi)$ *has superlinear growth in* ξ*, in the sense*

$$\lim_{|\xi| \to +\infty} \frac{\varphi(y, \xi)}{|\xi|} = +\infty, \text{ uniformly in } y.$$

Then, there is a minimizer for

$$J(y) = \int_a^b \varphi(y(t), y'(t))dt$$

in the class

$$C(\alpha, \beta) = \{y \in AC((a, b), \mathbf{R}) : y(a) = \alpha, y(b) = \beta\}.$$

Here, $\mathcal{L} \otimes \mathbf{B}$-measurable means measurable with respect to $\mathcal{L} \otimes \mathbf{B}$, the smallest σ-algebra of subsets of $\mathbf{R} \times \mathbf{R}$ that contains all the product sets $A \times B$ where A is a Lebesgue set and B a Borel set. We represent by $AC((a, b), \mathbf{R})$ the set of absolutely continuous functions from (a, b) to $\mathbf{R}$.

Theorem 3.2 is a general remarkable result, independent of our application here. Its scope is thus beyond our contribution. It is somehow the final stage of a series of improvements of existence results for scalar, one-dimensional problems. See ([Or07]) for the proof, and related references.

To apply this result to (VP), we must derive conditions to ensure $\varphi^{**}(y, 0) = \varphi(y, 0)$, for every y.

Let us consider the following fact.

Proposition 3.1. *For a real function,* $\varphi : \mathbf{R} \longrightarrow [0, +\infty]$*, the condition*

$$\varphi^{**}(0) = \varphi(0) < +\infty$$

is equivalent to

$$\sup_{\xi<0} \frac{\varphi(\xi) - \varphi(0)}{\xi} \leq \inf_{\xi>0} \frac{\varphi(\xi) - \varphi(0)}{\xi}\ . \tag{3.6}$$

Proof. Put $m^- \leq m^+$ to designate the slopes on the left hand and right hand sides of (3.6), respectively. Notice that we have $m^- = -\infty$ or $m^+ = +\infty$ if and only if $\varphi(\xi) = +\infty$ for all $\xi < 0$ or $\varphi(\xi) = +\infty$ for all $\xi > 0$ respectively. It is easy to see that

$$\forall\ \xi > 0, \quad m^-\xi + \varphi(0) \leq m^+\xi + \varphi(0) \leq \varphi(\xi)$$

and

$$\forall\ \xi < 0,\quad m^+\xi + \varphi(0) \le m^-\xi + \varphi(0) \le \varphi(\xi).$$

Using the definition of φ^{**}, recalled in the Introduction, we can conclude

$$\varphi^{**}(0) = \varphi(0).$$

Conversely, suppose we have an affine function

$$a(\xi) = m(\xi) + \varphi(0)$$

such that

$$m\xi + \varphi(0) \le \varphi(\xi),\ \forall\ \xi.$$

By setting again m^- and m^+ as above, it is easy to check that

$$m^- \le m \le m^+,$$

and consequently

$$\sup_{\xi<0} \frac{\varphi(\xi) - \varphi(0)}{\xi} \le \inf_{\xi>0} \frac{\varphi(\xi) - \varphi(0)}{\xi}.$$

□

We also have a parallel result when $\varphi(0)$ attains the value $+\infty$.

Proposition 3.2. *If* $\varphi(0) = +\infty$ *then*

$$\varphi^{**}(0) = \varphi(0) \Leftrightarrow$$

$$(\varphi(\xi) = +\infty,\ \forall\xi < 0) \quad or\ \ (\varphi(\xi) = +\infty,\ \forall\xi > 0)\,.$$

Proof. We shall prove the non trivial implication ($\Leftarrow$).

Consider, without lost of generality, that

$$\varphi(\xi) < +\infty$$

for certain $\xi < 0$ (consequently $\varphi(\xi) = +\infty\ \forall\xi > 0$). Let

$$\hat{\xi} = \max\{\xi : \varphi(\xi) < +\infty\}.$$

Then

$$\varphi^{**}(0) = \sup\{A(0) : A(\xi) < \varphi(\xi),\ \forall\xi\} =$$

$$= \sup\{A(0) : A(\hat{\xi}) = \varphi(\hat{\xi})\} = +\infty.$$

□

To connect these results with our situation, we have the following lemma.

Lemma 3.2. *Let $\varphi : \mathbf{R} \times \mathbf{R} \longrightarrow [0, +\infty]$ be the new density function built using F and f as described in Section 3.2. Then the fact*

$$\varphi^{**}(y, 0) = \varphi(y, 0) \ \text{ for all } y$$

is equivalent to

$$\sup_{u \in K,\ f(y,u)<0} \frac{F(y,u) - m(y)}{f(y,u)} \leq \inf_{u \in K,\ f(y,u)>0} \frac{F(y,u) - m(y)}{f(y,u)}, \tag{3.7}$$

where $m(y) := \min_{u \in K}\{F(y,u) : f(y,u) = 0\}$, if $\varphi(y,0) < +\infty$; or to

$$f(y,u) < 0, \ \forall u \in K$$

or

$$f(y,u) > 0, \ \forall u \in K$$

when $\varphi(y,0) = +\infty$.

Proof. The case $y \notin L$ is immediate since we will have $\varphi(y,\xi) = +\infty$ everywhere.

Let $y \in L$ and $m(y)$ be as indicated above. We realize that in fact $m(y) = \varphi(y,0)$, and hence

$$\begin{aligned}
&\max_{\xi<0} \frac{\varphi(y,\xi) - m(y)}{\xi} \\
&= \max_{\xi<0} \frac{1}{\xi} \left(\min_{u \in K,\ f(y,u)=\xi} (F(y,u) - m(y)) \right) \\
&= \max_{\xi<0} \left(\max_{u \in K,\ f(y,u)=\xi} \frac{F(y,u) - m(y)}{f(y,u)} \right) \\
&= \max_{u \in K,\ f(y,u)<0} \frac{F(y,u) - m(y)}{f(y,u)}
\end{aligned}$$

and

$$\begin{aligned}
&\min_{\xi>0} \frac{\varphi(y,\xi) - m(y)}{\xi} \\
&= \min_{\xi>0} \frac{1}{\xi} \left(\min_{u \in K,\ f(y,u)=\xi} (F(y,u) - m(y)) \right) \\
&= \min_{\xi<0} \left(\min_{u \in K,\ f(y,u)=\xi} \frac{F(y,u) - m(y)}{f(y,u)} \right) \\
&= \min_{u \in K,\ f(y,u)>0} \frac{F(y,u) - m(y)}{f(y,u)},
\end{aligned}$$

so the first part of the lemma comes directly from an application of Proposition 3.1.

The second part of the lemma is a direct application of the definition of φ in Section 3.2 and Proposition 3.2. □

The following result will also be helpful.

Proposition 3.3. (a) *Let φ be as in Proposition 3.2. If φ has a global minimum at $\xi = 0$, then $\varphi^{**}(0) = \varphi(0)$.*

(b) *Let φ be as in Theorem 3.2. If $\varphi(y,.)$ has a global minimum at $\xi = 0$ for every y, then $\varphi^{**}(y,0) = \varphi(y,0)$.*

Proof. If φ has a global minimum at $\xi = 0$ then $\varphi(0) \leq \varphi(\xi)\ \forall\ \xi$, and so

$$\xi < 0 \Rightarrow \frac{\varphi(\xi) - \varphi(0)}{\xi} \leq 0,$$

$$\xi > 0 \Rightarrow \frac{\varphi(\xi) - \varphi(0)}{\xi} \geq 0,$$

and, by Proposition 3.1, (a) follows. When we consider y fixed and $\varphi(0) := \varphi(y,0)$ like in (a), then we have (b). □

These conclusions lead us to our main result.

Theorem 3.3. *Let F, f, K and L be as above. Suppose that F is lower semi-continuous, f is continuous, K and $L \subseteq \mathbf{R}$ are closed sets. Suppose that the following conditions are verified:*

(i) *F is coercive with respect to K in the sense*

$$\lim_{|u|\to\infty,\ u\in K} F(y,u) = +\infty, \qquad \textit{uniformly in } y;$$

(ii) *For $y \in L$, put $m(y) = \min_{u\in K}\{F(y,u) : f(y,u) = 0\}$. Then*

$$\sup_{u\in K,\ f(y,u)<0} \frac{F(y,u) - m(y)}{f(y,u)} \leq \inf_{u\in K,\ f(y,u)>0} \frac{F(y,u) - m(y)}{f(y,u)};$$

(iii) *F is bounded from below, and*

$$\lim_{|f(y,u)|\to\infty,\ u\in K} F(y,u)/|f(y,u)| = +\infty,\ \textit{uniformly in } y\ .$$

Then, there is an absolutely continuous function y, and a measurable function u, such that the pair (y,u) is an optimal solution for (P).

Proof. Our strategy is to apply Theorem 3.2 to the equivalent variational problem (VP). First, let us see that the integrand function φ defined as above verifies condition (i) of Theorem 3.2. It will be enough to check if for all $\alpha \in \mathbf{R}$ the set

$$D_\alpha = \{(y,\xi) : \varphi(y,\xi) \leq \alpha\}$$

$$= \{(y,\xi) : y \in L, \exists u \in K, F(y,u) \leq \alpha, \xi = f(y,u)\}$$

is closed (see [Ce83b] or [Ru87b]).

Consider

$$(y_n, \xi_n)_n \subseteq D_\alpha \text{ such that } (y_n, \xi_n) \to (\bar{y}, \bar{\xi}).$$

Then, $\bar{y} \in L$ and $\exists u_n \in K : \quad F(y_n, u_n) \leq \alpha$ and $\xi_n = f(y_n, u_n)$. By the coercivity of F (condition (i)) $(u_n)_n$ is bounded and so, it has a convergent subsequence to a certain $\bar{u} \in K$. Since F is lower semicontinuous, and f is continuous,

$$F(\bar{y}, \bar{u}) \leq \alpha$$

and

$$\bar{\xi} = f(\bar{y}, \bar{u}), \quad \bar{u} \in K,$$

so $(\bar{y}, \bar{\xi}) \in D_\alpha$.

Condition (ii) in our statement comes directly from Lemma 3.2. The convexity condition (ii) of Theorem 3.2 follows immediately. Notice that if condition (ii) in our statement is void, when we have

$$f(y, u) < 0, \ \forall u \in K$$

or

$$f(y, u) > 0, \ \forall u \in K$$

by the second part of Lemma 3.2, the convexity at the origin still follows.

Finally, to check that φ has superlinear growth, consider the growth condition (iii) on F. We have in particular that

$$\forall M > 0, \ \exists L > 0, \ \forall u \in K, \ \forall \xi : \ |\xi| = |f(y, u)| > L, \ \varphi(y, \xi) = F(y, u)$$

$$\Rightarrow \frac{F(y, u)}{|f(y, u)|} = \frac{\varphi(y, \xi)}{|\xi|} > M,$$

and so

$$\lim_{|\xi| \to +\infty} \frac{\varphi(y, \xi)}{|\xi|} = +\infty.$$

Since F is bounded from below so is φ, and it is easy to see that φ has superlinear growth in the sense of condition (iii) of Theorem 3.2. We are now in a position to apply Theorem 3.2 to our reformulated problem (VP), and conclude. □

If K is bounded, conditions (i) and (iii) are void, so we can also state the following result.

Corollary 3.1. *Assume that K is bounded, and the condition*

$$\sup_{u \in K, \ f(y,u) < 0} \frac{F(y, u) - m(y)}{f(y, u)} \leq \inf_{u \in K, \ f(y,u) > 0} \frac{F(y, u) - m(y)}{f(y, u)}$$

is verified for $y \in L$, where $m(y)$ is as in Theorem 3.3. Then we have existence of optimal solutions for our original optimal control problem.

In this situation, if we apply Proposition 3.3 (b), and consider the definition of $\varphi(y,0)$, we also have the useful result that follows.

Corollary 3.2. *Let F, f, K and L as indicated above. If K is bounded, and for all $y \in L$ we have*

$$\min_u\{F(y,u) : \ f(y,u)=0, \ u\in K\} = \min_u\{F(y,u) : \ u\in K\},$$

then, the associated optimal control problem admits optimal solutions.

3.4 Examples

We will next study some examples to illustrate our results.

Example 3.1. *Consider the problem*

$$\textit{Minimize} \quad I(y,u) = \int_0^1 |u(t)|^2 dt$$

subject to

$$y'(t) + y(t) = u(t)$$

and

$$|u(t)| \le 1, \ \forall t \in [0,1],$$

with boundary conditions

$$y(0) = y_0 \ \textit{and} \ y(1) = y_1.$$

This is an easy example where standard results can be applied since we have a convex function of u in the cost, a linear dependence of the state equation on u, and a convex set for the constraints on u.

If we define

$$\varphi(y,\xi) = \begin{cases} \min\limits_{u\in[-1,1]}\{u^2 : \xi = -y+u\} & \textit{if } \{u\in[-1,1] \ : \xi = -y+u\} \neq \emptyset \\ +\infty & \textit{else} \end{cases}$$

$$= \begin{cases} (\xi+y)^2 & \textit{if } -1 \le \xi + y \le 1 \\ +\infty & \textit{else} \end{cases}$$

we obtain the reformulated problem:

$$\textit{Minimize} \quad J(y) = \int_0^1 \varphi(y(t), y'(t))dt$$

with

$$y(0) = y_0 \text{ and } y(1) = y_1.$$

To ensure the existence we must check the conditions of Corollary 3.1, since $K = [-1, 1]$ *is compact. We have*

$$F(y, u) = u^2,\ f(y, u) = u - y,\ L = \mathbf{R}.$$

It is easy to see that for $y \in \mathbf{R} \setminus [-1, 1]$ *we have*

$$\varphi^{**}(y, 0) = \varphi(y, 0) = +\infty.$$

The insteresting case corresponds to $y \in]-1, 1[$*. One must have*

$$m(y) = \min_{u\in[-1,1]} \{u^2 : 0 = u - y\} = y^2,$$

and condition iii) becomes

$$\sup_{u\in[-1,1],\ u-y<0} \frac{u^2 - y^2}{u - y} \leq \inf_{u\in[-1,1],\ u-y>0} \frac{u^2 - y^2}{u - y} \Leftrightarrow$$

$$\sup_{-1\leq u<y} \{u + y\} \leq \inf_{y<u\leq 1} \{u + y\} \Leftrightarrow$$

$$2y \leq 2y$$

which is obviously true, and the existence of an optimal pair (y, u) *is ensured.*

Example 3.2. *Let us look at the problem*

$$\textit{Minimize} \quad I(y, u) = \int_0^1 u(t)^4 dt$$

subject to

$$y'(t) + u(t)y(t) = u^2(t)$$

and

$$|u(t)| \leq 1,\ \forall t \in [0, T],$$

with initial conditions

$$y(0) = y_0 \text{ and } y(1) = y_1.$$

The density function is

$$\varphi(y, \xi) = \begin{cases} \min\limits_{u\in[-1,1]} \{u(t)^4 : u(t) = \dfrac{y \pm \sqrt{y^2 + 4\xi}}{2}\} & \text{if } \xi > -\frac{y^2}{4} \\ +\infty & \text{else} \end{cases}$$

$$= \begin{cases} (\frac{y+\sqrt{y^2+4\xi}}{2})^4 & \text{if } -1 \leq \frac{y+\sqrt{y^2+4\xi}}{2} \leq 1 \text{ e } y \leq 0 \text{ e } \xi > -\frac{y^2}{4} \\ (\frac{y-\sqrt{y^2+4\xi}}{2})^4 & \text{if } -1 \leq \frac{y-\sqrt{y^2+4\xi}}{2} \leq 1 \text{ e } y > 0 \text{ e } \xi > -\frac{y^2}{4} \\ +\infty & \text{else} \end{cases}.$$

In spite of this complicated form of φ, if we consider

$$F(y,u) = u^4, \ f(y,u) = u^2 - uy,$$

$$K = [-1,1], \ L = \mathbf{R},$$

we can easily see that

$$\min_{u\in[-1,1]} u^4 = 0 = \min_{u\in[-1,1]} \{u^4 : u = 0 \vee u = y\},$$

and one applies Corollary 3.2 to conclude the existence of optimal solutions. With a bit of careful calculations, φ above can in fact be checked to be convex not only at 0. But to conclude the existence of optimal solutions one need not care about the full convexity, but just apply directly Corollary 3.2.

Example 3.3. *If we consider the problem*

$$\text{Minimize} \quad I(y,u) = \int_0^1 u(t)^2 dt$$

subject to

$$y'(t) = u(t)y(t)$$

and

$$|u(t)| \leq 1, \ 0 \leq y(t) \leq 1 \ \forall t \in [0,T],$$

with initial conditions $y(0) = y_0$ and $y(1) = y_1$, we have

$$F(y,u) = u^2, \ f(y,u) = uy,$$

$$K = [-1,1] \text{ and } \ L = [0,1].$$

Again we can apply Corollary 3.2 by verifying that

$$\min_{u\in[-1,1]} u^2 = 0 = \min_{u\in[-1,1]} \{u^2 : u = 0 \vee y = 0\}.$$

Example 3.4. *Concerning the problem*

$$\text{Minimize} \quad I(y,u) = \int_0^1 u(t)^2 dt$$

subject

$$y'(t) = -u(t)^2 y(t) + u(t)$$

and

$$|u(t)| \leq 1, \ \forall t \in [0,T],$$

with initial conditions $y(0) = y_0$ *and* $y(1) = y_1$*, once more the condition of convexity is easily verified because*

$$F(y,u) = u^2, \ \ f(y,u) = u - u^2 y,$$

$$K = [-1,1] \ \textit{and} \ \ L = \mathbf{R},$$

and

$$\min_{u \in [-1,1]} u^2 = 0 = \min_{u \in [-1,1]} \{u^2 : u = 0 \vee uy = 1\}.$$

Chapter 4

Application to Underwater Vehicles Models

4.1 Introduction

In this chapter we turn over the existence of solution for the model of manoeuvrability control of a submarine which has been recently proposed in [GOP09]. It corresponds to a real-life engineering problem so that all the hypotheses and ingredients that we will consider in the sequel are motivated by real (non-academic) requirements. To describe such model a state vector is defined

$$x = (\mathtt{x}, \mathtt{y}, \mathtt{z}, \phi, \theta, \psi, \mathtt{u}, \mathtt{v}, \mathtt{w}, \mathtt{p}, \mathtt{q}, \mathtt{r}) \in \Omega \subset \mathbf{R}^{12}, \tag{4.1}$$

where $X_{world} = (\mathtt{x}, \mathtt{y}, \mathtt{z}; \phi, \theta, \psi)$ indicates the position and orientation of the submarine in the world fixed coordinate system, and $V_{body} = (\mathtt{u}, \mathtt{v}, \mathtt{w}; \mathtt{p}, \mathtt{q}, \mathtt{r})$ is the vector of linear and angular velocities measured in the body coordinate system. Throughout this chapter we follow the usual SNAME [1] notation [Fs94]. Permitted ranges of Euler angles are

$$-\pi < \phi < \pi, \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}, \quad 0 < \psi < 2\pi, \tag{4.2}$$

so that

$$\Omega = \mathbf{R}^3 \times]-\pi, \pi[\times \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\times]-0, 2\pi[\times \mathbf{R}^6.$$

The control vector is

$$u = (\delta_b, \delta_s, \delta_r)\,, \tag{4.3}$$

where δ_b and δ_s represent, respectively, the angle of the bow and stern coupled planes, and δ_r is deflection of rudder. These controls act on the system in linear and quadratic form. Therefore,

[1] Society of Naval Architects and Marine Engineers

it is convenient to consider the mapping

$$\Phi(u) = \left(u, u^2\right) \equiv \left(\delta_b, \delta_s, \delta_r, \delta_b^2, \delta_s^2, \delta_r^2\right) \in \mathbf{R}^6.$$

Admissible controls u are measurable functions that should lie in a certain set $K \subset \mathbf{R}^3$, which, in our case, is given by

$$K = [-a_1, a_1] \times [-a_2, a_2] \times [-a_3, a_3],$$

with $0 < a_1, a_2, a_3 < \pi/2$. Finally, the state law is described by a system of twelve ordinary differential equations

$$x'(t) = Q(x(t))\,\Phi(u(t)) + Q_0(x(t)) \tag{4.4}$$

where

$$Q : \mathbf{R}^{12} \to \mathcal{M}^{12\times 6} \quad \text{and} \quad Q_0 : \mathbf{R}^{12} \to \mathbf{R}^6$$

will be described in Section 3. At this point, we just indicate that the right-hand side of (4.4) includes both kinematic and dynamic equations of motion (see [Fe79, Fs94, GOP09, GH67] for more details).

The manoeuvrability control problem for an underwater vehicle describes a situation where we want to reach (or to be very close to) a final state x^T in time T, while minimizing the use of control during the time interval $[0, T]$. The latter can be understood as minimizing the typical cost

$$\int_0^T \|u(t)\|^2\, dt$$

while the first aspect can be seen as minimizing

$$\begin{aligned}\frac{1}{2}\left\|x(T) - x^T\right\|^2 &= \frac{1}{2}\int_0^T \frac{d}{dt}\left\|x(t) - x^T\right\|^2 dt + \frac{1}{2}\left\|x(0) - x^T\right\|^2 \\ &= \int_0^T < x(t) - x^T,\ Q(x(t))\,\Phi(u(t)) + Q_0(x(t)) > dt + \frac{1}{2}\left\|x(0) - x^T\right\|^2.\end{aligned}$$

Hence, we consider the cost

$$\begin{aligned}&\int_0^T \left[< x(t) - x^T,\ Q(x(t))\,\Phi(u(t)) + Q_0(x(t)) > + \|u(t)\|^2\right] dt \\ =\ &\int_0^T \left[c(x(t))\,\Phi(u(t)) + c_0(x(t))\right] dt\end{aligned}$$

where the vector c is given by

$$\begin{cases} c_i(x) = \sum_{j=1}^{12} \left(x - x^T\right)_j Q_{ji}, & i = 1, 2, 3, \\ c_i(x) = \sum_{j=1}^{12} \left(x - x^T\right)_j Q_{ji} + 1, & i = 4, 5, 6, \end{cases}$$

and

$$c_0(x) =< x - x^T,\ Q_0(x) > .$$

Typically, some penalty parameters are introduced to weigh at convenience the above two goals, but for simplicity and since it does not change mathematically the problem we have not considered such weights.

To sum up, we can write the manoeuvrability control problem as

$$\text{(UVP)} \quad \begin{cases} \text{Minimize in } u: & \int_0^T \left[c\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right) + c_0\left(x\left(t\right)\right)\right] dt \\ \text{subject to} & \\ & x'\left(t\right) = Q\left(x\left(t\right)\right)\Phi\left(u\left(t\right)\right) + Q_0\left(x\left(t\right)\right), \quad 0 < t < T \\ & x\left(0\right) = x^0 \in \Omega \\ & x\left(t\right) \in \Omega \quad \text{and} \quad u\left(t\right) \in K,\ 0 \leq t \leq T. \end{cases}$$

The main goal of this chapter is to prove the following local existence result.

Theorem 4.1. *For $T > 0$, small enough, there exists an optimal solution of (UVP).*

We notice that the constraint on T is imposed to be able to guarantee that the state law is well-posed. The existence of T will be established during the proof of Theorem 4.1. As we will see later on, the fundamental question for this existence result is the relation between the vector c, the matrix Q, the mapping Φ and the set K. The role played by Q_0 is related to the existence and uniqueness of solution for the state law, and c_0 does not influence at all. To prove Theorem 4.1 we will apply a very recent general existence result [PT09] which requires some modifications to adapt the specific structure of our model. Section 2 is devoted to present this general result (Theorem 2.1) with its corresponding changes. In Section 3 we will check that our model satisfies the hypotheses required by this last theorem.

4.2 A general existence and uniqueness result for some specific optimal control problems

Throughout this section we basically follow the same ideas as in [PT09], but since our problem is slightly different from the one considered there and to make the chapter easier for readers we include detailed statements and proofs.

To study the existence of solution for (UVP) we will turn ourself over the general optimal control problem of the type

$$(CP) \qquad \text{Minimize in } u: \quad \int_0^T c(x) \cdot \Phi(u) + c_0(x) dt \tag{4.5}$$

subject to

$$\mathtt{x}' = Q(x)\Phi(u) + Q_0(x) \tag{4.6}$$

$$x(0) = x_0 \in \mathbf{R}^N,$$

and

$$u(t) \in K, \tag{4.7}$$

where $K \subset \mathbf{R}^m$. We search a control u in $L^\infty((0,T),K)$ corresponding to an absolutely continuous state function $\mathtt{x} : (0,T) \to \mathbf{R}^N$.

The mappings

$$\Phi(u) \in \mathbf{R}^s,$$

$$Q : \mathbf{R}^N \to \mathcal{M}^{N\times s},$$

$$Q_0, c : \mathbf{R}^N \to \mathbf{R}^s$$

should be such that the cost function is defined and takes finite values for admissible pairs $(\mathtt{x}, u)$ and the state system is well-posed.

As we will see, the fundamental question for the existence result is the relation between the vector c, the matrix Q, the application Φ and set K. For a better understanding of such relations we consider additionally a $\mathcal{C}^1$ mapping

$$\Psi : R^s \to R^{s-m}, \quad \Psi = (\psi_1, ..., \psi_{s-m}), \quad (s > m), \tag{4.8}$$

so that $\Phi(K) \subset \{\Psi = 0\}$. This means that we are embedding the image space $\Phi(K)$ into a level surface (submanifold) defined by Ψ. Notice for example that for problem (UVP) where

$$\Phi(u) = (u_1, u_2, u_3, (u_1)^2, (u_2)^2, (u_3)^2) \in \mathbf{R}^6$$

we have

$$\Psi(v) = ((v_1)^2 - v_4, (v_2)^2 - v_5, (v_3)^2 - v_6) \in \mathbf{R}^3.$$

Also we define for every pair (c, Q) the set

$$\mathcal{N}(c,Q) = \{v \in \mathbf{R}^s : Qv = 0, cv \leq 0\}. \tag{4.9}$$

Similarly, we consider

$$\mathcal{N}(K,\Phi) = \tag{4.10}$$

$$\{v \in \mathbf{R}^s : \text{ for each } u \in K, \text{ either } \nabla\Psi(\Phi(u))v = 0 \text{ or } \exists i \text{ s. t. } \nabla\psi_i(\Phi(u))v > 0\},$$

the set of "growth directions" of Ψ over $\Phi(K)$. We are now in conditions to state the existence result proved in [PT09] adapted to our frame.

Theorem 4.2. *Assume that the mapping Ψ as above is component-wise convex and $\mathcal{C}^1$. If for each $x \in \mathbf{R}^n$, we have*

$$\mathcal{N}\left(c\left(x\right), Q\left(x\right)\right) \subset \mathcal{N}\left(K, \Phi\right), \tag{4.11}$$

then the corresponding optimal control problem (CP) has at least one solution. If, in addition, Φ is component-wise one to one, convex and strictly convex for at least one component over K, c and c_0 are constant functions, then the solution of (CP) is unique.

Notice that in the statement of Theorem 2.1, we have dropped the strictly convexity of Ψ as it was asked in [PT09]. Also we have included a sufficient condition which ensures the uniqueness of such a solution. For more general sufficient conditions for uniqueness see [PerT09].

An essential tool to the proof of this result is the verification of the assumption

Assumption 4.1. *For each fixed $x \in \mathbf{R}^N$, and $\xi \in Q(x)\Lambda + Q_0(x) \subset \mathbf{R}^N$, the minimum*

$$\min_{m \in \Lambda} \{c(x) \cdot m + c_0(x) : \xi = Q(x)m + Q_0(x)\}$$

is only attained in L, where $L = \Phi(K)$ and $\Lambda = co(L)$.

In fact this hypothesis has a very simple geometrical meaning, as we show in Figure 4.1 for the simple case were $N = n = 1$, $K = [a_1, a_2]$ and $\Phi(u) = (u, u^2)$. The set L is part of the parameterized curve by Φ and Λ is its convex hull. In this example, the fixed ξ and x are such that $Q_0(x)$ and $c_0(x)$ are both nulls and $\{\xi = Qm\}$ intersects L at the origin, precisely at m_2. It is easy to see that for these x and ξ, the vectors c verifying Assumption 4.1 are those such that

$$c(x) \cdot m_1 > c(x) \cdot m_2 = 0$$

for any m_1 simultaneously at Λ and at $\{\xi = Qm\}$. Such condition is not verified, for example by vector $\bar{c}(x)$ as

$$\bar{c}(x) \cdot m_1 < \bar{c}(x) \cdot m_2 = 0.$$

This assumption allows us to proceed through a relaxation process using Young measures (as in [MP01], [P97b], [PT09], [Rb96] and [Rb97b]) and conclude that there is a Dirac-type solution of the relaxed problem which corresponds to a solution of the original problem.

Before starting the proof of the existence result, let us first consider the following Lemma.

Lemma 4.1. *Let Ψ be as in Theorem 4.2. If c, Q, Φ and K in (CP) are such that condition (4.11) is satisfied, then Assumption 4.1 holds.*

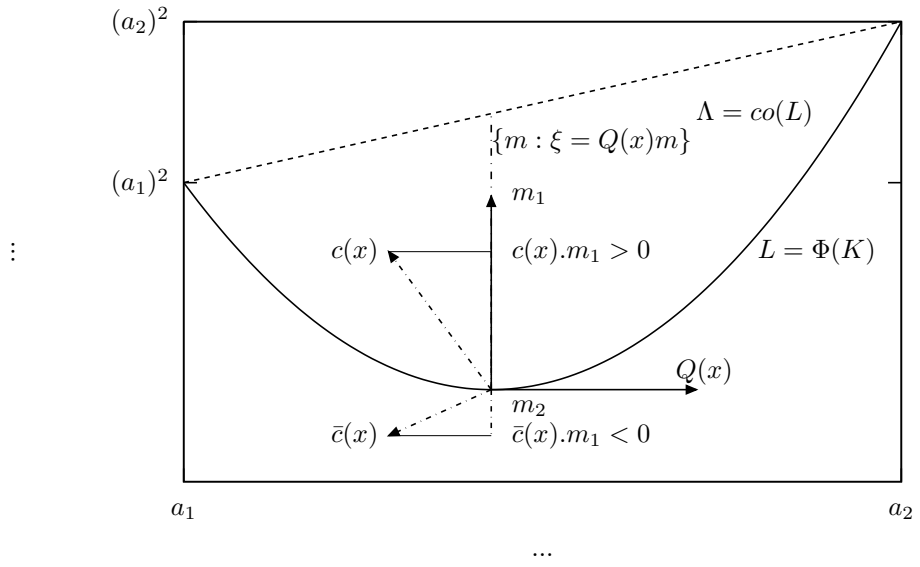


Figure 4.1: Admissible c for $\Phi = (u, u^2)$, $K = [a_1, a_2]$

Proof. We want to see that for every fixed $x \in \mathbf{R}^N$ and $\xi \in Q(x)\Lambda + Q_0(x)$ the minimizer of $c(x) \cdot v + c_0(x)$ over the set of vectors in Λ verifying the restriction $\xi = Q(x)v + Q_0(x)$ can only be in L, where both L and Λ are as in Assumption 4.1.

Suppose that $v_0 \in L$ and $v_1 \in \Lambda$ both belong to the manifold

$$\{\xi = Q(x)v + Q_0(x)\}$$

but they verify

$$c(x)v_1 + c_0(x) \leq c(x)v_0 + c_0(x).$$

As Ψ is component-wise convex and $L \subset \{\Psi = 0\}$, we have $\Lambda = co(L) \subset \{\Psi \leq 0\}$. Hence,

$$\Psi(v_0) = 0,\ \Psi(v_1) \leq 0,\ c \cdot v_1 \leq c \cdot v_0,\ \text{and } Qv_1 = Qv_0\ (= \xi - Q_0).$$

Therefore it is obvious that $v = v_1 - v_0 \in \mathcal{N}(c(x), Q(x))$. Due to condition (4.11), $v \in \mathcal{N}(K, \Phi)$. Accordingly to the definition of $\mathcal{N}(K, \Phi)$ either $\nabla\psi_i(v_0)v > 0$ for some i or $\nabla\Psi(v_0)v = 0$. Suppose we are in the first situation. Because of the convexity of Ψ,

$$\psi_i(v_1) - \psi_i(v_0) - \nabla\psi_i(v_0)v \geq 0 \Leftrightarrow$$

$$\psi_i(v_1) \geq \nabla\psi_i(v_0)v > 0.$$

But this is impossible because $\psi_i(v_1) > 0$ cannot happen for a vector in Λ.

Suppose now that $\nabla\Psi(v_0)v = 0$. Again by convexity of each component of Ψ, we have

$$\Psi(v_1) - \Psi(v_0) - \nabla\Psi(v_0)v \geq 0,$$

that is,

$$0 = \Psi(v_0) \leq \Psi(v_1) \leq 0.$$

Hence, as $v_1 \in \Lambda = (\Lambda \setminus L) \cup L$ and

$$\Lambda \setminus L \subset \{\Psi(v) \leq 0, \quad \exists i \text{ s.t. } , \ \psi_i(v) < 0\}$$

we conclude that $v_1 \in L$ and Assumption 4.1 holds. □

We can now prove Theorem 4.2.

Proof. We begin by the relaxation of (CP) using Young measures associated with sequences of admissible controls. Consider the problem

$$(RP) \quad \text{Minimize in } \mu = \{\mu_t\}_{t\in(0,T)} : \quad \tilde{I}(\mu) = \int_0^T [\int_K c(x(t)) \cdot \Phi(\lambda) d\mu_t(\lambda)] + c_0(x(t))dt$$

subject to

$$\mathtt{x}'(t) = \int_K Q(x(t))\Phi(\lambda)d\mu_t(\lambda) + Q_0(x(t))$$

and

$$\text{supp}(\mu_t) \subset K, \quad x(0) = x_0 \in \mathbf{R}^N.$$

Notice that the theory of Young measures ([MP01], [P97b], [Rb96], [Rb97b]) allows us to conclude that this formulation is, in particular, well posed, as having $u \in L^\infty([0,T], K)$ for K bounded implies (see [P03]) that the associated Young measures $\{\mu_t\}_t$ belongs to

$$\mathcal{Y}^p((0,T), P(K)) =$$

$$\left\{\mu = \{\mu_t\}_{t\in(0,T)} : \ \int_0^T \int_K \|\lambda\|^p d\mu_t(\lambda)dt < \infty, \ \mu_t \in P(K)\right\} \quad \text{for every } p > 1,$$

where $P(K)$ is the space of probability measures supported in K. The existence of an optimal measure for this problem is immediately established by applying the existence result in [MP01] for the particular case where K is bounded.

In addition, (RP) can be rewritten by taking advantage of the moment structure of the cost density and the state equation. If we consider the set

$$\Lambda = \{m \in \mathbf{R}^s : \ m = \int_K \Phi(\lambda)d\nu(\lambda), \nu \in P(K)\},$$

then for each Young measure $\mu = \{\mu_t\}_t$ we can associate a function in $L^\infty([0,T], \Lambda)$ given by

$$m(t) = \int_K \Phi(\lambda)d\mu_t(\lambda).$$

This relation is not one-to-one but we can also associate at least one Young measure to each function in $L^\infty([0,T], \Lambda)$. The set Λ is very especial. Indeed, notice that L defined above as

$L = \Phi(K)$ is part of Λ as it corresponds to generalized moments associated to Dirac-type Young measures. Moreover, in [Me04] it was shown that when K is a compact and convex set we have

$$\Lambda = \overline{co(L)} = co(L)$$

so that Λ is a convex, compact set, defined as

$$\Lambda = co(\Phi(K)).$$

This considerations allow us to conclude that the relaxed problem (RP) is equivalent to the linear optimal control problem

$$(LP) \quad \text{Minimize in } m \in \Lambda : \quad \int_0^T c(x(t)) \cdot m(t) + c_0(x(t))dt$$

subject to

$$\mathtt{x}'(t) = Q(x(t))m(t) + Q_0(x(t)), \quad x(0) = x_0,$$

whose optimal solution (for the existence of such a solution see [Ce74]) corresponds to a Young measure which is an optimal solution (not necessarily unique) of (RP). Next, we will characterize this optimal solution, say $\tilde{m}(.)$ of (LP). To that purpose consider the function

$$\varphi(\mathtt{x}, \xi) =$$

$$\begin{cases} \min_{m\in\Lambda}\{c(x) \cdot m + c_0(x) \ : \xi = Q(x)m + Q_0(x)\} & \text{if } \xi \in Q(x)\Lambda + Q_0(x) \\ +\infty & \text{else.} \end{cases}$$

This density function is the typical integrand of the cost which defines the equivalent variational problem (VP)

$$\text{Minimize in } x(t) : \quad \int_0^T \varphi(x(t), \mathtt{x}'(t))dt$$

subject to $x(0) = x_0$, $x(t) \in AC([0,T], \mathbf{R}^N)$. The equivalence between problems (VP) and (LP) is well known and can be found in [Rk75] , [Ce74] and in more recent works under a similar framework [P97b], [PT07]. Accordingly, there is a solution for (VP), let us say $\tilde{\mathtt{x}}(.)$, whose connection to $\tilde{m}(.)$ is established through the relation

$$\varphi(\tilde{\mathtt{x}}(t), \tilde{\mathtt{x}}'(t)) = \min_{m\in\Lambda}\{c(\tilde{\mathtt{x}}(t)) \cdot m(t) + c_0(\tilde{\mathtt{x}}(t) : \ \tilde{\mathtt{x}}'(t) = Q(\tilde{\mathtt{x}}(t))m(t) + Q_o(\tilde{\mathtt{x}}(t))\}$$

$$= c(\tilde{\mathtt{x}}(t)) \cdot \tilde{m}(t) + c_0(\tilde{\mathtt{x}}(t)) \quad a.e. \quad t \in (0,T).$$

This means that for almost every t, $\tilde{m}(t)$ is the minimizer of

$$\{c(\tilde{\mathtt{x}}(t)) \cdot m(t) + c_0(\tilde{\mathtt{x}}(t) : \ \tilde{\mathtt{x}}'(t) = Q(\tilde{\mathtt{x}}(t))m(t) + Q_o(\tilde{\mathtt{x}}(t))\}.$$

By Lemma 4.1,

$$\tilde{m}(t) \in L = \Phi(K)$$

so that there is a Dirac-type Young measure μ solution of (RP), associated to $\tilde{m}$. As a consequence, (CP) has an optimal solution $u \in L^\infty([0,T],K)$ such that $\mu = \{\delta_{u(t)}\}_{t\in(0,T)}$.

Let us now prove the second part of the theorem. Suppose that $u_1(.)$ and $u_2(.)$ are different optimal solutions of (CP). Then $\mu_1 = \{\delta_{u_1(t)}\}_t$ and $\mu_2 = \{\delta_{u_2(t)}\}_t$ are optimal solutions of (RP). As Φ is component-wise one to one, the corresponding generalized moments defined by $m_1(t) = \Phi(u_1(t))$ and $m_2(t) = \Phi(u_2(t))$ are different optimal solutions of (LP). Hence for $\lambda \in]0,1[$, we have that $m = \lambda m_1 + (1-\lambda)m_2$ is also an optimal solution of the linear problem (LP) and therefore $m \in L$. But since $L = \Phi(K)$ and Φ is strictly convex for some component i, m does not belong to L. A contradiction. Therefore we must have $u_1 = u_2$. □

4.3 Proof of Theorem 4.1

In this section we will apply the first part of Theorem 4.2 to the optimal control problem (UVP). In our case, Φ is not injective so that we cannot conclude about uniqueness. In fact, some numerical simulations (see [GOP09]) suggest that the solution of (UVP) is not unique. We proceed in several steps:

4.3.1 Step 1: the matrices Q and Q_0

We start by paying some attention to the matrices Q and Q_0 of the control system, as it is fundamental to verify the well-posedness character of the state law and condition (4.11) of Theorem 4.2. We recall the notation introduced in Section 4.1 where we have set

$$\mathtt{x} = (\mathtt{x}, \mathtt{y}, \mathtt{z}, \phi, \theta, \psi, \mathtt{u}, \mathtt{v}, \mathtt{w}, \mathtt{p}, \mathtt{q}, \mathtt{r}) \in \Omega \subset \mathbf{R}^{12},$$

with $X_{world} = (\mathtt{x}, \mathtt{y}, \mathtt{z}; \phi, \theta, \psi)$ and $V_{body} = (\mathtt{u}, \mathtt{v}, \mathtt{w}; \mathtt{p}, \mathtt{q}, \mathtt{r})$. Using this notation, accordingly to what we have seen also in Section 4.1 and using the data in [GOP09] we know that Q is given by

$$Q = \begin{pmatrix} 0_{6\times 6} \\ M^{-1}F(V_{body}) \end{pmatrix}$$

where the matrix M is given by

$$M =$$

$$\begin{pmatrix} m-\frac{\rho}{2}L^3X'_{\dot{u}} & 0 & 0 & 0 & mZ_G & -mY_G \\ 0 & m-\frac{\rho}{2}L^3Y'_{\dot{v}} & 0 & -mZ_G-\frac{\rho}{2}L^4Y'_{\dot{p}} & 0 & mX_G-\frac{\rho}{2}L^4Y'_{\dot{r}} \\ 0 & 0 & 0 & m-\frac{\rho}{2}L^3Z'_{\dot{w}} & -mX_G-\frac{\rho}{2}L^4Z'_{\dot{q}} & mY_G \\ 0 & -mZ_G-\frac{\rho}{2}L^4K'_{\dot{v}} & mY_G & I_x-\frac{\rho}{2}L^5K'_{\dot{p}} & -I_{xy} & -I_{xz}-\frac{\rho}{2}L^5K'_{\dot{r}} \\ mZ_G & 0 & -mX_G-\frac{\rho}{2}L^4M'_{\dot{w}} & -I_{xy} & I_y-\frac{\rho}{2}L^5M'_{\dot{q}} & -I_{yz} \\ -mY_G & mX_G-\frac{\rho}{2}L^4N_{\dot{v}} & 0 & -I_{xz}-\frac{\rho}{2}L^5N'_{\dot{p}} & -I_{yz} & I_z-\frac{\rho}{2}L^5N'_{\dot{r}} \end{pmatrix}$$

and $F = (G,\ H)$, $G, H \in \mathcal{M}^{6\times 3}$, with

$$G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{\rho}{2}l^2(Y'_{\delta_r}+Y'_{\delta_r\eta}(\eta-\frac{1}{C})C)u^2 \\ \frac{\rho}{2}l^2(Z'_{\delta_b})u^2 & \frac{\rho}{2}l^2(Z'_{\delta_s}+Z'_{\delta_s\eta}(\eta-\frac{1}{C})C)u^2 & 0 \\ 0 & 0 & \frac{\rho}{2}l^3(K'_{\delta_r})u^2 \\ \frac{\rho}{2}l^3(M'_{\delta_b})u^2 & \frac{\rho}{2}l^3(M'_{\delta_s}+M'_{\delta_s\eta}(\eta-\frac{1}{C})C)u^2 & 0 \\ 0 & 0 & \frac{\rho}{2}l^3(N'_{\delta_r}+N'_{\delta_r\eta}(\eta-\frac{1}{C})C)u^2 \end{pmatrix}$$

and

$$H = \begin{pmatrix} \frac{\rho}{2}l^2(X'_{\delta_b\delta_b})u^2 & \frac{\rho}{2}l^2(X'_{\delta_s\delta_s})u^2 & \frac{\rho}{2}l^2(X'_{\delta_r\delta_r})u^2 \\ 0 & 0 & \\ 0 & 0 & \\ 0 & 0 & \\ 0 & 0 & \\ 0 & 0 & \end{pmatrix}.$$

Finally, considering the dimensionless hydrodynamic coefficients in [GOP09, Appendix] gives

$$Q(x) = \mathtt{u}^2 \begin{pmatrix} & 0_{6\times 6} & & & & \\ Q_{11} & Q_{12} & Q_{13} & Q_{14} & Q_{15} & Q_{16} \\ 0 & 0 & Q_{23} & 0 & 0 & 0 \\ Q_{31} & Q_{32} & Q_{33} & Q_{34} & Q_{35} & Q_{36} \\ 0 & 0 & Q_{43} & 0 & 0 & 0 \\ Q_{51} & Q_{52} & Q_{53} & 0 & 0 & 0 \\ 0 & 0 & Q_{63} & 0 & 0 & 0 \end{pmatrix}$$

$$= (x_7)^2 \begin{pmatrix} & 0_{6\times 6} & & & & \\ -0.0056307 & -0.0056219 & 0.0002292 & -0.0028418 & -0.0011310 & -0.0037067 \\ 0 & 0 & -0.0001291 & 0 & 0 & 0 \\ 1.527832 & 1.4903911 & -0.0617573 & -0.0001656 & -0.0000659 & -0.0002160 \\ 0 & 0 & 0.0001049 & 0 & 0 & 0 \\ -0.0162938 & -0.0162684 & 0.0006631 & 0 & 0 & 0 \\ 0 & 0 & -0.0002773 & 0 & 0 & 0 \end{pmatrix}.$$

We remark that Q, the 12×6 matrix of the coefficients interacting with the control, only depends on the surge velocity. Such particularity allows us to verify condition (4.11) quite easily, as we will see after.

As for Q_0, it is given by

$$Q_0 = \begin{pmatrix} \mathcal{T}(\phi,\theta,\psi)V_{body} \\ M^{-1}F_0(V_{body},\phi,\theta,\psi) \end{pmatrix} \in \mathbf{R}^{12}.$$

where $\mathcal{T}$ is the transformation matrix in the kinematic equations

$$(X_{world})' = \mathcal{T}(\phi,\theta,\psi)V_{body}$$

defined by

$$\mathcal{T} = \begin{pmatrix} J_1(\phi,\theta,\psi) & 0_{3\times 3} \\ 0_{3\times 3} & J_2(\phi,\theta,\psi) \end{pmatrix}$$

with

$$J_1(\phi,\theta,\psi) = \begin{pmatrix} \cos\psi\cos\theta & -\sin\psi\cos\theta+\cos\psi\sin\theta\sin\phi & \sin\psi\sin\phi+\cos\psi\cos\phi\sin\theta \\ \sin\psi\cos\theta & \cos\psi\cos\phi+\sin\phi\sin\theta\sin\psi & -\cos\psi\sin\phi+\sin\theta\sin\psi\cos\phi \\ -\sin\theta & \cos\theta\sin\phi & \cos\theta\cos\phi \end{pmatrix}$$

and

$$J_2(\phi,\theta,\psi) = \begin{pmatrix} 1 & \sin\phi\tan\theta & \cos\phi\tan\theta \\ 0 & \cos\phi & -\sin\phi \\ & \sin\phi/\cos\theta & \cos\phi/\cos\theta \end{pmatrix}.$$

Concerning F_0, it is defined in [GOP09] through the ordinary differential system of six equations

$$MV'_{body} = F_0(V_{body},\phi,\theta,\psi) + F(V_{body})\Phi(u)$$

so that it corresponds to the terms independent of the controls. To obtain Q_0 we write F_0 with the data given in [GOP09] and multiply it by M^{-1}, just as we have done for Q. Using the state notation

$$\mathtt{x} = (x_j), \quad \bar{F}_0(x) = ((\bar{F}_0)_j) = M^{-1}F_0(x), \quad 1 \leq j \leq 6,$$

we obtain

$$(\bar{F}_0)_1 = 0.21 \sin x_4 \cos x_5 + 5.593 x_{12} |x_{12}| - 10.68 x_{12}^2 - 7.234 x_{11}x_{12} + 2.905 x_{10}x_{12} - 0.93 x_8 x_{12} - 0.11 x_7 x_{12}$$
$$- 19.65 x_{11} |x_{11}| + 5.658 x_{11}^2 + 0.015 x_{10}x_{11} - 1.809 x_9 x_{11} + 0.61 x_7 x_{11} + 7.252 x_{10} |x_{10}| - 0.4 x_{10}^2 + 0.14 x_9 x_{10}$$
$$- 2.477 x_8 x_{10} + 0.21 x_7 x_{10} - 0.0085 \sqrt{x_9^2 + x_8^2}\, |x_9| - 0.0022 x_7 |x_9| - 0.0056 x_8 \sqrt{x_9^2 + x_8^2} + 0.0074 x_9^2$$
$$- 0.015 x_8 x_9 - 0.022 x_7 x_9 + 0.012 x_8 |x_8| + 0.22 x_8^2 + 0.013 x_7 x_8 - 0.0012 x_7^2 - 0.014 x_7 + 0.2$$

$$(\bar{F}_0)_2 = 0.032 \sin x_4 \cos x_5 + 4.918 x_{12} |x_{12}| - 1.028 x_{11}x_{12} - 0.21 x_7 x_{12} + 0.064 x_{10}x_{11} + 1.101 x_{10} |x_{10}|$$
$$+ 0.41 x_9 x_{10} - 0.0073 x_7 x_{10} - 0.023 x_8 \sqrt{x_9^2 + x_8^2} - 0.061 x_8 x_9 + 0.0017 x_8 |x_8|$$
$$- 0.01 x_7 x_8 + 2.4985 \times 10^{-7} x_7^2 - 5.6213 \times 10^{-5} x_7 + 0.0012$$

$$(\bar{F}_0)_3 = -0.43 \sin x_5 - 57.56 \sin x_4 \cos x_5 - 1508. x_{12} |x_{12}| + 5212. x_{12}^2 + 1951. x_{11}x_{12} + 98.94 x_{10}x_{12} + 884.9 x_8 x_{12}$$
$$+ 30.0 x_7 x_{12} + 5149. x_{11} |x_{11}| + 108.1 x_{11}^2 - 4.058 x_{10}x_{11} - 0.047 x_9 x_{11} - 166.7 x_7 x_{11} - 1956. x_{10} |x_{10}|$$
$$+ 107.8 x_{10}^2 - 38.2 x_9 x_{10} + 667.5 x_8 x_{10} - 57.7 x_7 x_{10} + 2.215 \sqrt{x_9^2 + x_8^2}\, |x_9| + 0.59 x_7 |x_9| + 1.501 x_8 \sqrt{x_9^2 + x_8^2}$$
$$+ 4.2833 \times 10^{-4} x_9^2 + 3.913 x_8 x_9 + 6.062 x_7 x_9 - 3.109 x_8 |x_8| - 54.46 x_8^2 - 3.376 x_7 x_8 + 0.088 x_7^2 + 0.099 x_7 - 2.205$$

$$(\bar{F}_0)_4 = -0.098 \sin x_4 \cos x_5 - 2.562 x_{12} |x_{12}| + 3.317 x_{11}x_{12} + 0.051 x_7 x_{12} - 0.0069 x_{10}x_{11} - 3.325 x_{10} |x_{10}|$$
$$- 0.065 x_9 x_{10} - 0.098 x_7 x_{10} + 0.0025 x_8 \sqrt{x_9^2 + x_8^2} + 0.0066 x_8 x_9 - 0.0053 x_8 |x_8| - 0.0057 x_7 x_8$$
$$- 7.5427 \times 10^{-7} x_7^2 + 1.697 \times 10^{-4} x_7 - 0.0038$$

$$(\bar{F}_0)_5 = 0.62 \sin x_4 \cos x_5 + 16.2 x_{12} |x_{12}| - 56.57 x_{12}^2 - 20.96 x_{11}x_{12} - 9.622 x_8 x_{12} - 0.32 x_7 x_{12} - 56.86 x_{11} |x_{11}|$$
$$- 1.157 x_{11}^2 + 0.044 x_{10}x_{11} + 1.76 x_7 x_{11} + 21.01 x_{10} |x_{10}| - 1.157 x_{10}^2 + 0.41 x_9 x_{10} - 7.167 x_8 x_{10} + 0.62 x_7 x_{10}$$

$$-0.025\sqrt{x_9^2+x_8^2}\,|x_9| - 0.0064x_7\,|x_9| - 0.016x_8\sqrt{x_9^2+x_8^2} - 0.042x_8x_9 - 0.065x_7x_9 + 0.033x_8\,|x_8|$$
$$+0.59x_8^2 + 0.036x_7x_8 - 9.4993\times 10^{-4}x_7^2 - 0.0011x_7 + 0.024$$

$$(\bar{F}_0)_6 = 0.0037\sin x_4\cos x_5 + 2.308x_{12}\,|x_{12}| - 0.12x_{11}x_{12} - 0.079x_7x_{12} - 1.91x_{10}x_{11} + 0.12x_{10}\,|x_{10}| + 0.0063x_9x_{10}$$
$$-0.0073x_7x_{10} - 0.0043\,|x_8|\,\sqrt{x_9^2+x_8^2} - 3.2111\times 10^{-4}x_8\sqrt{x_9^2+x_8^2} - 0.071x_8x_9 + 1.9811\times 10^{-4}x_8\,|x_8|$$
$$-0.0042x_7x_8 + 2.8285\times 10^{-8}x_7^2 - 6.3637\times 10^{-6}x_7 + 1.412\times 10^{-4}$$

Notice that in fact $\bar{F}_0$ does not depend on (x_1, x_2, x_3), but for simplicity we will still consider Q_0 as a vector function from $\mathbf{R}^{12}$ to $\mathbf{R}^{12}$ which is described by

$$Q_0(x) = \begin{pmatrix} \begin{pmatrix} J_1(x_4,x_5,x_6) & 0_{3\times 3} \\ 0_{3\times 3} & J_2(x_4,x_5,x_6) \end{pmatrix} \begin{pmatrix} x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \end{pmatrix} \\ \bar{F}_0(x) \end{pmatrix}$$

where J_1, J_2 and $\bar{F}_0$ are as above.

4.3.2 Step 2: local existence and uniqueness of solutions for the state law

Let us now show that it is possible to find a time interval $I = [0, T]$ for which the initial value problem

$$\text{(IVP)} \quad \begin{cases} x'(t) = Q(x(t))\,\Phi(u(t)) + Q_0(\mathtt{x}(t)), & 0 < t < T \\ x(0) = x^0 \in \Omega \end{cases}$$

is well posed in the sense that for every control function $u \in L^\infty(0, T; K)$ there is a unique solution. We start by recalling the classical theory on this subject and therefore we rewrite (IVP) in the standard way

$$\begin{cases} x'(t) = \mathbf{f}(t, x(t)), & 0 < t < T \\ x(0) = x^0 \in \Omega \subset \mathbf{R}^N, \end{cases} \tag{4.12}$$

with $\mathbf{f} : I \times \Omega \to \mathbf{R}^N$, $N = 12$ in our case. A (Carathéodory) solution of (4.12) is an absolutely continuous function

$$\mathtt{x} : (0, T_1) \to \Omega, \text{ with } T_1 \le T,$$

such that for all $t \in (0, T_1)$

$$x(t) = x^0 + \int_0^t \mathbf{f}(s, \mathtt{x}(s))\, ds.$$

The solution $\mathtt{x} : (0, T_1) \to \Omega$ is said to be maximal if for another solution $\overline{\mathtt{x}} : (0, T_2) \to \Omega$ of (4.12) the two following conditions hold:

(i) $T_2 \leq T_1$, and

(ii) $\overline{\mathtt{x}}(t) = x(t)$ for all $0 \leq t \leq T_2$.

As is well-known (see for instance [S90, Appendix C]), if $\mathbf{f}$ satisfies conditions (H1)-(H4) below, then we can ensure the existence and uniqueness of a maximal solution for (4.12).

(H1) For each $x \in \Omega$, the function $\mathbf{f}(\cdot, x) : I \to \mathbf{R}^N$ is measurable,

(H2) for each $t \in I$, the function $\mathbf{f}(t, \cdot) : \Omega \to \mathbf{R}^N$ is continuous,

(H3) $\mathbf{f}$ is locally Lipschitz on $\mathtt{x}$, that is, for each $x^0 \in \Omega$ there are a real number $\rho > 0$ and a locally integrable function

$$\alpha : I \to \mathbf{R}^+$$

such that the ball $B_\rho(x^0)$ of radius ρ centered at x^0 is contained in Ω and

$$\|\mathbf{f}(t, x) - \mathbf{f}(t, \mathbf{y})\| \leq \alpha(t) \|\mathtt{x} - \mathbf{y}\|$$

for each $t \in I$ and $\mathtt{x}, \mathbf{y} \in B_\rho(\mathtt{x}^0)$, and

(H4) $\mathbf{f}$ is locally integrable on t, that is, for each $x^0 \in \Omega$ there exists a locally integrable function $\beta : I \to \mathbf{R}^+$ such that

$$\left\|\mathbf{f}(t, x^0)\right\| \leq \beta(t) \quad \text{a. e. } t \in I.$$

Our next task is to check that (H1)-(H4) hold in our particular case. For any $u \in L^\infty(\mathbf{R}; K)$, since the control variable u appears in linear and quadratic form, it is clear that the function

$$\mathbf{f}(t, x) = Q(x)\, \Phi(u(t)) + Q_0(x) \tag{4.13}$$

is measurable with respect to t for each fixed $x \in \Omega$. In addition, looking at the particular form of (4.13), it is clear that for each t, the function $x \to \mathbf{f}(t, \mathbf{x})$ is continuous. With respect to conditions (H3) and (H4), again the form in which the controls appear let us conclude that (H4) is satisfied. As for the local Lipschitz condition (H3), since $\mathbf{f} = (f_1, \cdots, f_{12})$ is a vector function, we should check that condition for each component. Due to the constraints (4.2) and

taking into account that the first six components of $\mathbf{f}$ only include the transformation matrix between body and world references frames, we have that $f_1, \cdots, f_6 \in C^\infty(\Omega)$ and therefore they are locally Lipschitz with respect to $\mathtt{x}$. As for the remaining $f_7, \cdots, f_{11}$, we notice that these components include by one side, polynomial terms, terms in the form of absolute value, terms with the structure of $\sqrt{x_j^2 + x_k^2}$, where $\mathtt{x} = (x_1, \cdots, x_{12})$, all of them locally Lipschitz, and products of locally Lipschitz functions, also locally Lipschitz, by the other.

Therefore we may state that for each $x^0 \in \Omega$ and $\mathbf{u} \in L^\infty(\mathbf{R}; K)$ there exists a maximal time $T\left(x^0, u\right)$ and a unique maximal solution of (IVP) defined on $\left[0,\ T\left(x^0, u\right)\right]$. In fact, looking at the proof of the mentioned existence result (see [S90]), we can see that $T\left(x^0, u\right)$ depends on both $\alpha(t) = \alpha(u(t))$ and $\beta(t) = \beta(u(t))$ in the sense that

$$\int_0^t \alpha(\tau)\, d\tau < 1 \quad \forall t \in [0, T(x^0, u)]$$

and

$$\int_0^t \rho\alpha(\tau) + \beta(\tau)\, d\tau < \rho \quad \forall t \in [0, T(x^0, u)].$$

Since Φ is continuous on the compact set K and taking into account the particular structure of matrices Q and Q_0, we can choose $\alpha(t)$ and $\beta(t)$ such that (H1)-(H4) are satisfied simultaneously to all $u \in L^\infty(\mathbf{R}_+; K)$ and consequently we can choose T (uniformly in u) such that problem (IVP) has a unique solution in $I = [0, T]$, with $T = T(x^0)$, for every $u \in L^\infty(I; K)$.

Remark 4.1. *It is not difficult to convince ourselves that for some suitable inputs* $\mathbf{u}$*, the corresponding solution* $\mathbf{x}$ *of the state law is not defined for all* $t > 0$ *because of the constraints (4.2). That is, we can not expect to have a global solution for all admissible* $\mathbf{u}$*. Moreover, in a real situation we also must impose some constraints on the state variables* (x, y, z) *due to the finite dimension of ocean. These restrictions, which are specially important in a situation in which the submarine is moving in littoral waters, may let the solution* $\mathbf{x}$ *blow-up in finite time.*

4.3.3 Step 3: checking condition (4.11) in Theorem 4.2

We need to describe for every $x \in \mathbf{R}^{12}$ (and corresponding pair $(c(x), Q(x))$) the set

$$\mathcal{N}(c(x), Q(x)) = \left\{v \in \mathbf{R}^6 : Q(\mathbf{x})v = 0, c(x) \cdot v \leq 0\right\},$$

and check that such set is contained in

$$\mathcal{N}(K, \Phi) =$$

$$\left\{v = (v_1, \cdots, v_6) \in \mathbf{R}^6 :\ \text{for each } u \in K,\ \text{either } \nabla\Psi(\Phi(u))v = 0 \text{ or there is } i \text{ with } \nabla\Psi_i(\Phi(u))v > 0\right\},$$

where Q is like described in the beginning of this section, where the data from [GOP09] were used. Notice that in this model, as the propulsion coefficients are considered to be constant, the surge velocity $\mathtt{u} = x_7$ is always positive.

Let us first find the solution of $Qv = 0$. Assuming that $x_7 \neq 0$ we have

$$\begin{cases} v_3 = 0 \\ v_6 = -\frac{1}{Q_{16}}(Q_{11}v_1 + Q_{12}v_2 + Q_{14}v_4 + Q_{15}v_5) \\ v_6 = -\frac{1}{Q_{36}}(Q_{31}v_1 + Q_{32}v_2 + Q_{34}v_4 + Q_{35}v_5) \\ v_2 = -\frac{Q_{51}}{Q_{52}}v_1. \end{cases}$$

Thus

$$\begin{cases} \dots \\ \frac{1}{Q_{16}}(Q_{11} - \frac{Q_{51}}{Q_{52}}Q_{12})v_1 + \frac{Q_{14}}{Q_{16}}v_4 + \frac{Q_{15}}{Q_{16}}v_5 = \\ \frac{1}{Q_{36}}(Q_{31} - \frac{Q_{51}}{Q_{52}}Q_{32})v_1 + \frac{Q_{34}}{Q_{36}}v_4 + \frac{Q_{35}}{Q_{36}}v_5 \\ \dots \end{cases}$$

but

$$\frac{Q_{14}}{Q_{16}} = 0.7666667 = \frac{Q_{34}}{Q_{36}}$$

and

$$\frac{Q_{15}}{Q_{16}} = 0.3051282 = \frac{Q_{35}}{Q_{36}}$$

so that

$$\begin{cases} \dots \\ (Q_{11} - \frac{Q_{51}}{Q_{52}}Q_{12})v_1 = (Q_{31} - \frac{Q_{51}}{Q_{52}}Q_{32})v_1 \\ \dots \end{cases} .$$

Since

$$Q_{11} - \frac{Q_{51}}{Q_{52}}Q_{12} = 0 \neq 0.0348637 = Q_{31} - \frac{Q_{51}}{Q_{52}}Q_{32},$$

we have

$$Qv = 0 \Leftrightarrow \begin{cases} v_1 = 0 \\ v_2 = 0 \\ v_3 = 0 \\ v_6 = -\frac{1}{Q_{16}}(Q_{14}v_4 + Q_{15}v_5) = \\ -\frac{1}{Q_{36}}(Q_{34}v_4 + Q_{35}v_5). \end{cases}$$

Before completing the characterization of $\mathcal{N}(c,Q)$ notice that the function Ψ used in describing $\mathcal{N}(K,\Phi)$ is given by

$$\Psi(m) = (m_1^2 - m_4, m_2^2 - m_5, m_3^2 - m_6), \quad m = (m_1, \cdots, m_6)\,,$$

so that Ψ is obviously $\mathcal{C}^1$ and convex. Moreover,

$$\nabla\Psi(m) = [2diag(m_1, m_2, m_3), -I_3].$$

Hence, for v such that $Qv = 0$ the vector $\nabla\Psi(m)\cdot v$ is in fact

$$2\ \mathtt{diag}[m_1, m_2, m_3] \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} - I_3 \begin{pmatrix} v_4 \\ v_5 \\ v_6 \end{pmatrix}$$

$$= - \begin{pmatrix} v_4 \\ v_5 \\ v_6 \end{pmatrix}.$$

This means that for a vector v (in the manifold $Qv = 0$) to belong to $\mathcal{N}(K,\Phi)$, it must satisfy

$$v_4 = v_5 = v_6 = 0$$

or else one of those three components must be negative.

As a consequence, condition (4.11) can only hold if the vectors in $\mathcal{N}(c,Q)$ have one of the last three components strictly negative or either all null. But as we have seen, for the case where the surge velocity $u = x_7 \neq 0$ we have

$$v_6 = -\frac{1}{Q_{16}}(Q_{14}v_4 + Q_{15}v_5) = \ -\frac{1}{Q_{36}}(Q_{34}v_4 + Q_{35}v_5).$$

Hence, if both v_4 and v_5 are positive or null, we have v_6 necessarily negative or also null. Consequently

$$\mathcal{N}(c,Q) \subset \mathcal{N}(K,\Phi),$$

and applying Theorem 4.2 the proof is complete.

Appendix A

Some Classical Theory

A.1 Measurable Selection

Definition A.1 ([AF90b] pp. 307)**.** *Let X be a separable metric space, $(\Omega, \mathcal{A})$ a measurable space and $F : \Omega \to X$ a set valued function taking nonempty and closed subsets of X as values. We say that F is measurable if for every $\mathcal{O} \subset X$ we have*

$$F^{-1}(\mathcal{O}) = \{\omega \in \Omega : F(\omega) \cap \mathcal{O} \neq \emptyset\} \in \mathcal{A} \tag{A.1}$$

Definition A.2 ([AF90b] pp. 308)**.** *Let X be a separable complete metric space, $(\Omega, \mathcal{A})$ a measurable space, F a measurable set valued function from Ω to nonempty and closed subsets of X. We call $f : \Omega \to X$ a measurable selection of F if it verifies*

$$\forall \omega \in \Omega,\ f(\omega) \in F(\omega) \tag{A.2}$$

Theorem A.1 (Measurable Selection [AF90b] p. 308)**.** *Let X be a separable complete metric space, $(\Omega, \mathcal{A})$ a measurable space, F a measurable set valued function from Ω to nonempty and closed subsets of X.*

Then there exists a measurable selection of F.

A.2 Parametrized Measures

A function $f \in X'$ is weakly* measurable if $x \to < T, f(x) >$ is measurable $\forall T \in X$. Let $X = C_0(\mathbf{R}^M)$ and $X' = \mathcal{M}(\mathbf{R}^m)$ where $\mathcal{M}(\mathbf{R}^m)$ is the set of Radon Measures, the σ-additive measures with respect to the Borel σ-algebra of $\mathbf{R}^m$, which are regular and have bounded variation (see [Rb97b] pp.39-40 for the corresponding definitions).

Let $|\mu|$ be defined as in [Rb97b] pp.39 and $\|\mu\|_{\mathcal{M}(\mathbf{R}^m)} = |\mu|(\mathbf{R}^m)$.

$$L^\infty_{w^*}((0,T),\mathcal{M}(\mathbf{R}^m)) =$$

$$\{\nu : (0,T) \to \mathcal{M}(\mathbf{R}^m) \text{ s.t. } \nu \text{ is w* measurable and } \|\nu\|_{L^\infty_{w^*}((0,T),\mathcal{M}(\mathbf{R}^m))} = esssup_{t\in(0,T)}\|\nu_t\|_{\mathcal{M}(\mathbf{R}^m)} < \infty\}$$

$$\mathcal{Y}^\infty((0,T),P(K)) =$$

$$\left\{\mu = \{\mu_t\}_{t\in(0,T)} \in L^\infty_{w^*}((0,T),\mathcal{M}(\mathbf{R}^m)) : \mu_t \in P(K),\ \forall t \in (0,T)\right\},$$

Remark A.1. *If K is unbounded the theory can also be establish to quad for every $p > 1$ ([Rb97b],[P97b]). For that we consider*

$$L^p_w((0,T),\mathcal{M}(\mathbf{R}^m)) =$$

$$\{\nu : (0,T) \to \mathcal{M}(\mathbf{R}^m) \text{ s.t. } \nu \text{ is } w \text{ measurable and } \|\nu\|_{L^p_w((0,T),\mathcal{M}(\mathbf{R}^m))} = \int_0^T \int_K \|\lambda\|^p d\mu_t(\lambda)dt < \infty\}$$

$$\mathcal{Y}^p((0,T),P(K)) =$$

$$\left\{\mu = \{\mu_t\}_{t\in(0,T)} \in L^p_w((0,T),\mathcal{M}(\mathbf{R}^m)) :\ \mu_t \in P(K), \forall t \in (0,T)\right\}$$

Theorem A.2. *[Ta79] (See also [Rb97b] pp.114) For any sequence $\{z_j\} \in L^\infty((0,T),B)$ where $B \subset \mathbf{R}^s$ is a compact set, there exists a subsequence, again denoted by $\{z_j\}$ and a Young measure $\nu \in \mathcal{Y}^\infty((0,T),P(B))$ such that, for any $v \in C(B)$,*

$$v(z_j) \rightharpoonup v_\nu \quad \text{wealky}^* \text{ in} \quad L^\infty(0,T),$$

where

$$v_\nu(t) = \int_B v(s)d\nu_t(s).$$

Reciprocally, for each $\nu \in \mathcal{Y}^\infty((0,T),P(B))$ there is a sequence $\{z_j\} \in L^\infty((0,T),B)$ such that

$$v(z_j) \rightharpoonup v_\nu \quad \text{wealky}^* \text{ in} \quad L^\infty(0,T),$$

for every $v \in C(B)$.

Theorem A.3. *([P97b] pp.112)*

Let $z_j = (x_j, u_j) : (0,T) \to \mathbf{R}^N \times \mathbf{R}^m$ be a bounded sequence in $L^\infty((0,T),\mathbf{R}^{N\times m})$ such that

$$\{x_j\} \to x \text{ in } L^\infty((0,T),\mathbf{R}^N).$$

If $\mu = \{\mu_t\}_t$ is the parametrized measure associated with $\{z_j\}$, then

$$\mu_t = \delta_{x(t)} \otimes \nu_t \text{ a.e. } t \in (0,T)$$

where $\nu = \{\nu_t\}_t$ is the parametrized measure corresponding to $\{u_j\}$.

A.3 Control Theory

Theorem A.4. *[Ce83b] pp.310 Consider the Mayer optimal control problem*

$$\min J(x) = g(x(T))$$

subject to

$$x'(t) = f(x(t), u(t))$$

$$x(0) = x_0, \quad u \in K$$

$$u \in L^{\infty}(0,T), \mathbf{R}^s), \quad x \in AC((0,T), \mathbf{R}^d)$$

where g is a lower semicontinuous function and f is Lipschitz with respect to the state variable and continuous with respect to the control variable, such that the set of admissible solutions is non empty.

Assume also that the set

$$Q(x) = \{z : z = f(x,u), \ \text{certain } u \in K\}$$

is convex and that K is compact. Then there is an optimal solution (x,u) for the optimal control problem.

Remark A.2. *The convexity of $Q(x) = f(x,K)$ is essential to apply the Filippov existence result for the differential inclusion $x' \in f(x,K)$. The Lipschitz condition on f also allows us to conclude that the solution of the control system must lie in a compact set. See [Ce83b] and [AF90b] for more details.*

Theorem A.5. *[Ce83b] Consider the Lagrange optimal control problem*

$$\min I(x,u) = \int_0^T F(x(t), u(t)dt$$

subject to

$$x'(t) = f(x(t), u(t))$$

$$x(0) = x_0, \quad u \in K$$

$$u \in L^{\infty}(0,T), \mathbf{R}^m), \quad x \in AC((0,T), \mathbf{R}^N)$$

where F is a continuous function and f is Lipschitz with respect to the state variable and continuous with respect to the control variable, such that the set of admissible solutions is non empty.

Consider also that the set

$$\bar{Q}(x) = \{(v,z) : \ v \geq F(x,u), \ z = f(x,u), \ u \in K\}$$

is convex and K is compact. Then there is an optimal pair (x,u).

Proof. We shall apply Theorem A.4. Let us consider $\tilde{u} = (u^0, u)$ with values in $\mathbf{R}^{1+m}$ and $\tilde{x} = (x^0, x)$ in $\mathbf{R}^{1+N}$, where

$$x^0(0) = 0 \text{ and } (x^0)' = u^0.$$

This means, in particular that $x^0(T) = \int_0^T u^0 dt$.

Due to the Lipschitz condition on f, the fact that K is compact and by Gronwall's Lemma, we can conclude that an admissible state x must belongs to a compact set, say Ω, so that $|F(\Omega \times K)| \leq R$ for a certain radius R. Using this, consider the viability set

$$\tilde{K} = \{(u^0, u) : \ F(x, u) \leq u^0 \leq R, \ u \in K\}$$

which is also a compact set. Taking $\tilde{f}(\tilde{x}, \tilde{u}) = (u^0, f)$ we define also the set

$$\tilde{Q}(\tilde{x}) = \{\tilde{z} = (z^0, z) : \ \tilde{z} = \tilde{f}(\tilde{x}, \tilde{u}), \ \tilde{u} \in \tilde{K}\}$$

$$= \{(z^0, z) : \ F(x, u) \leq z^0 \leq R, \ z = f(x, u), \ u \in K\}$$

which can be easily seen to be convex using the fact that $\bar{Q}$ has the same property.

Therefore, if we consider the Mayer optimal control problem

$$\min \tilde{J}(\tilde{x}) = x^0(T)$$

subjected to

$$\tilde{x}'(t) = \tilde{f}(\tilde{x}(t), \tilde{u}(t))$$

$$x(0) = x_0, \ x^0(0) = 0 \quad \tilde{u} \in \tilde{K}$$

$$u \in L^\infty((0, T), \mathbf{R}^{1+m}), \quad x \in AC((0, T), \mathbf{R}^{1+N}),$$

we can see that $g(x) = x$ is a continuous function and $\tilde{f}$ is Lipschitz with respect to the state variable and continuous with respect to the control variable.

As we have seen above $\tilde{Q}(\tilde{x})$ is a convex set so that by applying Theorem A.4 we conclude that this Mayer problem has an optimal solution. In fact as

$$\tilde{J}(\tilde{x}, \tilde{u}) = x^0(T) = \int_0^T u^0 dt \geq \int_0^T F(x, u) dt = I(x, u)$$

it is easy to see that both problems are equivalents in the sense that if $(\tilde{x}, \tilde{u})$ is an optimal pair for the Mayer problem, the corresponding pair (x, u) is optimal for the Lagrange problem. Reciprocally, if (x, u) is optimal for the Lagrange problem, taking $u^0 = F(x, u)$ and x^0 as above we get $(\tilde{x}, \tilde{u})$ optimal for the Mayer problem. This concludes our proof. □

Consider now the problems (P) and (RP) to be problems (P_1) and (P_2) in chapter 2.

Theorem A.6. *Let $I(u)$ represent the value of the cost function in (P) applied to the control u and $\bar{I}(\nu)$ the value of the cost function in (RP) applied to the Young measure ν. Let $\mathcal{I}$ and $\bar{\mathcal{I}}$ be the corresponding infimums. Under the above assumptions problem (RP) as an optimal solution, so that $\bar{\mathcal{I}}$ is in fact a minimum and*

$$\mathcal{I} = \bar{\mathcal{I}}.$$

Proof. Let u be an admissible control. Taking $\nu_t = \delta_{u(t)}$ then

$$\nu \in \mathcal{Y}^\infty((0,T), P(K))$$

and

$$I(u) = \bar{I}(\nu)$$

so that

$$\mathcal{I} \geq \bar{\mathcal{I}}.$$

If $\nu \in \mathcal{Y}^\infty((0,T), P(K))$ since K is a compact set, by Theorem A.2 we know that exists a sequence of controls $\{u_j\}$ admissible for problem (P). Hence, for each j there is a unique absolutely continuous solution for the initial value problem

$$x_j'(t) = Q(x_j(t))\Phi(u_j(t)) + Q_0(x_j(t))$$

$$x_j(0) = x_0.$$

Putting (Assuming uniform Lipchitz conditions on $(0,T)$)

$$|x_j(t) - x_0| = \left| \int_0^t Q(x_j(s))\Phi(u_j(s)) + Q_0(x_j(s))ds \right| \leq$$

$$\left| \int_0^t Q(x_0)\Phi(u_j(s)) + Q_0(x_0)ds \right| + \int_0^t |Q(x_j(s))\Phi(u_j(s)) + Q_0(x_j(s)) - Q(x_0)\Phi(u_j(s)) - Q_0(x_0)|ds =$$

$$\left| \int_0^t Q(x_0)\Phi(u_j(s)) + Q_0(x_0)ds \right| + \int_0^t L|x_j(t) - x_j(0)|ds \leq C + \int_0^t L|x_j(t) - x_0|ds$$

where L is the Lipchitz constant, and applying Gronwall's Lemma we conclude that $\{x_j\}$ is a bounded sequence in $L^\infty((0,T), \mathbf{R}^N)$. We consider therefore that $\{x_j\} \in \Omega$ where Ω is a compact subset of $\mathbf{R}^N$.

Consider the sequence $\{(x_j, u_j)\}$ bounded in $L^\infty((0,T), \mathbf{R}^{N\times m})$.

Applying Theorem A.2 we know that there is an associated family of probability measures in $\mathcal{Y}^\infty((0,T), P(\Omega \times K))$, say $\mu = \{\mu_t\}_t$ which verifies, for

$$f(x,u) = Q(x)\Phi(u) + Q_0(x) \in C(\Omega \times K)$$

$$f(x_j, u_j) \rightharpoonup f_\nu \ w^* \text{ in } \quad L^\infty(0,T),$$

where

$$f_\nu(t) = \int_{\Omega \times K} f(s) d\mu_t(s).$$

Defining

$$x(t) = x_0 + \int_0^t f_\mu(\tau) d\tau$$

we have

$$x_j(t) - x(t) = \int_0^t [f(x_j(\tau), u_j(\tau)) - f_\mu(\tau)] d\tau.$$

Also $f_j \rightharpoonup f_\mu \ w^*$ in $L^\infty(0,T)$ tells us (see [P97b] pp. 23) that

$$\int_D [f(x_j(\tau), u_j(\tau)) - f_\mu(\tau)] d\tau \to 0$$

for all "cubes" $D = [0,t] \subset [0,T]$. Thus

$$x_j(t) - x(t) \to 0 \text{ a.e. } t \in [0,T].$$

As (x_j) are continuous functions over $[0,T]$ (Lebesgue's Theorem) and both (x_j) and (u_j) are equibounded (they belong to $\Omega \times K$). It is inmediate to see that (x_j) is an equicontinuous family and by Ascoli-Arzela's Theorem (see Remark A.3) we can conclude that there is a subsequence (x_j) such that

$$x_j \to x \in L^\infty((0,T), \mathbf{R}^N).$$

Applying Theorem A.3 we know that μ is given by

$$\mu_t = \delta_{x(t)} \otimes \nu_t \text{ a.e. } t \in (0,T).$$

This means that

$$f_\mu(t) = \int_K f(x(t), \lambda) d\nu_t(\lambda) = x'(t).$$

Let $F(x,u) = c(x)\Phi(u) + c_0(x)$, thus $F \in C(\Omega \times K)$ so that by Theorem A.2

$$F(x_j, u_j) \rightharpoonup F_\mu \ w^* \text{ in } L^\infty(0,T)$$

where

$$F_\mu(t) = \int_{\Omega \times K} F(\lambda) d\mu_t(\lambda).$$

Thus

$$\int_0^T F(x_j(t), u_j(t)) dt \to \int_0^T F_\mu(t) dt = \int_0^T \int_K F(x(t), \lambda) d\nu_t(\lambda) dt = \bar{I}(\nu).$$

Hence $\bar{\mathcal{I}} \geq \mathcal{I}$ and therefore $\bar{\mathcal{I}} = \mathcal{I}$.

Let us know check that actually there is a minimizer ν for problem (RP).

Consider a minimizing sequence $\{u_j\} \in L^\infty((0,T),K)$ for problem (P). Arguing like before (using Theorem A.2) let $\nu \in \mathcal{Y}^\infty((0,T),P(K))$ be the associated Young measure and $\{x_j\}$ the corresponding state sequence. Then

$$\mathcal{I} = \mathcal{I} = \lim_{j\to\infty} \int_0^T F(x_j(t),u_j(t))dt = \int_0^T \int_K F(x(t),\lambda)d\nu_t(\lambda)dt = \bar{I}(\nu).$$

Also we can check like before that ν verifies the differential constraints in (RP) so that ν is an optimal solution for this problem.

□

Remark A.3. *If we had considered the case with K unbounded, the convergence of*

$$\int_0^t f(x_j(\tau),u_j(\tau))d\tau$$

would require additionally conditions in order to check the equiintegrability and the equicontinuity of the sequence $\{f(x_j,u_j)\}$, essential to apply Ascoli-Arzelá theorem. See [MP01] for more details. As (x,u) must lie in the compact set $\Omega \times K$ and f is continuous those properties are trivial in our case.

Appendix B

Future Directions

B.1 On the numerical approximation for optimal control problems via a steepest descent method

B.1.1 Introduction

The main purpose of this chapter is to describe some detailed ideas for future work on numerical aspects related to optimal control problems. We would like to improve standard steepest descent methods used to approximate the solution of the following problem

(P) Minimize in u

$$\mathcal{I}(x,u) = \int_I F(t,x(t),u(t))dt$$

subject to

$$x'(t) = f(t,x(t),u(t)),\ x(t0) = x_0 \tag{B.1}$$

$$u(t) \in K \subset \mathbf{R}^n,\ t \in I$$

where $I = (t0, t1)$ and K is the set of the vectors u such that $k_i^0 \leq u_i \leq k_i^1$. We consider as an admissible pair (x,u) with $x \in AC(I,\mathbf{R}^m)$ and $u \in L^\infty(I,\mathbf{R}^n)$ where AC stands for the set of absolutely continuous functions. The explicit time dependence in F and f is included because we will analyze some examples of this kind.

Many different techniques have been used to approximate the optimal solutions of (P) (see for [Py99], [Sa00], [Sm06] and [Tr08] for detailed references). The techniques based on solving Hamilton-Jacobi-Bellman (HJB) partial differential equation

$$V_t(t,y) + H(t,y,u(t,y),V_y(t,y),-1) = 0$$

where $V(t,y)$ is the so called value function (see [Py99] and [Tr08]) and

$$H(t,y,u,p,\lambda) = p \cdot f(t,y,u) - \lambda F(t,y,u)$$

is the typical Hamiltonian, are well established for lower dimension problems. As to shooting methods, they are mainly based on solving the necessary conditions for optimality, the so called Pontryagin Maximum Principle (PMP)

$$p'(t) = -H_x(t, x(t), u(t), p(t), \lambda)$$

$$x'(t) = H_p(t, x(t), u(t), p(t), \lambda)$$

and

$$H(t, x(t), u(t), p(t), \lambda) = \max_v H(t, x(t), v, p(t), \lambda)$$

while special techniques are used to achieve both final and initial conditions on the pair of functions (x, p). These methods are known to guarantee a good precision but also to demand a previous analysis of the adjoint function p and the geometry of the problem ([Tr08]). Also indirect methods based on necessary conditions like the (PMP), require subsequent verification of the optimality for the found solution. The direct methods based on a full discretization of the problem have been recovered recently due to the last developments of computer capacities. Two approaches can be considered. One, the collocation method, consists in looking at $\mathcal{I}$ as a functional of both x and u, such that the discretized state and control will be the optimization variables in the finite dimension nonlinear problem corresponding to (P). Other, to consider x as a function of u, given by (B.1) so that the optimization variables will only correspond to u ([Sm06]). Although the implementation of these methods can be easier even with complicated constraints, the discretization tends to bring some problems in terms of accuracy ([Py99], [Tr08]). For other methods see for instance [Sm06].

Having in mind optimal control problems coming from underwater vehicles models, like problem (UVP) already described in Chapter 4, with nonlinear dependence on both state and control variables, we would like to implement an algorithm sufficiently stable to approximate optimal solutions with the state and control functions taking values in $\mathbf{R}^{12}$ and $\mathbf{R}^3$ respectively. For this purpose, we believe we should avoid either solving the partial differential equation (HJB), either the analysis required for the shooting method, or necessary optimality conditions, as they have been tested successfully on this kind of problems ([GOP09]) but shown to be quite demanding in terms of computational cost.

based on the opinions of Pytlak in [Py99], it seems that gradient (descent) methods can handle with state and control constraints quite well, and approximate solutions with the precision required for problems in engineering.

The methods based on steepest descent directions are classic and have been treated in different perspectives while complemented with other techniques. Among many references we

can cite [K60], [KKM63], [BSB63], [LMW67], [PW68], [MPD70], [JM70], [EP76], [PyV98] and [MW03].

The basic idea is the use of Gateaux's derivative whenever a gradient must be computed, which corresponds to first variation of the corresponding functional. Hence, in the spirit of standard techniques (see [MPD70], [PyV98], [MW03]), we consider an admissible pair (x, u), the original cost $\mathcal{I}(x, u)$, and we search a direction, say U, such that

$$\mathcal{I}(x_\varepsilon, u + \varepsilon(U - u)) \leq \mathcal{I}(x, u) \tag{B.2}$$

where x_ε is the solution of the initial value problem associated with the control $u_\varepsilon = u+\varepsilon(U-u)$ and and $u_\varepsilon \in K$. To this purpose, ε should lie in $(0, 1)$ and set to be sufficient small or found by one dimension minimization of the ε-parametrized cost

$$\mathcal{I}(x_\varepsilon, u + \varepsilon(U - u)).$$

Searching the direction U can be done by finding the solution of the linear optimal control problem

$$(LP) \quad \min_U \int_I \nabla F(t, x(t), u(t)) \cdot (X(t), U(t))dt$$

subject to

$$X'(t) = \nabla f(t, x(t), u(t)) \cdot (X(t), U(t) - u(t))$$

$$X(0) = 0, \quad U(t) \in K \ (\Leftrightarrow k_i^0 \leq U_i(t) \leq k_i^1) \ \forall t \in I.$$

Actually, this means that $(X, u - U)$ minimize the first variation of $\mathcal{I}$ around (x, u)

$$\int_I \nabla F(t, x, u) \cdot (X, U - u)dt$$

where X solves the linearization of the control system on the neighborhood of (x, u),

$$X' = \nabla f(t, x, u) \cdot (X, U - u).$$

The main (iterative) step of this method relates to solving (LP). This can be done via several well established techniques among those already mentioned. Either indirect methods based on the maximum principle (PMP), or going through a direct discretization. The later option leads us to solving a linear programming problem for which powerful methods are available. While seeking to make this step less heavy, we would like to try a different approach. It consists in seeing that at each step, (LP) can be written as

$$\text{Minimize in } U: \quad \int_I (a(t)X(t) + b(t)U(t))\, dt,$$

subject to

$$X'(t) = A(t)X(t) + B(t)(U(t) - u(t)), t \in I, \quad X(t0) = 0, \tag{B.3}$$

$$U(t) \in K, \quad t \in I.$$

After that, proceed to the Rockafellar's variational reformulation

$$\varphi(t, X, \xi) = \begin{cases} +\infty \quad \text{if} \quad \xi \notin A(t)X + B(t)(K - u(t)) \\ \min\limits_{U \in K} \{a(t)X + b(t)U : \xi = A(t)X + B(t)(U - u(t))\} \quad \text{else} \end{cases}$$

and consider the genuine variational problem

$$(VP) \quad \text{Minimize in } X: \quad \int_I \varphi(t, X(t), X'(t))\, dt$$

subject to $X(t0) = x_0$. The equivalence between the two problems is well established ([Ce83b], [P03]). In particular the well-posedness of (VP) is related to the fact that K in (LP) is a compact set.

One advantage of dealing with the variational reformulation is that we will only need to search for the optimal state X, while the corresponding control can be obtained after by solving the algebraic system of equations (B.3). We wonder if this option can be executed with less computational cost than via the optimality conditions (PMP) and with more accuracy than by solving the linear problem coming from the discretized (LP). Notice also that even if φ is defined through a minimum, evaluating it consists in a lower dimension linear optimization problem.

After obtaining the discretized direction U from the iterative step, we must determine ε such that (B.2) is verified. For this purpose, we must find the state x_ε corresponding to $u + \varepsilon(U - u)$, what can be done by integrating the system (B.1) using, for instance, the Runge-Kutta method of 4th order with the control approximated by either piecewise quadratic, linear or constant functions. Although this is the most common way of recovering the state, and the one we adopted here, other techniques can be used (see [MPD70]).

We would like to proceed in three stages:

- (I) Implementation of the global method with the iterative step based on direct discretization of (LP).
- (II) When the global method is considered to be stable, try to improve it by applying an appropriate method to approximate the solutions of (VP) instead of dealing with (LP). This can be done either by applying a descent method to the discretization of (VP) whose cost function takes infinite values, or using some process to make the problem smoother.

- (III) Implementation of the global method to the problem (UVP) related to the underwater vehicles model described in [GOP09].

Until the moment, we are still concluding the first stage. We have analyzed many academical examples where we tried to approximate the optimal solution with the method described above. For some of them, we also applied a direct discretization method for the original problem (P), in order to compare the solutions found. To solve the corresponding nonlinear optimization problems we used the package CONOPT3 compiled with GAMS (Generic Algebraic Modeling System).

We selected two class of vectorial problems to illustrate the results. One for a cost function with an integrand of the type

$$F(t, x, u) = \|u - g(t)\|^2 + (x - c) \cdot u$$

where $c = \int_{t_0}^{t_1} g(t)dt$, with the dynamics given by $f = u$. Another for an integrand function of the type

$$F = p_1(x - x_T) \cdot f(x, u) + p_2|u|^2$$

which measures the distance between $x(t1)$ and the desired x_T, and penalizes the use of the control accordingly to the weights p_1 and p_2.

With these examples we will see that, although we can obtain good results for some problems, when we increase the state and control dimensions, searching for a structure similar to the one of (UVP), some unexpected bad behavior occurs.

In addition we have made some primary attempts within stage (II) but we should go deeply into this subject in the future before presenting some results. Dealing with the non continuous infinite valued integrand φ in problem (VP) seems to be a major difficulty, for which we should find appropriate techniques.

In Section B.1.2 we describe in detail the general ideas we are dealing with and we state some classical concepts and results concerning steepest descent techniques. As for Section B.1.3, we turn over the question of solving problem (LP), the main iterative step. We propose an algorithm for it, either via direct discretization, or via a variational reformulation.

Section B.1.4 is devoted to put together the previous ideas and present a global algorithm to be implemented in some programming language, like Fortran 90/95. After that, in Section B.1.5 we present the results for the two class of academic examples mentioned above. Finally, we end the subject with some conclusions in Section B.1.6.

B.1.2 The general strategy: the descent method

For each given admissible u (and corresponding x) we want to find a direction U through which $\mathcal{I}$ is going to (locally) decrease.

We want to find a new control U_ε, obtained by a variation of u, parametrized by ε small enough, and the corresponding state X_ε, such that the pair $(X_\varepsilon, U_\varepsilon)$ is an admissible solution, $(U_\varepsilon)_{|\varepsilon=0} = u$ and

$$\mathcal{I}(X_\varepsilon, U_\varepsilon) \leq \mathcal{I}(x, u).$$

For this purpose we consider only the linearized variations of u

$$U_\varepsilon = u + \varepsilon(U - u) \in K$$

where $U \in K$ and $\varepsilon \in (0, 1)$. This allows us to trivially check that U_ε will an be admissible control. This means that we will search directions of the type $U - u$ instead of U.

We will look for the steepest descent direction in the following sense.

Definition B.1. *Let (u, x) be a fixed admissible solution. We say that $U(.)$ is a steepest descent direction from $\mathcal{I}(x, u)$ if there exists $s \in (0, 1)$ such that for every $\varepsilon \in (0, s)$, there is a function X_ε such that $(X_\varepsilon, u + \varepsilon(U - u))$ is an admissible solution for (P) and*

$$\mathcal{I}(X_\varepsilon, u + \varepsilon(U - u)) \leq I(Y_\varepsilon, u + \varepsilon(V - u))$$

for any other admissible solution of the type $(Y_\varepsilon, u + \varepsilon(V - u))$.

A way to achieve this purpose is to find U (in fact $U - u$) which minimizes the first variation of $\mathcal{I}$ (Gateaux's derivative in the direction $(X, U - u)$). For this purpose we assume F, and f to be of class $\mathcal{C}^1$ with respect to the state and the control. Hence we will try to minimize the quantity

$$dI(x, u).(X, U - u) =$$

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[I(x + \varepsilon X, u + \varepsilon(U - u)) - I(x, u) \right] =$$

$$\int_{t0}^{t1} \lim_{\varepsilon \to 0} \frac{F(t, x + \varepsilon X, u + \varepsilon(U - u)) - F(t, x, u)}{\varepsilon} dt =$$

$$\int_{t0}^{t1} \nabla F(t, x, u) \cdot (X, U - u) dt$$

where X satisfies $X(0) = 0$ and the linearized state system

$$X' = \nabla f(t, x, u) \cdot (X, U - u).$$

We will therefore try to solve the linear optimal control problem

$$(LP0) \quad \min_{X,U} \int_I \nabla F(t,x,u)\cdot(X,U-u)dt$$

subject to

$$X' = \nabla f(t,x,u)\cdot(X,U-u).$$

$$X(0) = 0, U \in K \Leftrightarrow k_i^0 \leq U_i \leq k_i^1$$

which is, in fact, equivalent to

$$(LP) \quad \min_{X,U} \int_I \nabla F(t,x,u)\cdot(X,U)dt$$

subject to

$$X' = \nabla f(t,x,u)\cdot(X,U-u).$$

$$X(0) = 0, U \in K \Leftrightarrow k_i^0 \leq U_i \leq k_i^1.$$

The existence of solution for this problem is given in classical results. See ([Ce83b] 10.8b,11.4c) or Theorem A.5 (autonomous case). The main tools are the linear dependence on X and U and the use of Gronwall's Lemma while checking the if the solutions X of the control system are bounded.

Formally, one should wait that if (x,u) is an admissible pair for (P) and U solves (LP), then it is a steespest descent direction for $\mathcal{I}$ from (x,u). In fact, first note that if U is an optimal control for (LP), then it is also an optimal control for $(LP0)$. Besides, consider $V(.)$ to be an admissible control such that there is a $s \in (0,1)$ and an unique absolutely continuous function Y_ε for every $\varepsilon \in [0,s]$, verifying all the constraints in (P) and the condition

$$\mathcal{I}(Y_\varepsilon, u+\varepsilon(V-u)) \leq \mathcal{I}(X_\varepsilon, u+\varepsilon(U-u)), \tag{B.4}$$

where X_ε is the unique solution of the control system in (P) associated to $u+\varepsilon(U-u)$.

As the pair $(Y_\varepsilon, u+\varepsilon(V-u))$ verifies the initial value problem (B.1) for every $\varepsilon \in [0,s]$, it should also verify the linearization of (B.1) for $\varepsilon = 0$,

$$Y' = \nabla f(t,x,u)\cdot(Y,V-u)$$

where $Y = (\frac{d}{d\varepsilon}Y_\varepsilon)_{|\varepsilon=0}$ verifies $Y(0) = 0$.

This means that (Y,V) respects the linear dynamics in problems (LP) or $(LP0)$.

Similarly, linearizing both members of inequality (B.4) around $\varepsilon = 0$ lead us to conclude that

$$\int_I \nabla F(t,x,u)\cdot(Y,V-u) < \int_I \nabla F(t,x,u)\cdot(X,U-u)$$

where $X = (\frac{d}{d\varepsilon}X_\varepsilon)_{|\varepsilon=0}$. But this contradicts the hypotesis of U being an optimal control for $(LP0)$. Consequently, we must have $U = V$ and $X = Y$.

Having in mind what we have comment above, we are going to adopt the following general strategy for approaching the solution of (P). Considering an admissible pair (x^0, u^0), we start an iteration process where at each step, from an admissible pair (x, u) fixed, the linear problem (LP) should be solved to obtain the descent direction U. As explained above this direction should be used to obtain the family of admissible controls $u + \varepsilon(U - u)$, where ε must be in $(0, 1)$, and from which the corresponding state X_ε can be obtained by integrating the control system in (P). The cost $\mathcal{I}$ computed over $(X_\varepsilon, u + \varepsilon(U - u))$ might be considered as a one dimensional cost function, which can be minimized on $\varepsilon \in (0, 1)$, or simply evaluated for ε small enough. A criterion on the distance between successive solutions, or costs, can be used for stooping the search.

This approach, which is in fact classical, requires the use of adequate numerical methods for solving each of the iterative steps described above. This is the subject we would like to explore further to understand how can we adapt the descent method to efficiently approximate the solutions of optimal control problems coming from underwater vehicles models.

B.1.3 The iteration step: the linear case

Solving the Linear Problem Directly

In order to solve Problem (LP) directly we discretize it as follows:

We assume the following:

- p the number of nodes of the partition of time interval $[t0, t1]$.
- $nstep = p - 1$ is the number of equal subintervals.
- $h = \frac{t1-t0}{nstep}$ is the size of each subinterval.
- $t = (t0 + hk, k = 0, nstep) \in \mathbf{R}^p$ is the partition (vector of p nodes) of the time interval.
- $u \in \mathbf{R}^{n\times p}$ is the value of the control function on the nodes in t, such that column u_i is given by $u(t_i)$. We define $x, X \in \mathbf{R}^{m\times p}$ and $U \in \mathbf{R}^{n\times p}$ the same way.
- We set

$$K = \prod_{i=1}^{n}[k_{i0}, k_{i1}],$$

$$a = \nabla_x F(t, x, u) \in R^{m\times p}$$

$$b = \nabla_u F(t, x, u) \in R^{n\times p}$$

and

$$A = \nabla_x f(t,x,u) \in R^{m\times m\times p}$$

$$B = \nabla_u f(t,x,u) \in R^{m\times n\times p}.$$

We want to solve problem (DLP)

$$\text{Minimize} \quad L(X,U)$$

$$= h\sum_{l=1}^{p-1}\left[\sum_{j=1}^{m}\left(\frac{a_{j,l+1}+a_{j,l}}{2}\right)\left(\frac{X_{j,l+1}+X_{j,l}}{2}\right)+\sum_{i=1}^{n}\left(\frac{b_{j,l+1}+b_{j,l}}{2}\right)\left(\frac{U_{i,l+1}+U_{i,l}}{2}\right)\right]$$

$$= \frac{h}{4}\sum_{l=1}^{p-1}\left[\sum_{j=1}^{m}(a_{j,l+1}+a_{j,l})(X_{j,l+1}+X_{j,l})+\sum_{i=1}^{n}(b_{j,l+1}+b_{j,l})(U_{i,l+1}+U_{i,l})\right]$$

subject to the constraints

$$\frac{X_{j,k+1}-X_{j,k}}{h}-\sum_{l=1}^{m}A_{j,l,k}(\frac{X_{l,k+1}+X_{l,k}}{2})-\sum_{i=1}^{n}B_{j,i,k}(\frac{U_{i,k+1}+U_{i,k}}{2}-u_{i,k})=0$$

$$\forall j\in\{1,...,m\},\quad k\in\{1,...,p-1\}$$

and

$$X_{j,1}=0 \quad \forall j=\overline{1,m}$$

$$U_{i,k}\in[k_{i0},k_{i1}] \quad \forall k=\overline{1,p},\quad \forall i=\overline{1,n}.$$

Notice that in the discretization of the cost we have approximated the integrals on each of the *nstep* intervals by a trapezoidal rule with respect to the time, state and control dependencies. The same thing can be done in the discretization of the control system.

Via variational reformulation

Here we propose a theoretical alternative approach to solve the linear problem (LP).

We have seen that the main iterative step of the descent direction method is a typical linear optimal control problem of the form

$$\min_{U} \quad \int_I (a(t)X(t)+b(t)U(t))\,dt,$$

subject to

$$X'(t)=A(t)X(t)+B(t)(U(t)-u(t)), t\in I,\quad X(t0)=0, \tag{B.5}$$

$$U(t)\in K,\quad t\in I$$

where $K\subset\mathbf{R}^n$ is a convex polytope (in particular a box) and u is a fixed measurable function with values in K.

Our treatment of this problem is based on the variational reformulation which eliminates the control in the problem and focuses on the integrand

$$\varphi(t,X,\xi)=\begin{cases}+\infty \quad \text{if} \quad \xi\notin A(t)X+B(t)(K-u(t))\\ \min\limits_{U\in K}\{a(t)X+b(t)U : \xi=A(t)X+B(t)(U-u(t))\} \quad \text{else}\end{cases}$$

and the genuine variational problem (VP)

$$\text{Minimize in } X: \quad \int_I \varphi(t,X(t),X'(t))\,dt$$

subject to $X(t0)=x_0$. The equivalence between the two problems is well established as we have seen in Chapter 2 and 3 (and in [MP01, P03, PT07]).

Following the notation above, the discretization of (VP) gives

$$(DVP)\quad \text{Minimize}\quad V(X)=h\sum_{k=1}^{p-1}\varphi\left(\frac{t_i+t_{i+1}}{2},\frac{X_k+X_{k+1}}{2},\frac{1}{h}(X_{i+1}-X_i)\right)$$

verifying the constraints

$$X_1=0,$$

where p, h, t and X_k are as in subsection B.1.3.

Notice that the computation of the φ amounts to solving a linear programming problem of low dimension which can be very efficiently solved.

In the continuous framework, for fixed $u\in\mathbf{R}^n$, $s\in\mathbf{R}$, $y\in\mathbf{R}^m$ and $\xi\in\mathbf{R}^m$, evaluating $\varphi(s,y,\xi)$ means to check if the linear problem

$$(SLP)\quad \min_{U\in K}\{a(s)y+b(s)U:\ \xi=A(s)y+B(s)(U-u)\}$$

has a solution, and set

$$\varphi(s,y,\xi)=a(s)y+b(s)\bar{U}$$

where $\bar{U}$ is such a solution, or

$$\varphi(s,y,\xi)=+\infty$$

if there is no such solution, that is, if the set defined by the constraints

$$\xi=A(s)y+B(s)(U-u)$$

$$k_i^0\le U_i\le k_i^1 \quad i\in\{1,...,n\}$$

is empty.

Remark B.1. *Notice that K being a compact set is crucial to the conclusion above.*

Solving (VP) instead of (LP) can be particularly useful when system (B.5) can be solved for U, explicitly for all times, that is the case when $m = n$, and B is constant square matrix. A simple situation illustrating this is given in example B.1.

Example B.1. *An easy example where this situation can be applied is the scalar optimal control problem where*

$$t0 = 0, \quad t1 = 1,$$

$$F = \frac{1}{2}(u-1)^2 + (x-1)u + \frac{1}{2},$$

$$f = u,$$

$$\nabla F(x,u) = (u, x+u-2),$$

$$\nabla f(x,u) = (0,1)$$

and $K = [0,2]$ *so that* φ *is given by*

$$\varphi(t,X,\xi) = \begin{cases} +\infty \quad \text{if} \quad \xi \notin K - u(t) \\ \min\limits_{U \in K}\{u(t)X + (x(t)+u(t)-2)U : \xi = U - u(t)\} \quad \text{else.} \end{cases}$$

In this situation (SLP) *has a trivial solution and* φ *is given by*

$$\varphi(t,X,\xi) = \begin{cases} +\infty \quad \text{if} \quad \xi \notin [-u(t), 2-u(t)] \\ u(t)X + (x(t)+u(t)-2)(\xi + u(t)) \quad \text{else.} \end{cases}$$

This means that φ *in* (DVP) *is explicitly found and this avoids solving* (SLP) *each time* φ *must be evaluated.*

A method to solve (DVP) must be chosen among those well fitted for finite dimension optimization problems with no smooth cost functions, in particular taking infinite values. This is an important issue on which we would like to go further. Descent methods adapted to this particular integrand type are a possibility, but more work is need before presenting any sufficiently stable approach.

After finding the optimal solution $X \in \mathbf{R}^{m\times p}$ we found the corresponding control U, such that (X,U) is an approximated optimal solution for (LP), verifying the algebraic constraints

$$\frac{X_{j,k+1} - X_{j,k}}{h} - \sum_{l=1}^{m} A_{j,l,k}\left(\frac{X_{l,k+1} + X_{l,k}}{2}\right) - \sum_{i=1}^{n} B_{j,i,k}\left(\frac{U_{i,k+1} + U_{i,k}}{2} - u_{i,k}\right) = 0$$
$$\forall j \in \{1,...,m\}, \quad k \in \{1,...,p-1\}$$

and

$$U_{i,k} \in [k_{i0}, k_{i1}] \quad \forall k = \overline{1,p}, \quad \forall i = \overline{1,n}.$$

However, in order to deal better with possible under-determinated systems, it is more secure to find U from X by solving the linear optimization problem

$$(DAP) \quad \text{Minimize in } U \quad L(X,U)$$

Subject to

$$\frac{X_{j,k+1}-X_{j,k}}{h}-\sum_{l=1}^{m}A_{j,l,k}(\frac{X_{l,k+1}+X_{l,k}}{2})-\sum_{i=1}^{n}B_{j,i,k}(\frac{U_{i,k+1}+U_{i,k}}{2}-u_{i,k})=0$$

$$\forall j\in\{1,...,m\},\quad k\in\{1,...,p-1\}$$

and

$$U_{i,k}\in[k_{i0},k_{i1}]\quad \forall k=\overline{1,p},\quad \forall i=\overline{1,n}.$$

where $X\in\mathbf{R}^{m\times p}$ is fixed.

Summarizing the method for the iterative step via variational reformulation:

- The parameters and variables are set as we did for solving directly (LP);
- From an initial admissible estimate for U we obtain X via numerical integration;
- We solve problem (DVP) using the adequate descent method. This is an $(m\times p)$ dimension, non smooth, optimization problem.
- After finding the optimal X we should solve (DAP) which is a $(n\times p)$ dimension linear optimization problem. This gives us the direction U to be used in the main iterative procedure.

B.1.4 A global implementation

In this section we will describe in detail the global descent method for finding an approximate solution for the nonlinear problem (P), while we will skip the details of the iterative step explained in the previous section.

As we have explained before our goal is to find at each step, with fixed

$$(x,u)\in AC(I,\mathbf{R}^m)\times L^\infty(I,K)$$

a steepest descent direction accordingly to which the cost $\mathcal{I}$ is going to decrease. This is done by solving the linear problem (LP) (or its equivalent reformulation (VP)) which implies solving the discretized problem (DLP). Hence, we will obtain an approximate descent direction $(U-u)$ along which we should move in order to decrease our cost. Therefore we will consider from the beginning that the desired solution should be approximated by a piecewise linear state and

a piecewise constant control considered to be good approximations for absolutely continuous and measurable functions respectively, and exactly described by finite dimension and coherent discretizations with (DLP). After finding the steepest direction, we will parameterize the cost in such a way that the parameter should tell us how much are we changing the previous control towards the direction found. This gives us a one dimension function that can also be minimized. Finding such minimizer allows us to compute the next approximated control which minimizes the variation of the cost in the descent direction. This control, together with the corresponding state obtained by numerical integration of the control system, form the approximated solution for the next iteration.

This finishes the main procedure and we should repeat it until the difference between either the consecutive costs, the state or the control are all smaller than a given precision. The method should also stop if the descent direction, solution of (DLP) is negligible.

An important question is how we integrate the control system. Notice that, when the control is given, we are dealing with initial value problems. Such problems can be solved by different methods. We have implemented the Runge-Kutta method of fourth order adapted from [PTVF96]. As for problems with both initial and final data different methods should be applied. Among them, we can apply appropriate shooting methods (see [PTVF96]).

As to the one dimensional minimization of the parameterized cost described above, although we can in fact replace it by just taking a parameter small enough, if we prefer to minimize we can use Brent's Method as in [PTVF96].

We can know summarize the method.

We start by specifying the optimal control problem by giving:

- The dimensions m for state, and n for the control.
- F , f, $\nabla_x F$, $\nabla_u F$, $\nabla_x f$ and $\nabla_u f$
- The starting point x_0
- The time interval $[t0, t1]$.
- The viability set

$$K = \prod_{i=1}^{n} [k_{i0}, k_{i1}],$$

To discretize the problem we proceed in a coherent way with what we have done in B.1.3. First, we set

- p the number of nodes of the partition of time interval $[t0, t1]$.

- $nstep = p - 1$ is the number of equal subintervals.
- $h = \frac{t1-t0}{nstep}$ is the size of each subinterval.

With this parameters we define

- $t = (t1 + hk, k = 0, nstep) \in \mathbf{R}^p$ is the partition (vector of p nodes) of the time interval.
- $u \in \mathbf{R}^{n\times p}$ is the value of the control function on the nodes in t, such that column u_i is given by $u(t_i)$. We define $x, X \in \mathbf{R}^{m\times p}$ and $U \in \mathbf{R}^{n\times p}$ in the same way.

This allows us to discretize the cost in (P)

$$DNL(x,u) = h\sum_{l=1}^{p-1} F(\frac{t_i + t_{i+1}}{2}, \frac{x_i + x_{i+1}}{2}, \frac{u_i + u_{i+1}}{2})$$

where $x_i = x(t_i)$ and $u_i = u(t_i)$.

Other necessary parameters are:

- the maximum admissible parameter $\varepsilon < 1$ for the one dimensional search,
- the precision *prec* for stopping criteria.

Finally, we need to choose:

- $u_0 \in \mathbf{R}^{np}$ the admissible discretized control to initialize the method.

After the essential data are given, and the variables discretizing the state and the control are defined, we are in condition to start the main process.

- We start by using Runge-Kutta of 4th order (function `rkdumb` from [PTVF96]) to solve the nonlinear ODE N-system and find the initial state $x_0 \in \mathbf{R}^{m\times p}$. For this purpose the control function can be approximated by a piecewise polynomial function. The method requires the derivative of the state so that we call the function `fdyn` which computes $f(t_i, x_i, u_i) \in R^m$ for every step of the integration;
- Compute $DNL(x_0, u_0)$ by calling the function `Discnlinear`;
- We set $(x, u) = (x_0, u_0)$;

Start the iterative process:

- We use all previous data to call CPLEX (via GAMS) to solve the linear problem (LP). This give us the U defining the descent direction $U - u$;

- We minimize in $s \in [0, \varepsilon]$ the parametrized cost over $(x_s, u + s(U-u))$, where x_s depends $u + s(U-u)$ via numerical integration. This parameterization is described by the function `f1dim` which is minimized through the subroutine `linmin` based on Brent's Method (combines quadratic interpolation with bisection);

- We check the stopping criteria for the newly computed $u_s = u + s(U-u)$ and respective x_s:

$$error = h\|U-u\| < prec,$$

$$error1 = \mathcal{I}(x,u) - \mathcal{I}(x_s,u_s) < prec,$$

$$error2 = h\|x_s - x\| < prec,$$

$$error3 = h\|u_s - u\| < prec,$$

  the magnitude of the descent direction and the variations in the cost function, in the state and in the control respectively. If one of the above errors is bigger than *prec* we set

$$(x,u) = (x_s,u_s)$$

  and we start another iteration.

Comparison

In order to compare our method with another well established procedure, we adopt the direct discretization of problem (P). Since we are going to analyze lower-medium dimensional examples, the discretized problem will not have more than a few thousands of variables, so that it can be solved by one of the many powerfull packages availables for mathematical programming problems.

Hence, for some examples, we will simultaneously solve the problem $(DNLP)$

$$\text{Minimize} \quad NL(x,u) = h\sum_{k=1}^{p-1} F(\frac{t_i + t_{i+1}}{2}, \frac{x_k + x_{k+1}}{2}, \frac{u_k + u_{k+1}}{2})$$

verifying the constrains

$$\frac{x_{k+1} - x_k}{h} - f(\frac{t_k + t_{k+1}}{2}, \frac{x_{k+1} + x_k}{2}, \frac{u_{k+1} + u_k}{2}) = 0$$

$$\forall\ k \in \{1, ..., p-1\}$$

$$x_1 = x_0,$$

$$u_{i,k} \in [k_{i0}, k_{i1}] \quad \forall k = \overline{1,p}, \quad \forall i = \overline{1,n},$$

where p, h, t are as above and $x_k = x(t_k) \in \mathbf{R}^m$ and $u_i = u(t_k) \in \mathbf{R}^n$.

B.1.5 Numerical experiments

Here we present some numerical results for the descent method implemented as we have described above. Such results were obtained by solving directly the linear problem (LP) to obtain the search direction U. The results obtained by the descent method are compared with the ones obtained with CONOPT3, the package that we have utilized to solve the discretized nonlinear problem $(DNLP)$. We have selected, what we expect to be some representative examples of the larger number of situations we have analyzed. In a first example good approximations of optimal solutions are shown, as it was obtained in several examples with low dimension for the state and control variables (2 and 3 dimensions). After that, another example shows some rare oscillatory behavior in the approximated control variables, by one side, and the way different methods can find different solutions, by another. Finally, when we increase the state dimension in order to analyze optimal control problems with a structure more close to the underwater vehicles problems (as in [GOP09]), lack of precision and failure arrive.

Remark B.2. *In order to use the package CONOPT3 we have compiled problem* $(DNLP)$ *with the software GAMS (Generic Algebraic Modeling System).*

A first class of examples:

We will consider examples of the type

$$F = \frac{1}{2}(u-g)^2 + (x-c)u + \frac{c^2}{2}$$

and $c = \int_0^1 f(t)dt$, where $g(t) \in [0,2], \quad \forall t \in [0,1]$, and so that $c \in [0,2]$.

Also $x(0) = x_0 = 0$, $T = 1$,

$$f(x,u) = u$$

and the admissible control set is

$$K = [0,2].$$

This problem has an optimal solution (measurable u and $x \in AC$) which can be found by using Pontryagin Maximum Principle, which for this particular type of problem (Lagrange cost function for a typical Cauchy problem) takes the form: the optimal x and u verify

$$-p = p\nabla_x f - \nabla_x \lambda_0 F$$

$$p(1) = 0 \quad p(0) = \lambda$$

$$x' = f \quad x(0) = 0$$

$$u = \arg\min_{K}(\lambda_0 F - pf)$$
$$\lambda_0 \geq 0, \quad \lambda_i^2 + \lambda^2 = 1,$$

which in our case become

$$-p' = -u \tag{B.6}$$

$$p(1) = 0 \quad p(0) = \lambda$$

$$x' = u \quad x(0) = 0 \tag{B.7}$$

$$u = \arg\min_{[0,2]}(\lambda_0(\frac{1}{2}(u-g)^2 + (x-c)u + \frac{c^2}{2}) - pu)$$

$$\lambda_0 \geq 0, \quad \lambda_i^2 + \lambda^2 = 1 \tag{B.8}$$

In fact, taking λ_0 is just to deal with possibility $\nabla_x F$ being null in the minimizer. In this case, if we take $\lambda_0 = 0$ we have

$$p = p(0) = \lambda = p(1) = 0$$

which contradicts condition (B.8). So we can take for instance $\lambda_0 = 1$. Hence

$$u = \arg\min\{\frac{u^2}{2} + (x - p - g - c)u + g^2 + \frac{c^2}{2}\}.$$

If we take the derivative of this expression and equal it to zero, we have

$$u = g + c + p - x$$

and by using equations (B.6) and (B.7)

$$u = g + c + p - (p - p(0)) = g + c + p(0).$$

So, whenever this last value belongs to $[0, 2]$, this will be the expression for the minimizer. Let us see that this is precisely the case.

If

$$p(0) + g + c < 0$$

then the minimizer is $u = 0$, and thus $p = p(1) = p(0) = 0$ so that $g + c < 0$ which cannot be. Else, if

$$p(0) + g + c > 2$$

then the minimizer is $u = 2$, and we have $p(t) = 2t - 2$ meaning that $p(0) = -2$, and consequently, due to the hypothesis on g and c,

$$p(0) + g + c \leq 2,$$

again a contradiction. We must have

$$u = g + c + p(0) \in [0, 2]$$

and consequently

$$p(t) = p(0) + \int_0^t u(t)dt = p(0) + \int_0^t (p(0) + g(t) + c)dt$$

$$= p(0) + t[p(0) + c] + \int_0^t g(t)dt.$$

Thus

$$p(1) = 0 = p(0) + p(0) + c + c$$

which implies that $p(0) = -c$, and the optimal control is

$$u = p(0) + g + c = g.$$

As a consequence, the optimal state is

$$x(t) = \int_0^t g(t)dt.$$

The minimum value of the cost function is

$$I(x, u) = \int_0^1 \frac{1}{2}(g - g)^2 + \int_0^1 (x - c)x' + \int_0^1 \frac{c^2}{2} =$$

$$\frac{1}{2}[x^2]_0^1 - c[x]_0^1 + \frac{c^2}{2} =$$

$$\frac{1}{2}x^2(1) - cx(1) + \frac{c^2}{2} = 0.$$

This class of examples can be easily extended to higher dimensional examples whenever the optimality conditions coming from the maximum principle give a system of non coupled equations, each one of them reducible to the previous situation. An example of this situation is the integrand of the cost function used in example B.2 where the cost has three different contributions of the previous type.

Example B.2. *Solve* (P) *where*

$$F = \frac{1}{2}(u_1 - t)^2 + (x_1 - \frac{1}{2})u_1 + \frac{1}{2}(u_2 - t^2)^2 + (x_2 - \frac{1}{3})u_2$$

$$+\frac{1}{2}(u_3 - 1)^2 + (x_3 - 1)u_3 + \frac{1}{8} + \frac{1}{18} + \frac{1}{2}$$

$x_0 = (0, 0, 0)$, $t0 = 0$, $t1 = 1$,

$$f(x,u) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

and the admissible control set is

$$K = [0,2]^3.$$

The optimal control and state are

$$u = \begin{pmatrix} t \\ t^2 \\ 1 \end{pmatrix}, \quad x = \begin{pmatrix} \frac{t^2}{2} \\ \frac{t^3}{3} \\ t \end{pmatrix}.$$

To approximate the solution, we implemented the descent method with a precision $prec = 1.0E-04$ *to be used on the stopping criteria, and* $p = 200$ *nodes for the partition of the time interval.*

The results, where the solutions are labeled `Desc Method`*, are shown in figures B.1-B.6. The solutions obtained with the descent method approximate very accurately either the solution obtained by CONOPT or the true exact solution, which we avoid represent for clarity. An exception is the control variable* u_3 *for which CONOPT gives a highly oscillating solution whose "average" corresponds to the exact solution, perfectly approximated with the descent method. This seems to be a consequence of the discretization used in* $(DNLP)$.

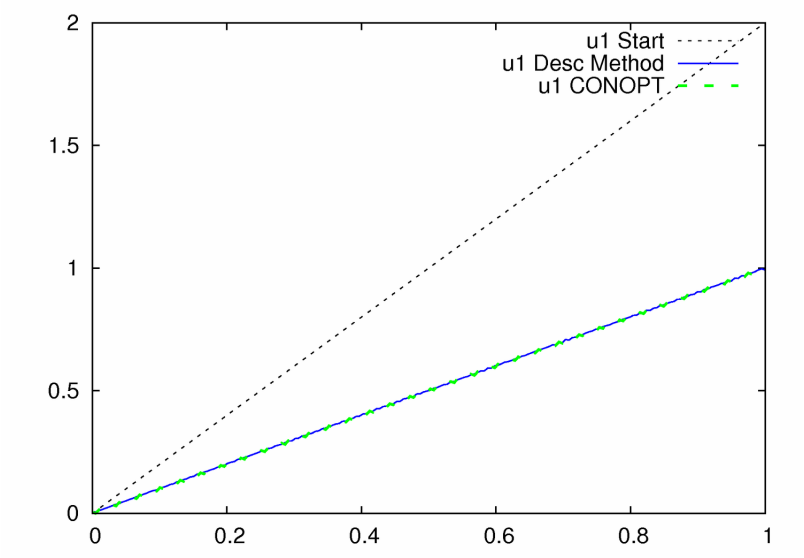


Figure B.1: Results for control variable u_1 in Example B.2

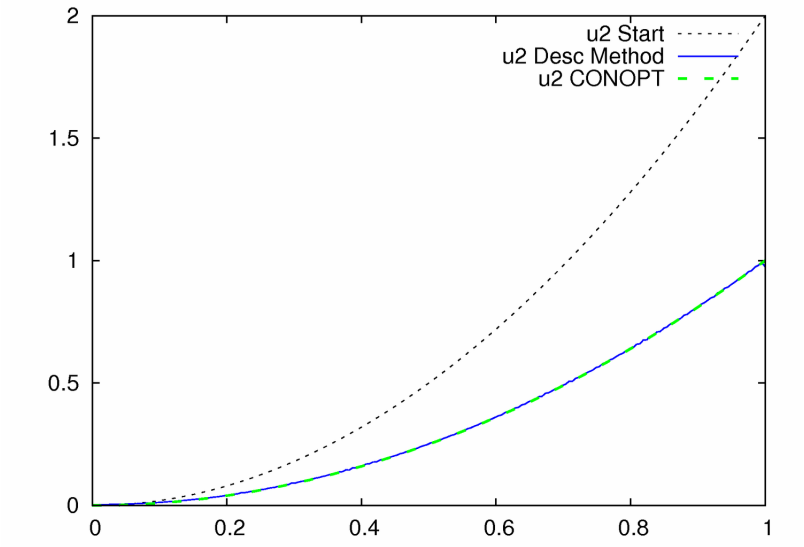


Figure B.2: Results for control variable u_2 in Example B.2

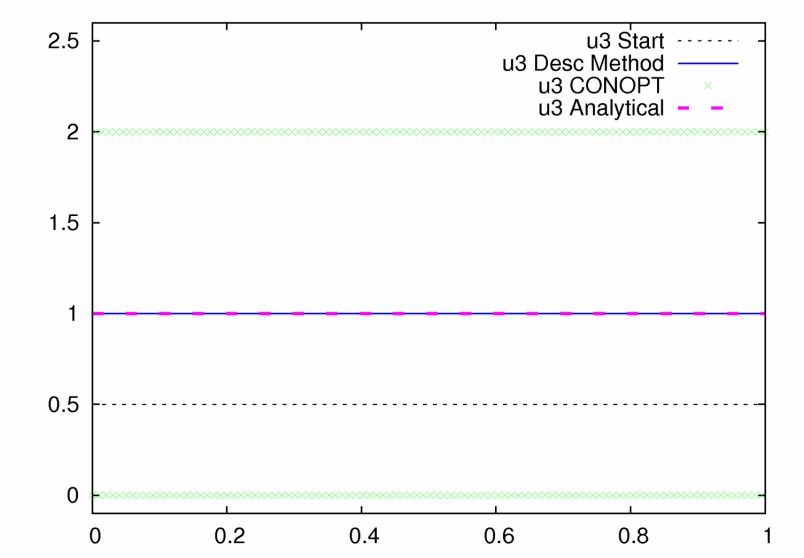


Figure B.3: Results for control variable u_3 in Example B.2

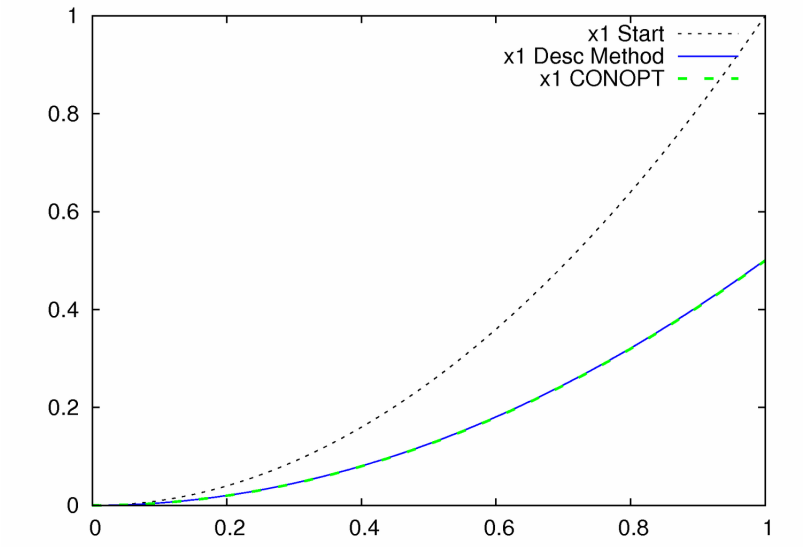


Figure B.4: Results for state variable x_1 in Example B.2

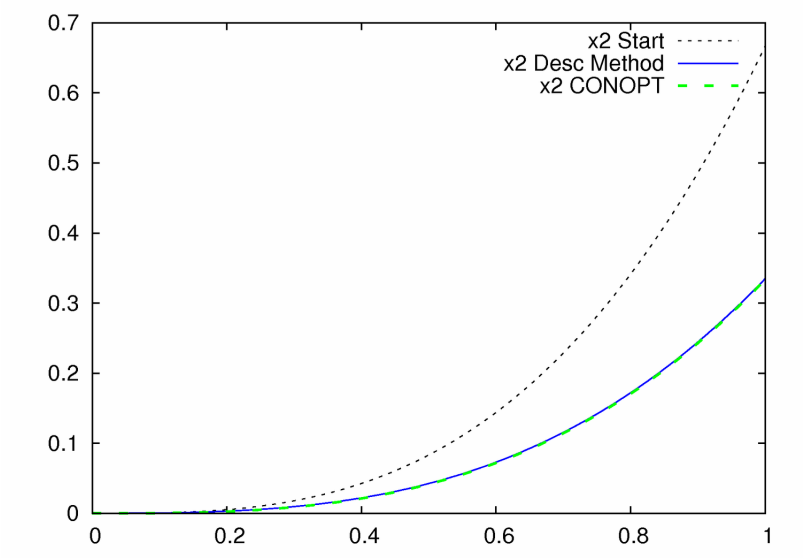


Figure B.5: Results for state variable x_2 in Example B.2

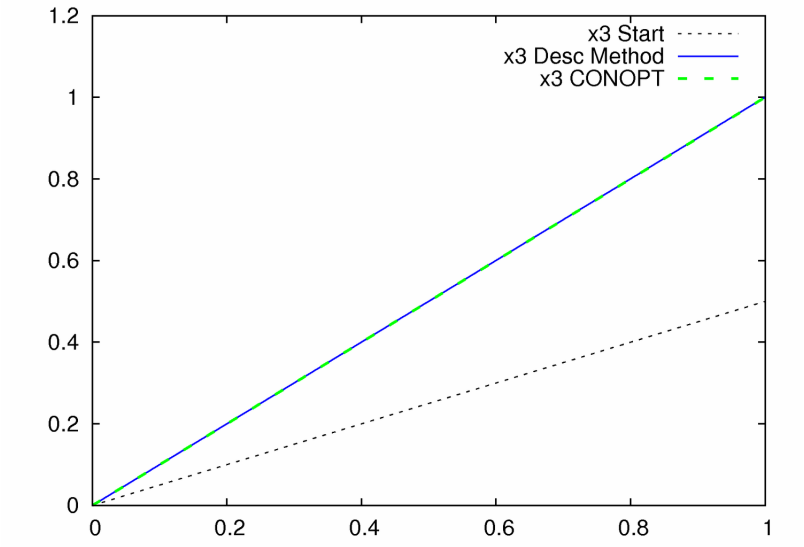


Figure B.6: Results for state variable x_3 in Example B.2

A second class of examples:

Here we are going to deal with a typical class of examples where the cost function measures the distance of the state at time $t1$ to a desired state x_T.

In fact, we want to minimize

$$\frac{1}{2}\|x(t1) - x_T\|^2 = \frac{1}{2}\int_{t0}^{t1} \frac{d}{dt}\|x(t) - x_T\|^2 dt + \frac{1}{2}\|x_0 - x_T\|^2 =$$
$$\int_{t0}^{t1} (x(t) - x_T) \cdot f(x(t), u(t))dt + \frac{1}{2}\|x_0 - x_T\|^2.$$

In addition, we consider the contribution of a penalizing term

$$\int_{t0}^{t1} \|u(t)\|^2 dt$$

weighted by a small coefficient to choose among those controls less expensive in case of non uniqueness of solution. Notice that adding such a penalizing term changes the optimal control problem, but in practice, choosing small weights for it should allow us to obtain an optimal solution close to the optimal solution of the original problem.

Example B.3. *The problem consists in trying to reach the vector*

$$x_T = (\exp(-\pi), \exp(-\pi), 1)$$

from $x_0 = (1, 1, 1)$ *in time* $T = \pi$*, where the dynamics is described by*

$$f(x, u) = \begin{pmatrix} -x_1 - 2x_2 + u_1 \\ -x_2 - x_3 + u_2 \\ -x_3 + u_3 \end{pmatrix},$$

and the admissible control set is

$$K = [\exp(-\pi), 1] \times [0, 2] \times [0, 2].$$

The cost function is the integral of

$$F = (x - x_T) \cdot f(x, u) + p_2|u|^2$$

which measures the distance between $x(T)$ *and the desired* x_T *and penalizes the use of the control accordingly to the value of* p_2*.*

In figures B.7-B.12 we can see the results for $p_2 = 0.1$*. The best results where obtained with a large number of nodes* $p > 700$*. We can see that the control variables approximated with the descent method tend to be close to the ones given by CONOPT, but they are not very accurate*

and present some strange oscillations which we couldn't avoid by changing search parameters. By contrary, the state variables seem to be very close with both methods, except for x_3 where some initial difference probably reflects the oscillating behavior of the control. Also the control approximated with the descent method is slightly more "expensive" since

$$||u \texttt{ DescMethod}||_{L^2} = 0.5862324$$

while

$$||u \texttt{ CONOPT}||_{L^2} = 0.5691115.$$

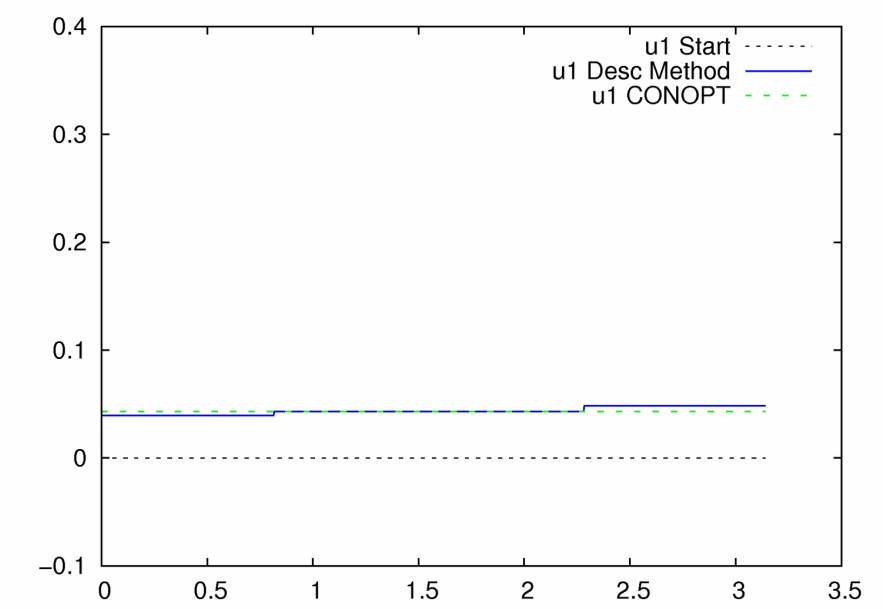


Figure B.7: Results for control variable u_1 in Example B.3 with $p_2 = 0.1$

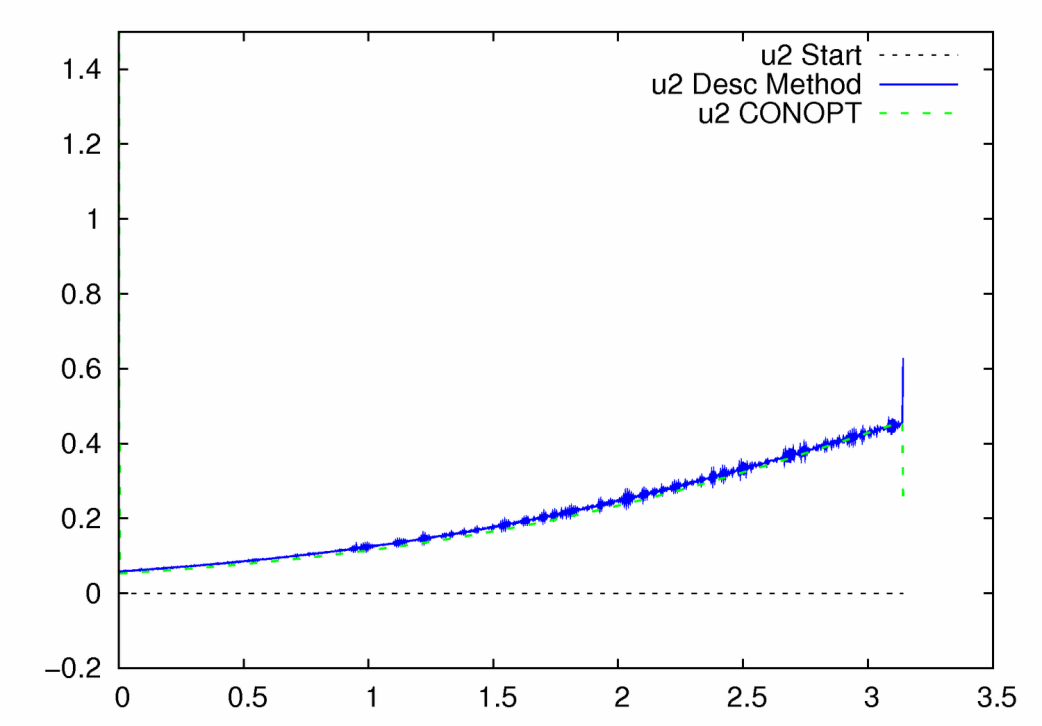


Figure B.8: Results for control variable u_2 in Example B.3 with $p_2 = 0.1$

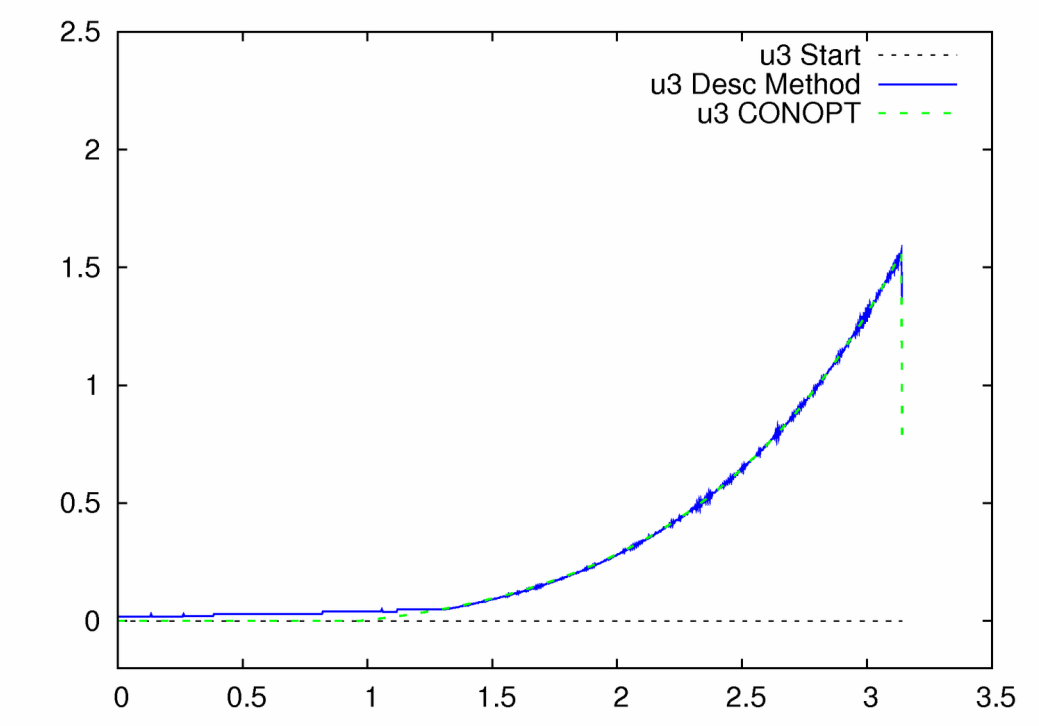


Figure B.9: Results for control variable u_3 in Example B.3 with $p_2 = 0.1$

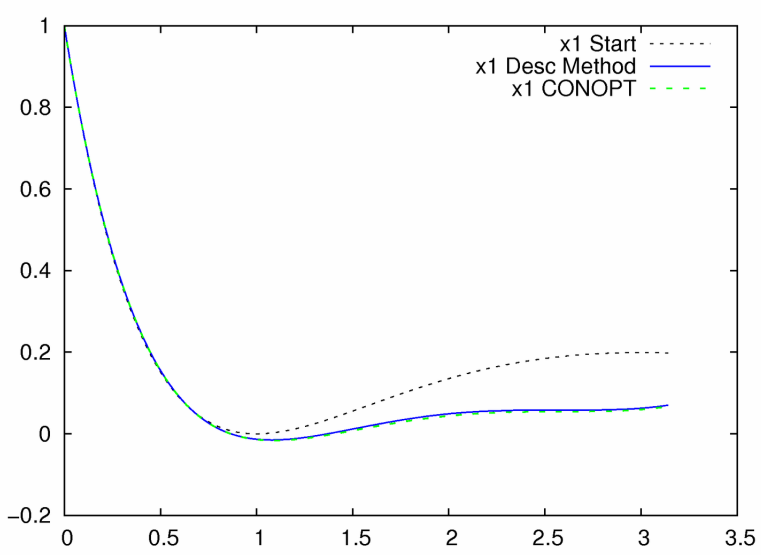


Figure B.10: Results for state variable x_1 in Example B.3 with $p_2 = 0.1$

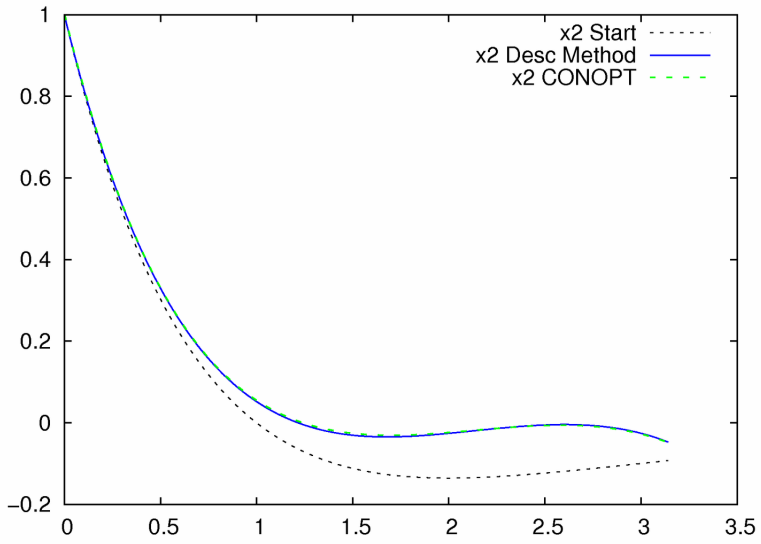


Figure B.11: Results for state variable x_2 in Example B.3 with $p_2 = 0.1$

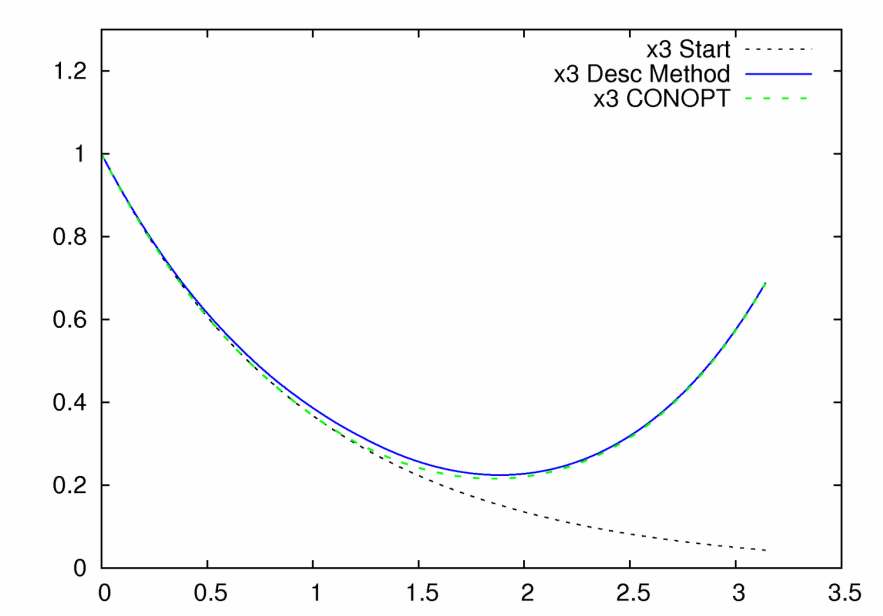


Figure B.12: Results for state variable x_3 in Example B.3 with $p_2 = 0.1$

For $p_2 = 0$ the problem doesn't have necessarily a unique solution, and actually, as we can see in figures B.13-B.18, the two methods found different solutions (which correspond to the same cost value). A remark just to say that although the state computed by the descent method seems to be more smooth, it corresponds to a more expensive control because

$$||u \texttt{ DescMethod}||_{L^2} = 2.158 \textit{ while } ||u \texttt{ CONOPT}||_{L^2} = 0.815.$$

These results are similar for a number of nodes p bigger than 500.

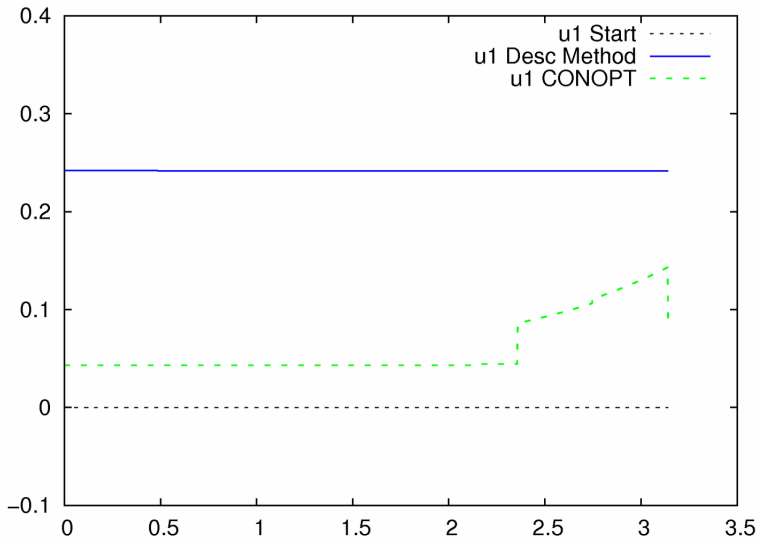


Figure B.13: Results for control variable u_1 in Example B.3 with $p_2 = 0$

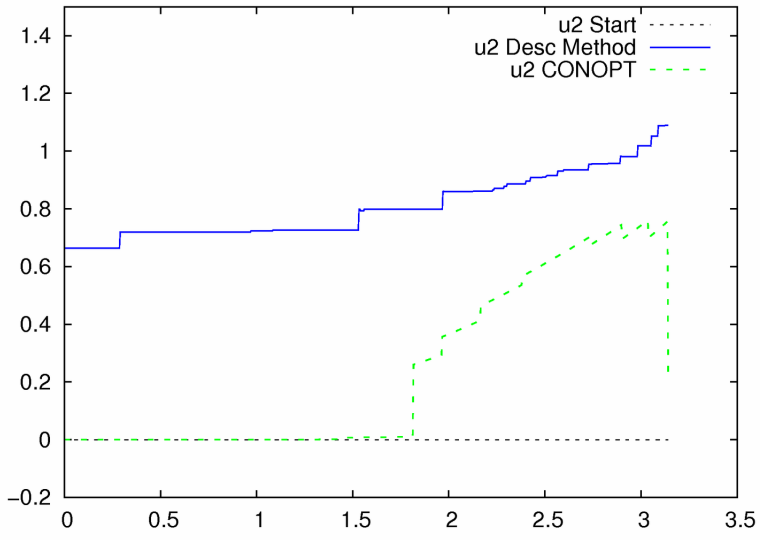


Figure B.14: Results for control variable u_2 in Example B.3 with $p_2 = 0$

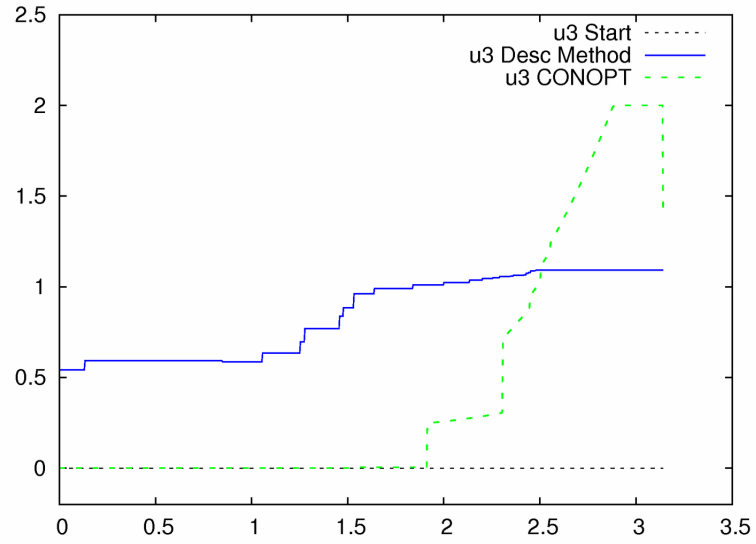


Figure B.15: Results for control variable u_3 in Example B.3 with $p_2 = 0$

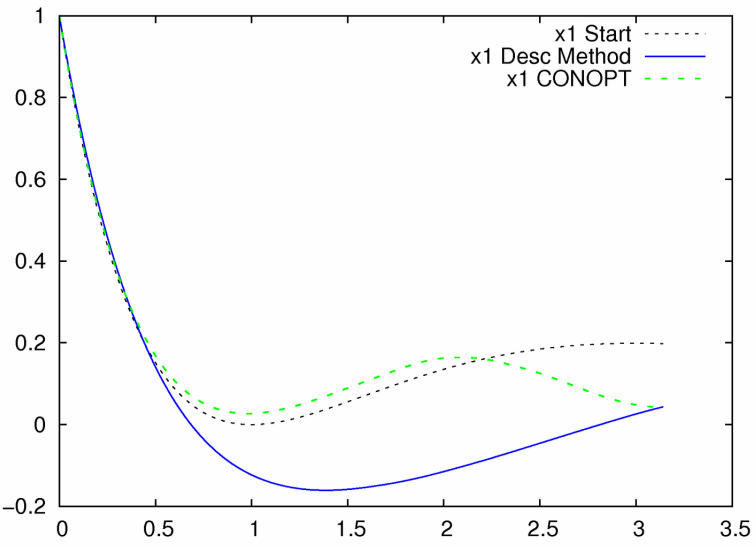


Figure B.16: Results for state variable x_1 in Example B.3 with $p_2 = 0$

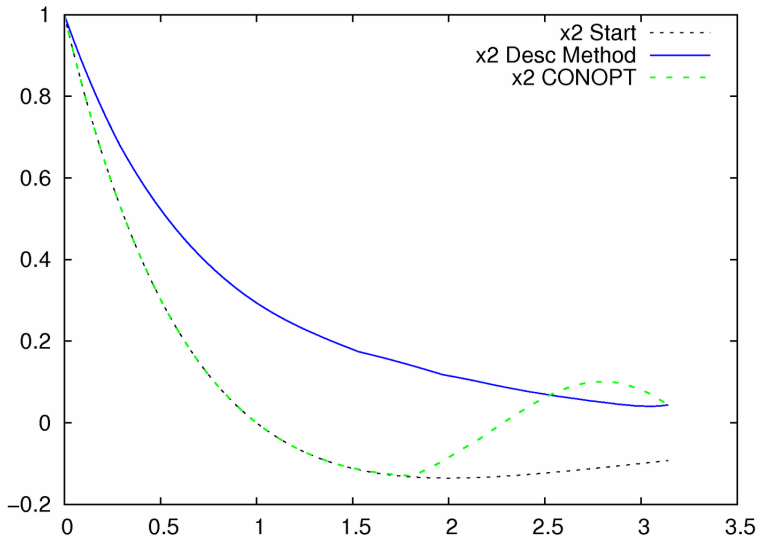


Figure B.17: Results for state variable x_2 in Example B.3 with $p_2 = 0$

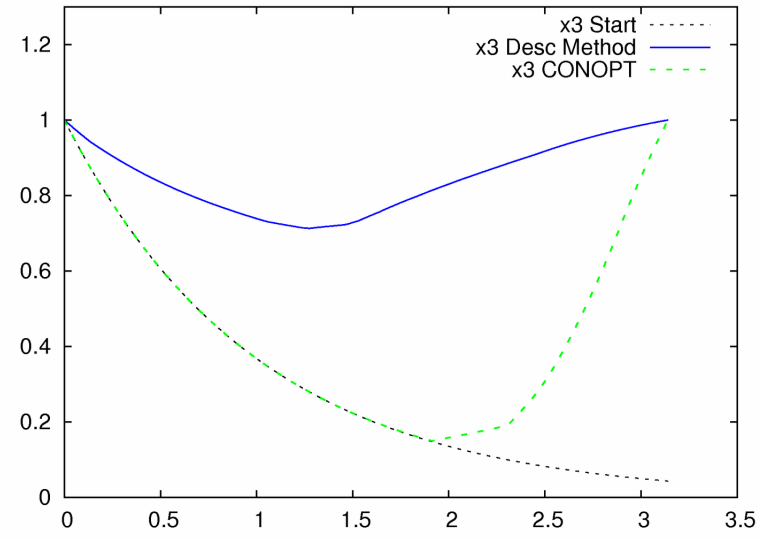


Figure B.18: Results for state variable x_3 in Example B.3 with $p_2 = 0$

Example B.4. *Try to reach the final configuration*

$$x_T = (0, 0, 0, 10)$$

from

$$x_0 = (0, 0, 0, 0)$$

in time $T = 100$*, where the dynamics is described by*

$$f(x,u) = \begin{pmatrix} -0.1828x_1 - 0.0085x_2 - 0.006u_1 - 0.0061u_2 \\ 0.084x_1 - 0.000x_2 - 0.0052u_1 + 0.0033u_2 \\ x_2 \\ x_1 - 0.1745x_3 \end{pmatrix},$$

and the control variable should lie in the admissibility set $K = [-1, 1]^2$*.*

The cost function is the integral of

$$F = (x - x_T) \cdot f(x,u) + p_2|u|^2$$

over the interval $[0, 100]$*, which like before, measures the distance between* $x(T)$ *and the desired* x_T*, and penalizes the use of the control accordingly to the value of* p_2*.*

Next we see some prints (figures B.19-B.24) for $p_2 = 0.1$*. In this particular case, and because* f *is linear, it can be seen that any two optimal solutions (thus both verifying the Maximum Principle) must be the same. However, within the small range of most of the variables, the solutions obtained by both methods are not exactly the same. This tendency seems to change when the range of the variable is bigger. That is the case of the state variable* x_4*. These results are only possible for a large number of nodes* $p \geq 500$*.*

The package CONOPT again find a less expensive control ($||u$ `CONOPT`$||_{L^2} = 0.1149597$*) than the descent method (*$||u$ `DescMethod`$||_{L^2} = 0.1191475$*).*

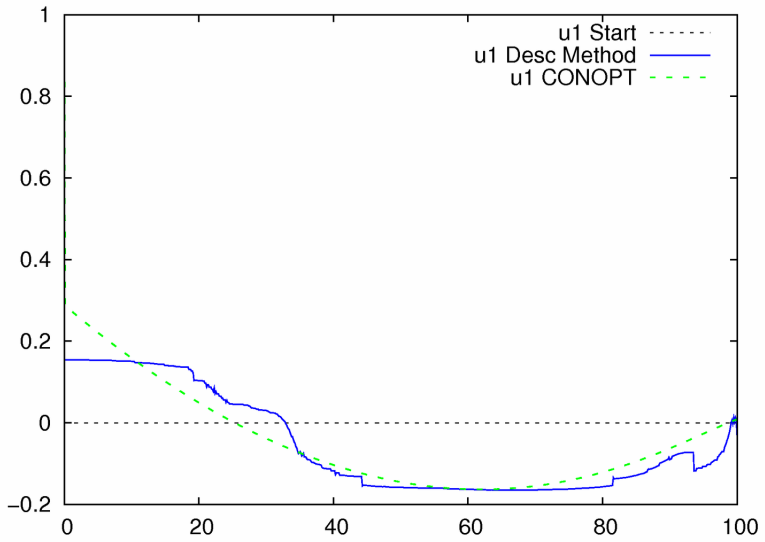


Figure B.19: Results for control variable u_1 in Example B.4 with $p_2 = 0.1$

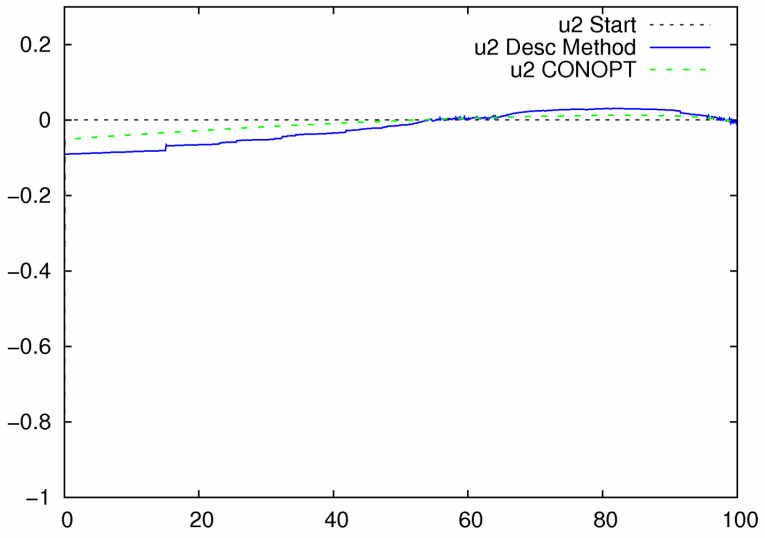


Figure B.20: Results for control variable u_2 in Example B.4 with $p_2 = 0.1$

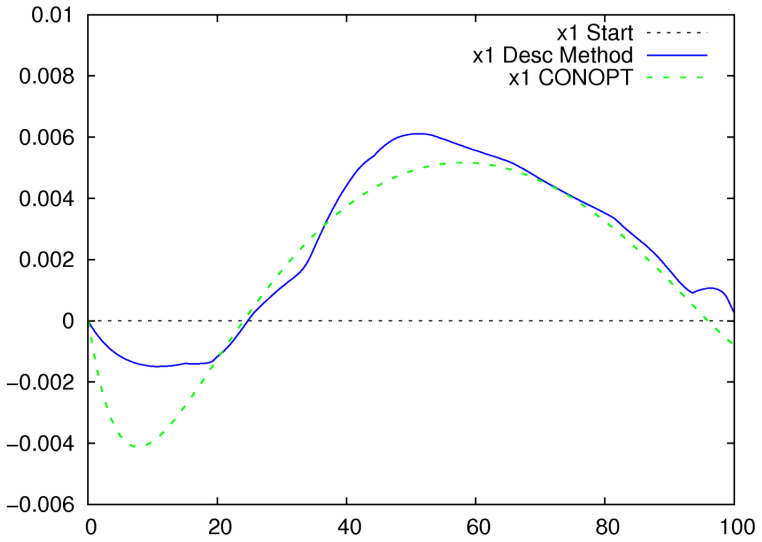


Figure B.21: Results for state variable x_1 in Example B.4 with $p_2 = 0.1$

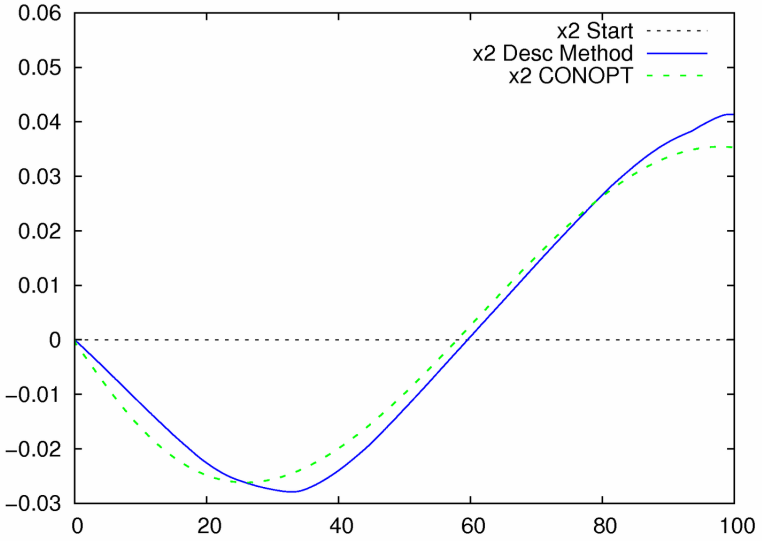


Figure B.22: Results for state variable x_2 in Example B.4 with $p_2 = 0.1$

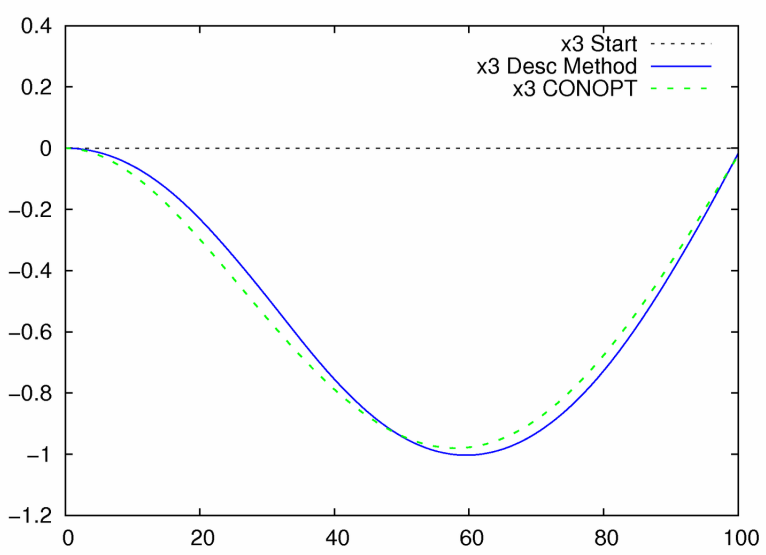


Figure B.23: Results for state variable x_3 in Example B.4 with $p_2 = 0.1$

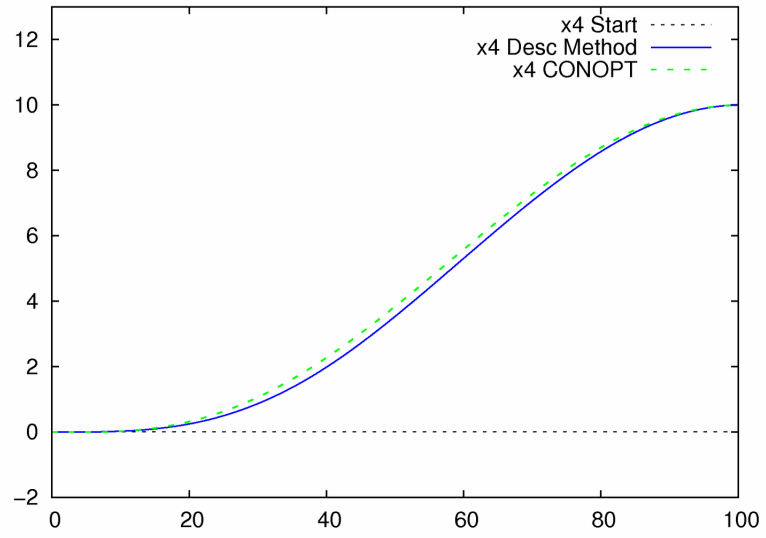


Figure B.24: Results for state variable x_4 in Example B.4 with $p_2 = 0.1$

We analyze now the results for $p_2 = 0$ *(figures B.25-B.30). Again different solutions are expected. In general we can say that a solution is approximated with the descent method, because the minimum value* (-50) *is approximated* $(-49,99873)$ *within an error smaller than* 0.0013*. We note however, that the variable* x_2 *doesn't approximate the value* $x_T^2 = 0$*, better than within an error of* 0.05*. This is because* x_2 *(and* x_1*) has an order of magnitude smaller than* x_4*. These aspects must be considered if we pretend to implement this method to more complicated problems modelling underwater vehicles. In such situation, the cost function should be adapted in order to avoid this type of anomalies. We also comment that the results improve with the number of nodes* p*. Here we considered* $p = 800$*.*

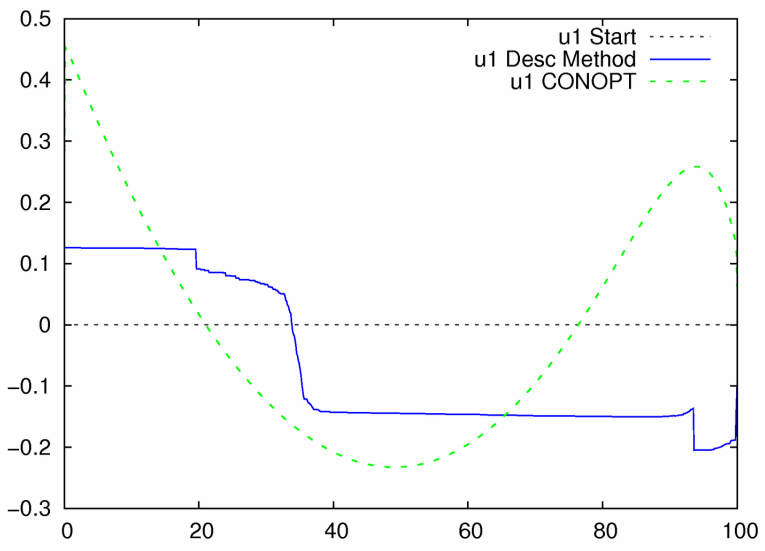


Figure B.25: Results for control variable u_1 in Example B.4 with $p_2 = 0$

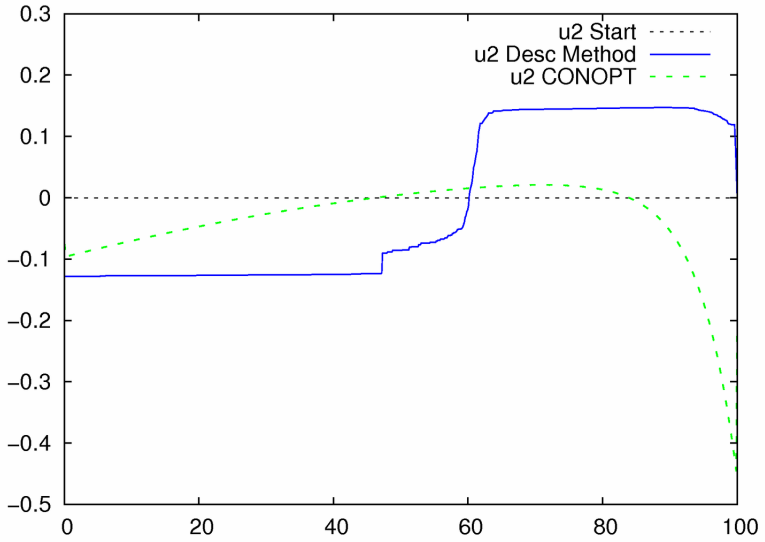


Figure B.26: Results for control variable u_2 in Example B.4 with $p_2 = 0$

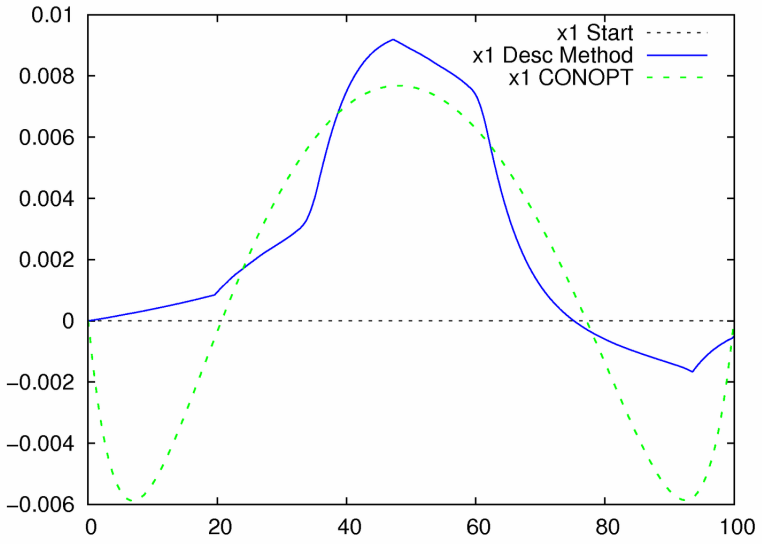


Figure B.27: Results for state variable x_1 in Example B.4 with $p_2 = 0$

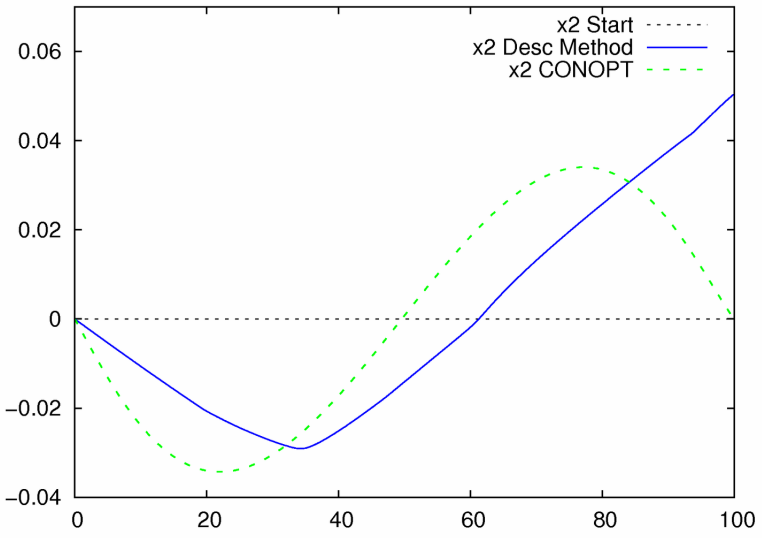


Figure B.28: Results for state variable x_2 in Example B.4 with $p_2 = 0$

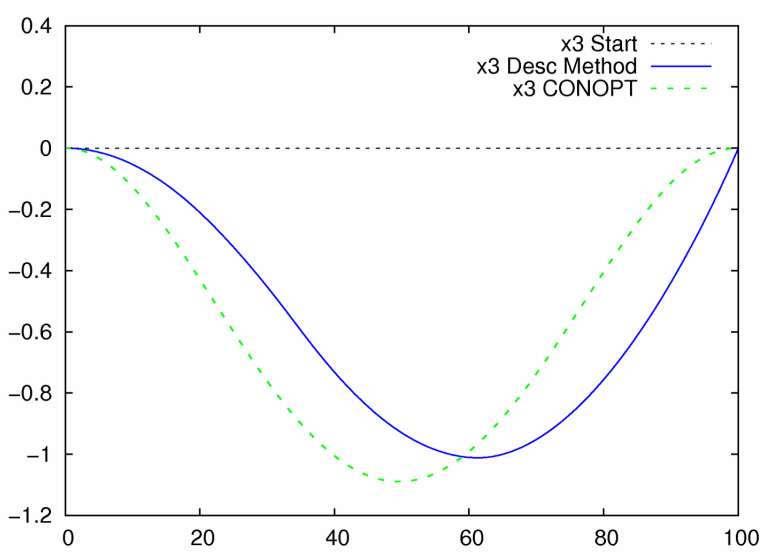


Figure B.29: Results for state variable x_3 in Example B.4 with $p_2 = 0$

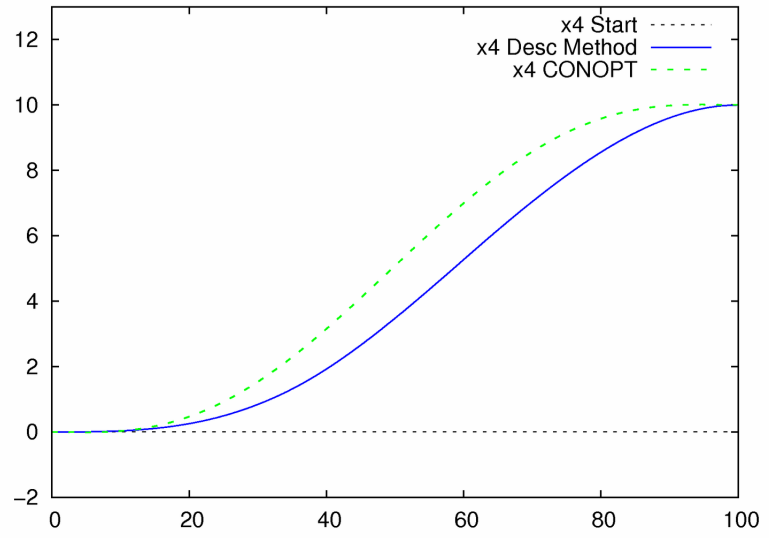


Figure B.30: Results for state variable x_4 in Example B.4 with $p_2 = 0$

B.1.6 Conclusions

As we have seen in the previous examples, the descent method like we have implemented it, is not stable yet for problems with the structure of the optimal control problem coming from the maneuvering of an underwater vehicle. Solutions can only be approximated when the time interval is partitioned with a large number of nodes, and even then, some rare oscillating behavior on the control appears. This must be a question related to the implementation of the

global method and not with the iterative step. In fact approximating the solution of problem (LP) doesn't seems to be an issue, as this is done by CPLEX, a well established package for linear optimization problems. Hence, we believe we should improve the global implementation before going further. After that, we can consider more complicated nonlinear dynamics, increase the state and control dimensions, and check the possibility of solving the variational problem (VP) instead of (LP), to search the descent directions. Finally we will be ready to implement the descent method for the problem proposed in [GOP09].

Bibliography

[Ak61] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Oliver and Boyd, Edinburgh and London (Translations from Russian, ed Moscow), 1961.

[AT79] G. Aubert and R. Tahraoui, *Théorémes dexistence pour des problémes du calcul des variations*, J. Diff. Equations, Vol. 33, (1979), pp. 1-15.

[AF90b] Aubin, J. and H. Frankowska, *Set-Valued Analysis*, Birkhauser, 1990.

[B94] E.J. Balder, *New existence results for optimal controls in the absence of convexity: the importance of extremality*, SIAM J. Control Optim., 32 (1994), pp. 890-916.

[BK90] J. M. Ball and G. Knowles, *Young measures and minimization problems of mechanics*, Elasticity, Mathematical Methods and Applications, G. Eason and R. W. Ogden, Ellis Horwood, London, England, pp.1-20,1990.

[Ba58] H. Bauer, *Minimalstellen von Funktionen und Extremalpunkte I and II*, Arch. Math., 9 (1958), pp. 389-393 and 11 (1960), pp. 200-205.

[BP01] J. Bellido and P. Pedregal, *Optimal Design via Variational Principles: The One-Dimensional Case*, Journal des Mathmatiques Pure e Appliques, Vol. 80, (2001),pp. 245-261.

[BSB63] J.V. Breakwell, J. Speyer and A.E. Bryson, *Optimization and control of nonlinear systems using the second variation*, SIAM J. Control Ser. A 1 2 (1963), pp. 193223.

[Br94] D. Brutzman, *A Virtual World for an Autonomous Underwater Vehicle*, Dissertation, Naval PostGraduated School, 1994.

[Bu87] G. Buttazzo, *Some Relaxation Problems in Optimal Control Theory*, Journal of Mathematical Analysis and Its Applications, Vol. 125, (1987), pp. 272-287.

[BGH98] G. Buttazzo, M. Giaquinta and S. Hildebrandt, *One-dimentional Variational Problems*, Oxford University Press, 1998.

[CO07] C. Carlota and A. Ornelas, *Existence of vector minimizers for nonconvex* 1*-dim integrals with almost convex Lagrangian*, J. Differential Equations, 243 (2007), pp. 414426.

[CRV04] C. Castaing, P. Raynaud de Fitte and M. Valadier, *Young Measures On Topological Spaces. With Applications in Control Theory and Probability Theory*, Mathematics and its Applications, 571. Kluwer Academic Publishers, Dordrecht, 2004.

[Ce74] L. Cesari, *An existence theorem without convexity conditions*, SIAM J. Control, 12 (1974), pp. 319-331.

[Ce83b] L. Cesari, *Optimization Theory and Applications: Problems with Ordinary differential Equations*, Springer-Verlag, Berlin, 1983.

[CC90] A. Cellina and G. Colombo, *On a classical problem of the calculus of variations*, Ann. Inst. H. Poincare Anal. Non Lineaire, 7 (1990), pp. 97-106.

[CFM06] A. Cellina, A. Ferriero and E. M. Marchini, *On the existence of solutions to a class of minimum time control problems and applications to Fermat's principle and to the brachystocrone*, Systems Control Lett., no. 2, 55 (2006), pp. 119-123.

[CO03] A. Cellina, A. Ornelas, *Existence of solutions to differential inclusions and to time optimal control problems in the autonomous case*, SIAM J. Control Optim. 42 (2003) 260265.

[Ck75] F. Clarke, *Admissible relaxation in variational and control problems*, J. Math. Anal. Appl., 51 (1975), pp. 557-576.

[EMeP03] J. J. Egozcue, R. Meziat and P. Pedregal, *From a Nonlinear, Nonconvex Variational Problem to a Linear, Convex Formulation*, Appl. Math. Optim., 47 (2003), pp. 27-44.

[EP76] E. R. Edge and W. F. Powers, *Function-Space Quasi-Newton Algorithms for Optimal Control Problems with Bounded Controls and Singular Arcs*, Journal of Optimization Theory and Applications: vol. 20, No. 4 (1976), pp. 455-479.

[ET74b] I. Ekeland and R. Temam, *Analyse convexe et problemes variationnels*, Dunod Bordas, 1974.

[Fe79] J. Felman, *Revised standard submarine equations of motion.* Report DTNSRDC/SPD-0393-09, David W. Taylor Naval Ship Research and Development Center, Washington D.C., 1979.

[F62] A. F. Filippov, *On certain questions in the theory of optimal control*, SIAM J. Control, 1 (1962) pp. 7684.

[Fs94] T. Fossen, *Guidance and Control of Ocean Vehicles*, John Wiley & Sons Ltd, UK, 1994.

[FMO98] N. Fusco, P. Marcellinia and A. Ornelas, *Existence of minimizers for some nonconvex onedimensional integrals*, Portugaliae Mathematica, 55 (2) (1998), pp. 167-184.

[GMk77] R. Gabasov and B. S. Mordukhovich, *Individual existence for optimal equations*, Soviet Math. Dokl. 15, 1974, pp. 576-581.

[Ga62] R. V. Gamkrelidze, *On sliding optimal states*, (Russian), Dokl. Akad. Nauk SSSR, 143 (1962), pp. 1243-1245.

[GOP09] J. García, D. M. Ovalle and F. Periago, *Optimal control design for the manoeuvrability of a submarine*, Preprint, Universidad Politécnica de Cartagena, (2009).

`http://www.dmae.upct.es/~fperiago/archivos_investigacion/submarine_paper`

[GH67] M. Gertler and G. R. Hagen, *Standard equations of motion for submarine simulations*, NSRDC Rep. 2510, 1967.

[HL93] A. J. Healey and D. Lienard, *Multivariable Sliding Mode Control for Autonomous Diving and Steering of Unmanned Underwater Vehicles*, IEEE Journal of Oceanic Engineering, no.3, 18 (1993), pp. 327-339.

[IT74] A. D. Ioffe and V. M. Tichomirov, *Theorie der Extremalaufgaben*, Nauka, Moscow, 1974;German transl., Deutscher Verlag der Wissenschaften, Berlin, 1979.

[JM70] D. H. Jacobson and D.Q. Mayne, *Differential Dynamic Programming*, Elsevier, New York, 1970.

[K60] H. J. Kelley, *Gradient theory of optimal flight paths*, J. Amer. Rocket Soc. 30(10) (1960), pp. 59-64.

[KKM63] H. J. Kelley, R.E. Kopp and G. Moyer, *A trajectory optimization technique based upon the theory of the second variation*, presented at the AIAA Astrodynamics Conference, New Haven, CT, 1963.

[KB76] Y. Kerbrat and J. M. Braemer, *Gometrie Des Courbes Et Des Surfaces*, Hermann, Paris, collection mthodes, 1976.

[KP94] D. Kinderlehrer and P. Pedregal, *Gradient Young measures generated by sequences in sobolev spaces*. J. Geom. Anal. 4 (1994), pp. 59-90.

[LMW67] L. S. Lasdon, S. K. Mitter and A. D. Waren, *The conjugate gradient methods for optimal control problems*. IEEE Trans. Automat Control AC12 (1967), pp. 132138.

[Ma80] P. Marcellini, *Alcune osservazioni sullesistenza del minimo di integrali del calcolo delle variazioni senza ipotesi di convessitd*, Rend. Mat., 13 (1980), pp. 271-281.

[M93] Mariconda, C., *A generalization of the Cellina-Colombo theorem for a class of non-convex variational problems*, J.Math. Ann. Appl.,Vol. 175,1993, pp. 514-522.

[MaPo80] D. Q. Mayne and E. Polak, *An exact penalty function algorithm for optimal control problems with control and terminal equality constraints, part 1 and part 2*, J. Optim. Theory Appl., 32 (1980), pp. 211246; pp. 345364.

[Ms40] E. J. McShane, *Curve-space topologies associated with variational problems*, Ann. Scuola Norm. Super. Pisa (2), 9 (1940), pp. 45-60.

[Ms67] E. J. McShane, *Relaxed controls and variational problems*, SIAM J. Control, 5 (1967), pp. 438-485.

[Me04] R. Meziat, *Analysis of Non Convex Polynomial Programs by the Method of Moments*, Frontiers in global optimization, Nonconvex Optim. Appl., 74, Kluwer Acad. Publ., Boston, MA, pp. 353-371, 2004.

[MPD70] A. Miele, R. E. Pritchard and J. N. Damoulakis, *Sequential gradient-restoration algorithm for optimal control problems*. J. Optimization Theory Appl., 5 (1970), pp. 235-282.

[MW03] A. Miele, T. Wang, *Multiple-Subarc gradient-restoration algorithm, Part 1: Algorithm Structure*, J. Optimization Theory Appl., 116 (2003), pp. 1-17.

[Mk77] B. S. Mordukhovich, *Existence of optimum controls*, J. Soviet Math, 7,1977, pp. 850-886.

[Mk99] B. S. Mordukhovich, *Existence theorems in nonconvex optimal control, in Calculus of Variations and Optimal Control*, A. Ioffe, S. Reich, I. Shafrir, eds., Chapman & Hall/CRC Press, Boca Raton, FL, 1999, pp. 173-197.

[MP01] J. Munoz and P. Pedregal, *A refinement on existence results in nonconvex optimal control*, Nonlinear Analysis, 46 (2001), pp. 381-398.

[N63] L. W. Neustadt, *The existence of optimal controls in the absence of convexity conditions*, J. Math. Anal. Appl. 7, (1963), 110-117,

[O70] C. Olech, *Integrals of set-valued functions and linear optimal control problems*, Colloque sur la Théorie Mathématique du Contrôlele Optimal, C.B.R.M., Vander Louvain, pp. 109-125, 1970.

[Or98] A. Ornelas, *Existence of scalar minimizers for nonconvex simple integrals of sum type*, J. Math. Anal. Appl. 221 (1998) pp. 559573.

[Or03] A. Ornelas, *Existence and regularity for scalar minimizers of affine nonconvex simple integrals*, Nonlin Anal 53 (2003), pp. 441-451.

[Or07] A. Ornelas, *Existence of scalar minimizers for autonomous simple integrals nonconvex except at zero*, Nonlinear Anal. (2007, in press).

[PW68] B. Pagurek and C. M. Woodside, *The Conjugate Gradient Method for Optimal Control Problems with Bounded Control Variables*, Automatica, Vol. 4, Nos. 5/6, (1968), pp. 337-349.

[P97b] P. Pedregal, *Parametrized measures and variational principles*, Progress in Nonlinear Partial Differential Equations, Birkhäuser, Basel, 1997.

[P03] P. Pedregal, *On the Generality of Variational Principles*, Milan J. Math., 71 (2003), pp. 319-356.

[PT07] P. Pedregal and J. Tiago, *A new existence result for autonomous non convex one-dimension optimal control problems*, J. Optimiz. Theory App., 134 (2007), pp. 241-255.

[PT09] P. Pedregal and J. Tiago, *Existence results for optimal control problems with some special non-linear dependence on state and control*, SIAM J. Control Opt., 48, n.2, pp. 415-437 (2009).

[PerT09] F. Periago and J. Tiago, *A local existence result for an optimal control problem modeling the manoeuvering of an underwater vehicle*, Nonlinear Analysis: Real World Applications, Accepted. Available online from September 2009. DOI: doi:10.1016/j.nonrwa.2009.09.002.

`http://dx.doi.org/10.1016/j.nonrwa.2009.09.002`

[PTVF96] W.H Press,S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical recipes in Fortran 90.The art of parallel scientific computing*, Cambridge University Press, Cambridge, 1996.

[Py99] R. Pytlak, *Numerical methods for optimal control problems with state constraints*, Springer-Verlag, Berlin, 1999.

[PyV98] R. Pytlak and R. B. Vinter, *A feasible directions algorithm for optimal control problems with state and control constrainsts: convergence analysis*, SIAM J. Control Optim., V.36, 6, (1998), pp. 1999-2019.

[Ra90] J. P. Raymond, *Existence theorems in optimal control theory without convexity assumptions*, J. Optim. Theory Appl., 67 (1990), pp. 109-132.

[Ra94] J. P. Raymond, *Existence and uniqueness results for minimization problems with nonconvex functionals*, J. Optim. Theory Appl., 82 (1994), pp. 571-592.

[Rk70] R. T. Rockafellar, *Convex Analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, New Jersey, 1970.

[Rk75] R. T. Rockafellar, *Existence theorems for general control problems of Bolza and Lagrange*, Advances in Math., 15 (1975), pp. 312-333.

[Rb96] T. Roubicek, *Relaxation of optimal control problems coercive in Lp-spaces*, K. Malanowski, Z. Nahorski, M. Peszynska (Eds.), Proceedings of the IFIP WG7.2 on Modelling and Optimization of Distributed Parameter Systems with Applications to Engineering, June 1995, Chapman and Hall, London, pp. 270277, 1996.

[Rb97b] T. Roubicek, *Relaxation in Optimization Theory and Variational Calculus*, W. De Gruyter, Berlin, 1997.

[RbS97] T. Roubicek and W. H. Smith, *Existence of solutions of certain nonconvex optimal control problems governed by nonlinear integral equations*, Optimization, no. 2, 42 (1997), pp. 91-108.

[R62] T. Roxin, *The existence of optimal controls*, Michigan Math. J., 9 (1962), pp. 109-119,

[Ru87b] Rudin, W., *Real and Complex Analysis*, Macgraw-Hill, 1987.

[Sa00] R. W. H. Sargent, *Optimal Control*, Centre for Process Systems Engineering, Imperial College, Prince Consort Road, London SW7 2BY, UK, 2000 (available online).

[ST70] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, Mathematical Surveys, n. 1, American Mathematical Society, Providence, RI, 1970.

[Sm06] W. H. Schmidt, *Numerical Methods for Optimal Control Problems with ODE or Integral Equations*, Large-Scale Scientific Computing ,Lecture Notes in Computer Science, Springer Berlin, 2006.

[S90] E. D. Sontag, *Mathematical control theory*, Texts in Applied Mathematics 6, Springer-Verlag, 1990.

[Ta79] L. Tartar, *Compensated compactness and applications to partial differential equations.* Nonlinear Analysis and Mechanics (R.J.Knops, ed.) Heriott-Watt Symposium IV, Pitman Res. Notes in Math. 39, San Francisco, 1979.

[Ti04] J. Tiago, *Controlo Óptimo Segundo Princípios Variacionais*, Tese de Mestrado, Universidade de Évora, 2004.

[To14] Tonelli, L. *Sur une methode directe du calcul des variations*, C. R. Acad. Sci. Paris, 158, (1914), 1776-1778, 1983-1985 = Oere Scelte 3, 177-182,

[Tr08] E. Trélat, *Contrôle Optimal : théorie et applications*, Vuibert, Paris, 2008.

[W62] J. Warga, *Relaxed variational problems*, J. Math. Anal. Appl., 4 (1962), pp. 111-128.

[W82] J. Warga, *Iterative procedure for constrained and unilateral optimization problems*, SIAM J. Control Optim., 20 (1982), pp. 360376.

[Y37] L. C. Young, *Generalized curves and the existence of an attained absolute minimum in the calculus of variations*, C.R. Soc. Sci. Lett. Varsovie Classe III, 30 (1937) pp. 212-234.

[Y42] L. C. Young, *Generalized surfaces in the calculus of variations, I and II*, Ann. Math., 43 (1942), pp. 84-103 and pp. 530-544.

[Y69] L. C. Young, *Lectures on Calculus of Variations and Optimal Control Theory*, W.B. Saunders, Philadelphia, 1969.