



DEPARTAMENTO DE
ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES

**NOVO PRINCÍPIO DO MÁXIMO
PARA PROBLEMAS DE CONTROLO
ÓPTIMO COM RESTRIÇÕES DE
ESTADO**

Maria Margarida de Amorim Ferreira

FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO
Rua dos Bragas 4099 Porto Codex - PORTUGAL

677

NOVO PRINCÍPIO DO MÁXIMO PARA PROBLEMAS DE CONTROLO ÓPTIMO COM RESTRIÇÕES DE ESTADO

Maria Margarida de Amorim Ferreira

621.3(043)Ferreira/Nov

UNIVERSIDADE DO PORTO
Faculdade de Engenharia
Distrito
N.º 20591-01-1
CDU
Data 1 6 95

p. n.º 6215

DISSERTAÇÃO DE DOUTORAMENTO APRESENTADA À
FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO
DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA E
COMPUTADORES

PORTO - OUTUBRO DE 1994

043 D
f 442 m
ex. 1

Ao Zé

Agradecimentos

Gostaria de agradecer aos meus supervisores Prof. F. Lobo Pereira e Prof. R. B. Vinter, pelo seu apoio, sugestões e paciência.

Agradecimentos são também devidos à JNICT, cujo apoio financeiro foi essencial.

Não posso ainda deixar de agradecer aos meus colegas, em particular, ao grupo do Imperial College.

A nível pessoal, aqui deixo um muito obrigado aos meus pais, família e amigos. Também ao Zé, cujo afecto e dedicação foram essenciais na elaboração deste trabalho.

Resumo

O Princípio do Máximo estabelece um conjunto de condições necessárias, a serem satisfeitas pela solução ótima do problema. Estas condições envolvem um conjunto de multiplicadores que têm como objectivo caracterizar a solução. Em determinados problemas de controlo ótimo com restrições, é possível definir à partida um conjunto de multiplicadores que, satisfazendo trivialmente as condições necessárias, não fornecem qualquer informação sobre a solução ótima do problema. Neste trabalho, é apresentada uma versão do Princípio do Máximo que evita tais multiplicadores mediante a satisfação de uma determinada condição, designada por "qualificação das restrições". Neste contexto, são estudados três tipos de problemas que, de acordo com as suas características, são abordados de maneira diferente. Assim, é tratado inicialmente um problema de controlo ótimo com restrições de estado, incluindo ainda restrições de desigualdade sobre o estado no instante final e satisfazendo propriedades mais ou menos regulares. O mesmo problema, mas com restrições de igualdade sobre o estado no instante final, é estudado em seguida através de uma nova abordagem. Finalmente, consideramos um problema satisfazendo propriedades mais fracas, através da consideração da não diferenciabilidade das funções intervenientes. Neste caso, são utilizados métodos de cálculo não diferenciável.

Índice

1	Introdução	7
2	Preliminares	10
2.1	Convergência forte, fraca e fraca*	11
2.2	Convexidade	15
2.3	Equações Diferenciais	16
2.4	Multifunções e Trajectórias	18
2.5	Cálculo não Diferenciável	21
3	Problema de Controlo Óptimo com Restrições	25
3.1	Problema de Controlo Óptimo	26
3.2	Princípio do Máximo	29
4	Princípio do Máximo não Trivial	32
4.1	Condições Necessárias não triviais	33
4.2	Demonstração dos Teoremas 4.1.1 - 4.1.4	36
4.2.1	Resultados Preliminares	36
4.2.2	Demonstração dos Teoremas 4.1.1 e 4.1.3	45
4.2.3	Demonstração dos Teoremas 4.1.2 e 4.1.4	54
5	Condições Necessárias no Cálculo Variacional	57
5.1	Problema do Cálculo Variacional	58

5.2	Condições Necessárias para um Problema do Cálculo Variacional . . .	60
6	Restrições de Igualdade no Estado Final	67
6.1	Condições Necessárias não triviais	68
6.2	Demonstração dos Teoremas 6.1.1 e 6.1.2	70
6.2.1	Resultados Preliminares	70
6.2.2	Demonstração do Teorema 6.1.1	76
6.2.3	Demonstração do Teorema 6.1.2	82
7	Problema de Controlo Óptimo Não Regular	84
7.1	Princípio do Máximo	85
7.2	Demonstração dos Resultados Principais	87
8	Conclusões	93

Capítulo 1

Introdução

A Teoria de Controlo Óptimo como moderna sucessora do Cálculo das Variações é a sua "herdeira" natural no tratamento de um elevado número de aplicações clássicas, desde os problemas isoperimétricos da antiguidade, passando pelos princípios variacionais dos séculos dezoito e dezanove.

Na tentativa de dar resposta aos problemas que surgiram com a evolução de áreas como a Engenharia, Robótica, Aeronáutica, Controlo de Processos, Economia etc, tem-se desenvolvido largamente, tornando-se uma ferramenta indispensável.

Neste contexto está a derivação de condições necessárias de optimalidade, sob a forma do Princípio do Máximo, para problemas com restrições de estado. De facto, restrições de estado surgem frequentemente associadas a problemas reais e a sua inclusão nas condições necessárias básicas permitiu que a Teoria de Controlo Óptimo pudesse ser aplicada a uma maior gama de problemas.

Tendo em vista o tipo de aplicação para o qual é dirigido e as propriedades dos elementos envolvidos, o Princípio do Máximo com restrições de estado aparece sob formas diversas. Foi estudado pela primeira vez por Pontryagin e al. [20] e versões posteriores foram provadas por Gamkrelidze, Neustadt, Warga e outros [14], [19], [27].

Os anos mais recentes continuaram a ser testemunhas de muitos resultados nesta área. Num ambiente de cálculo não diferenciável, foram derivadas condições necessárias de optimalidade para problemas em que as funções intervenientes eram não regulares [8]. Problemas de tempo final livre e problemas com restrições diversas de ponto final são também exemplos de problemas com restrições de estado largamente estudados [8], [7], [9], [22], [26].

Condições necessárias de optimalidade, sob a forma do Princípio do Máximo, para problemas de Controlo Óptimo com restrições de estado são ainda objecto de estudo neste trabalho. Em particular, são obtidas condições necessárias cuja não trivialidade é garantida. Observando o Princípio do Máximo para problemas deste tipo e em que o estado inicial está sobre a fronteira das restrições, é possível definir logo à partida um conjunto de multiplicadores que, satisfazendo trivialmente as condições necessárias, não fornece qualquer tipo de informação para o problema. O objectivo, neste trabalho, será apresentar hipóteses sobre os dados do problema, a que chamaremos *qualificação das restrições*, de forma a garantir a existência de multiplicadores não triviais para além dos triviais mencionados. Note-se que, não se trata

de assegurar que todos os multiplicadores sejam não triviais, tradicionalmente tem sido esse o papel desempenhado pelas qualificações das restrições, mas sim assegurar a existência de um não trivial.

Condições para a não degeneração do Princípio do Máximo restrito foram também estabelecidas por Arutyunov e Tynyanskiy, [2]. A abordagem efectuada por estes autores, baseada em técnicas de penalização, é contudo diferente da efectuada neste trabalho e os resultados não são directamente comparáveis. A qualificação das restrições por eles apresentada, expressa em termos de aproximações canónicas de primeira ordem ao conjunto de restrições do controlo, é bastante mais conservativa.

A tese está organizada da seguinte forma. No Capítulo 2 são introduzidas algumas noções e resultados básicos que serão utilizados nos capítulos posteriores. Uma descrição do Problema de Controlo Óptimo Restrito e discussão das condições necessárias de optimalidade na forma do Princípio do Máximo são apresentadas no Capítulo 3. É considerada a situação em que há degeneração dos multiplicadores e apresentado um problema tipo em que tal situação ocorre.

O Capítulo 4 inclui a principal contribuição, sob o ponto de vista de resultados, desta tese. Estabelece-se uma versão não trivial do Princípio do Máximo, mediante a satisfação da já mencionada qualificação das restrições. Variantes desta condição são ainda consideradas no mesmo capítulo. Em determinados problemas é fácil a verificação de tal condição. No Capítulo 5 é exposto um problema do Cálculo Variacional em que isso acontece.

A existência de restrições de igualdade no estado final impede uma abordagem do tipo considerado no capítulo 4. O estudo deste problema é efectuada no Capítulo 6, onde outras técnicas de abordagem são utilizadas.

Finalmente e no Capítulo 7, são consideradas hipóteses mais fracas sobre as funções intervenientes no problema inicial e é num ambiente de cálculo não diferenciável que o estudo do problema se vai processar.

É apresentado ainda um Capítulo 8 onde são revistos os principais resultados da tese e são expostas algumas ideias de trabalho futuro nesta área.

Capítulo 2

Preliminares

Neste capítulo são apresentadas algumas noções e resultados técnicos que servirão de base a vários desenvolvimentos efectuados nos capítulos posteriores.

Começamos por descrever as noções de convergência *fraca* e *fraca** em espaços topológicos. Tal descrição e o desenvolvimento de resultados associados podem ser encontrados em diversos autores dos quais referimos Luemberger [17] e M. Reed [21]. Ekeland e Temam, [13], fazem um estudo detalhado sobre convexidade. Do trabalho destes autores e de Warga [27], foram extraídos os resultados que apresentamos referentes a este assunto. Foi ainda Warga [27], a fonte utilizada para obter o teorema 2.3.1 sobre equações diferenciais.

Finalmente, Multifunções e Trajectórias e ainda o desenvolvimento do Cálculo Não Diferenciável foram desenvolvidos por Clarke, [8].

A menos que algo seja dito em contrário, $|\cdot|$ ou $\|\cdot\|$ em $\mathbb{R}^n$, representam a norma euclidiana no respectivo espaço. Em espaços normados mais gerais, $\|\cdot\|$ representa a norma definida no espaço.

2.1 Convergência forte, fraca e fraca*

Seja X um espaço vectorial linear e normado.

O *espaço dual* de X , que representamos por X^* , é o espaço de todas as funcionais lineares, limitadas, definidas em X . A norma de um elemento $f \in X^*$ é definida como

$$\|f\|_* = \sup_{\|x\| \leq 1} |f(x)|.$$

onde $\|x\|$ é calculada de acordo com a norma definida em X .

Munido desta norma, o espaço X^* é completo. Sendo um espaço normado completo, X^* é, por definição, um espaço de Banach.

Os conjuntos abertos de $\mathbb{R}$ auferem-lhe uma estrutura particular. Algumas das propriedades destes conjuntos postas sob a forma de axiomas, permitem definir uma estrutura análoga para espaços mais gerais.

Seja X um conjunto e $\mathcal{T}$ uma família de subconjuntos de X .

$(X, \mathcal{T})$ diz-se um espaço topológico se as propriedades seguintes são satisfeitas:

- (1) O conjunto vazio, que representamos por $\emptyset$, e X pertencem a $\mathcal{T}$;

- (2) A união de qualquer número de elementos de $\mathcal{T}$ está também em $\mathcal{T}$;
 (3) A interseção de um número finito de elementos de $\mathcal{T}$ é ainda um elemento de $\mathcal{T}$.

$\mathcal{T}$ diz-se uma *topologia* no conjunto X e os seus elementos são designados por *abertos*. A noção de conjunto aberto induz, de uma forma análoga ao caso real, as definições de conjunto fechado, vizinhança, continuidade, convergência, etc. Um conjunto fechado será um conjunto cujo complementar é um aberto. Uma sucessão $\{x_i\}$, de elementos de X , convergirá para o limite x_0 , se dada uma vizinhança V de x_0 , ou seja um aberto que contenha x_0 , existe um inteiro positivo N tal que, $x_n \in V$ para todo $n \geq N$.

A forma mais usual de definir uma topologia num dado espaço X é através de uma *base*. Define-se base como uma família, $\mathcal{B}$, de subconjuntos de X que satisfaz:

- (1) $\forall x \in X \exists B \in \mathcal{B} : x \in B$;
 (2) Se $x \in B_1 \cap B_2$, onde B_1 e B_2 são elementos de $\mathcal{B}$, então existe $B_3 \in \mathcal{B}$ tal que $B_3 \subset B_1 \cap B_2$ e $x \in B_3$.

Dada uma base $\mathcal{B}$, seja $\mathcal{T}$ a família de todos os conjuntos $T \subset X$ com a propriedade:

$$\forall x \in T \exists B \in \mathcal{B} : B \subset T \text{ e } x \in B.$$

É fácil de verificar que $\mathcal{T}$ assim definida, é uma topologia em X . Neste caso, $\mathcal{B}$ diz-se uma *base para a topologia* $\mathcal{T}$ e $\mathcal{T}$ será uma topologia gerada por $\mathcal{B}$.

Se o espaço X é um espaço de Banach e $\|\cdot\|$ a norma definida no espaço, os conjuntos

$$\{x \in X : \|x - x_0\| < r\}, \quad r > 0 \quad x_0 \in X,$$

designados por bolas abertas de raio r e centro x_0 , geram uma topologia em X . Esta topologia é designada por *topologia forte*. Sendo a mais usada em tais espaços, não é no entanto a única. Em muitas situações é necessário trabalhar com estruturas que forneçam maior espaço de manobra, topologias munidas de menos conjuntos abertos e, por isso, onde é possível garantir mais sucessões a convergir, etc.

São definidas em seguida duas topologias nestas condições. A partir de um espaço linear normado X , toma-se o seu dual X^* e define-se a *topologia fraca* em X em termos de X^* . A mesma técnica pode ser aplicada a X^* , sendo neste caso a topologia fraca definida em termos de X^{**} , dual de X^* . É possível no entanto definir uma topologia mais interessante em X^* , em termos de X , a *topologia fraca** (leia-se, fraca estrela).

No que se segue, a expressão $\langle x^*, x \rangle$ representa o valor da função x^* em x .

Definição 2.1.1 *Seja X um espaço linear normado. A topologia fraca em X é a topologia gerada pela base constituída por todos os conjuntos da forma*

$$\{x \in X : \langle x^*, x \rangle < \epsilon \text{ se } x^* \in A^*\}$$

onde A^ é um subconjunto finito de X^* e $\epsilon > 0$.*

Definição 2.1.2 *Seja X^* um espaço de Banach, dual do espaço X . A Topologia fraca*, em X^* , é a topologia gerada pela base constituída por todos os conjuntos da forma*

$$\{x^* \in X^* : \langle x^*, x \rangle < \epsilon \text{ se } x \in A\},$$

onde A é um subconjunto finito de X e $\epsilon > 0$.

A partir destas topologias definem-se conjuntos *fracamente* e *fracamente** fechados, como sendo conjuntos fechados na topologia *fraca* e *fraca**, respectivamente. De forma análoga se definem as noções de compacticidade *fraca* e *fraca**, continuidade *fraca* e *fraca**, convergência *fraca* e *fraca**, etc.

Assim, uma sucessão $\{x_i\}$ de elementos de X é *fracamente* convergente para $x_0 \in X$, se e só se para qualquer $x^* \in X^*$, a sucessão de números reais $\langle x^*, x_i \rangle$ converge para o número real $\langle x^*, x \rangle$.

Da mesma forma, uma sucessão $\{x_i^*\}$ de elementos de X^* é *fracamente** convergente para $x_0^* \in X^*$, se e só se para qualquer $x \in X$, a sucessão $\langle x_i^*, x \rangle$ converge para $\langle x_0^*, x \rangle$.

Em X^* ficam assim definidas três tipos de convergência diferentes: *forte*, *fraca* e *fraca**. Pode ser verificado que a convergência *forte* implica a convergência *fraca* e esta implica a *fraca**.

Um subconjunto $K \subset X^*$ será *fracamente** compacto se toda a sucessão de elementos em K contém uma subsucessão *fracamente** convergente.

O resultado que a seguir apresentamos revela-se de especial importância na teoria de existência de solução em problemas de optimização, quando o espaço das variáveis é de dimensão infinita.

Teorema 2.1.1 *Seja X um espaço linear normado, real. A esfera unitária, fechada, em X^* é um conjunto fracamente* compacto.*

Tome-se o espaço das funções contínuas, definidas no intervalo $[a, b]$ e tomando valores em $\mathbb{R}$. Munido da norma $\|x\| = \max_{a \leq t \leq b} |x(t)|$, este espaço, que representamos por $C([a, b]; \mathbb{R})$ ou mais simplesmente por $C([a, b])$, é um espaço linear normado e completo, ou seja é um espaço de Banach.

O Teorema da Representação de Riez [17], afirma que se f é uma funcional linear e limitada, definida em $X = C([a, b])$, então existe uma medida de Borel regular e finita, μ , tal que para cada $x \in X$,

$$f(x) = \int_a^b x(t) d\mu(t)$$

e tal que a norma de f é a variação total de μ em $[a, b]$. Inversamente, qualquer função de variação limitada em $[a, b]$ define uma funcional linear limitada em $C([a, b])$.

- Uma sucessão de medidas μ_i converge *fracamente** para uma medida μ em $C^*([a, b])$, dual de $C([a, b])$, se para cada $\phi \in C([a, b])$,

$$\lim_{i \rightarrow \infty} \int_a^b \phi(t) d\mu_i = \int_a^b \phi(t) d\mu.$$

Do Teorema 2.1.1 resulta que, qualquer sucessão de medidas, limitada, em $C^*([a, b])$ contém uma subsucessão *fracamente** convergente.

Compacticidade *fraca* e *fraca** são propriedades menos restrictivas do que a compacticidade definida em termos da topologia usual num espaço. Um resultado bem conhecido da teoria clássica de optimização, estabelece que uma função g , definida num espaço normado X , tomando valores em $\mathbb{R}$ e semicontínua inferiormente, atinge o seu mínimo quando restricta a um subconjunto não vazio e compacto. Uma forma de alargar a aplicabilidade deste resultado é a utilização de uma topologia diferente, em relação à qual seja mais fácil estabelecer compacticidade. Uma escolha possível será a topologia *fraca*. Pela forma como está definida, esta topologia é a mais *fraca*, para a qual todas as funcionais lineares definidas no espaço são contínuas. Um conjunto $K \subset X$ *fracamente compacto* é caracterizado pelas propriedades: limitado em norma e *fracamente* fechado. Uma vez que, para conjuntos convexos, as propriedades *fracamente* fechado e fortemente fechado, são equivalentes, é possível utilizar argumentos análogos de demonstração em termos da topologia *fraca* e afirmar que a função dada, g , atinge o seu mínimo quando restricta a um conjunto convexo fechado, limitado e não vazio. Daqui resulta, por exemplo, que qualquer subconjunto convexo e fechado de X tem um elemento de norma mínima.

2.2 Convexidade

A convexidade desempenha um papel importante nos problemas de controlo óptimo. A presença desta propriedade em certos conjuntos e funções, não só simplifica a determinação das condições necessárias de optimalidade, mas também pode garantir a suficiência das mesmas condições.

O Teorema da Separação de Eidelheit que apresentamos, é um resultado clássico largamente utilizado neste contexto.

Um hiperplano $H \subset X$, é um conjunto do tipo $H = \{x \in X : f(x) = c\}$, onde f é uma função linear não nula, definida em X e com valores em $\mathbb{R}$ e c é uma constante real. Os conjuntos $\{x \in X : f(x) < c\}$, $\{x \in X : f(x) \leq c\}$, $\{x \in X : f(x) > c\}$ e $\{x \in X : f(x) \geq c\}$ são designados por *semiespaços* determinados por H .

Teorema 2.2.1 *Seja A um subconjunto limitado e convexo de $\mathbb{R}^n$. Então A tem a propriedade de ponto fixo, ou seja, para qualquer função contínua $f : A \rightarrow A$ existe um ponto $x \in A$ tal que $x = f(x)$.*

Teorema 2.2.2 (Teorema da Separação de Eidelheit) *Sejam K_1 e K_2 subconjuntos convexos de X tais que K_1 tem pontos interiores e K_2 não contém nenhum ponto interior de K_1 .*

Então, existe um hiperplano fechado H que separa os conjuntos K_1 e K_2 , isto é, existe $x^ \in X^*$ tal que*

$$\sup_{x \in K_1} \langle x^*, x \rangle \leq \inf_{x \in K_2} \langle x^*, x \rangle.$$

Por outras palavras, K_1 e K_2 estão em semiespaços opostos determinados por H .

Teorema 2.2.3 *Seja X um espaço topológico, C um subconjunto aberto e convexo de X e W um subconjunto convexo de $\mathbb{R} \times \mathbb{R}^m \times X$. Suponha-se que $0 \in W$ e $0 \in \bar{C}$. Então, uma das proposições 1) ou 2) seguintes é verdadeira:*

1) *Existe $l = (l_0, l_1, l_2) \in [0, \infty) \times \mathbb{R}^m \times X^*$, $l \neq 0$, tal que, $\forall w = (w_0, w_1, w_2) \in W$ e $\forall c \in \bar{C}$, tem-se*

$$l(w) = l_0 w_0 + l_1 \cdot w_1 + l_2(w_2) \geq 0, \quad l_2(c) \leq 0$$

2) *Existem pontos $\xi^i = (\xi_0^i, \xi_1^i, \xi_2^i) \in W$ e números $\beta^i > 0$, $i = 0, 1, \dots, m$ tais que $\sum_{i=0}^m \beta^i = 1$, $\xi_2^i \in C$, o conjunto $\{\xi_{0,1}^0, \dots, \xi_{0,1}^m\}$, onde $\xi_{0,1}^i = (\xi_0^i, \xi_1^i)$, é linearmente independente, $\xi_0^i < 0$ e $\sum_{i=0}^m \beta^i \xi_1^i = 0$.*

Os conjuntos W e C são dimensionalmente diferentes. O elemento 0 deve ser interpretado como o zero do espaço a que se refere.

A proposição 1) é imediatamente satisfeita, com $l_2 = 0$, se 0 é ponto fronteira da projecção do conjunto W em $\mathbb{R} \times \mathbb{R}^m$ e $(1, 0, 0)$ pertence a W .

A situação 2) acontece quando, entre outras condições, o ponto 0 é um ponto interior da projecção de W em $\mathbb{R} \times \mathbb{R}^m$.

O estudo das projecções de W , respectivamente em $\mathbb{R} \times \mathbb{R}^m$ e em X , argumentos de separação envolvendo estes conjuntos, os conjuntos C e W e ainda a consideração do conjunto $] -\infty, 0] \times \{0\} \times C$, permitem concluir o teorema na sua totalidade.

Lembramos em seguida a definição de função convexa e apresentamos uma caracterização alternativa destas funções.

Definição 2.2.1 *Seja A um subconjunto convexo de um espaço vectorial V e F uma função definida em A e tomando valores em $\mathbb{R}$.*

F diz-se convexa se, para todo u e v em A ,

$$F(\lambda u + (1 - \lambda)v) \leq \lambda F(u) + (1 - \lambda)F(v) \quad \forall \lambda \in [0, 1]$$

Para tais funções, o conjunto $\{(u, a) \in V \times \mathbb{R} : f(u) \leq a\}$, designado por *epigrafo* de F , é um conjunto convexo. Inversamente, se o epigrafo da função é um conjunto convexo, então a função será convexa.

Proposição 2.2.1 *Seja A um conjunto convexo e $F : A \rightarrow \mathbb{R}$, diferenciável em A . F é convexa sse*

$$F(v) \geq F(u) + F'(u) \cdot (v - u), \quad \forall v, u \in A.$$

2.3 Equações Diferenciais

Problemas de controlo óptimo envolvem de uma forma geral uma equação diferencial que traduz a dinâmica do sistema, representado em termos das variáveis de estado e do controlo. O teorema seguinte é um resultado de existência, unicidade e sensibilidade de tais equações a determinadas perturbações.

$\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ representa o espaço das funções lineares, definidas em $\mathbb{R}^m$ e tomando valores em $\mathbb{R}^n$.

$B(T, S; \mathcal{X})$, onde $T = [t_0, t_1] \subset \mathbb{R}$, S é um espaço topológico e $\mathcal{X}$ é um espaço de Banach, representa o espaço das funções $\phi : T \times S \rightarrow \mathcal{X}$ tais que:

- 1) Para todo $t \in T$, $\phi(t, \cdot) : S \rightarrow \mathcal{X}$ é uma função contínua;
- 2) Para todo $s \in S$, $\phi(\cdot, s) : T \rightarrow \mathcal{X}$ é uma função mensurável;
- 3) Existe uma função integrável $\psi_\phi : T \rightarrow \mathbb{R}$, tal que $\|\phi(t, \cdot)\|_{\sup} \leq \psi_\phi(t)$.

As propriedades dos elementos de $B(T, S; \mathcal{X})$ foram aparentemente estudadas, de forma sistemática, pela primeira vez por Caratheodory, no seu teorema de existência de solução para equações diferenciais, e por isso são designadas por funções de Caratheodory. Mais do que um espaço de funções, $B(T, S; \mathcal{X})$ representa um espaço vectorial de classes de equivalência de funções, na medida em que se duas funções ϕ_1 e ϕ_2 são tais que $\phi_1(t, \cdot) = \phi_2(t, \cdot)$ a menos de um conjunto de medida nula, serão identificadas como a mesma função.

Teorema 2.3.1 *Seja $f : T \times V \times A \rightarrow \mathbb{R}^n$, onde $T = [t_0, t_1] \subset \mathbb{R}$, V é um subconjunto aberto de $\mathbb{R}^n$ e A é um subconjunto convexo de $\mathbb{R}^k$. Suponha-se que, para todo $(t, v, a) \in T \times V \times A$, $f(t, \cdot, \cdot)$ é diferenciável em relação a (v, a) , e que*

$$f \in B(T, V \times A; \mathbb{R}^n), \quad f_{v,a} \in B(T, V \times A; \mathcal{L}(\mathbb{R}^{n+k}, \mathbb{R}^n)).$$

onde $f_{v,a}$ representa a derivada de f em ordem a (v, a) .

Para $\eta_0 \in V$ e $a_0 \in A$ seja $y_0 \in C(T, V)$ definido por

$$y_0(t) = \eta_0 + \int_{t_0}^t f(\tau, y_0(\tau), a_0) d\tau \quad (t \in T)$$

Então, existem vizinhanças $\tilde{V}$ de η_0 em V e $\tilde{A}$ de a_0 relativa a A , tais que a equação

$$y(t) = \eta + \int_{t_0}^t f(\tau, y(\tau), a) d\tau \quad (t \in T) \quad (2.1)$$

tem uma solução única $y(\cdot) = u(\eta, a)(\cdot)$ em $C(T, V)$, para todo $(\eta, a) \in \tilde{V} \times \tilde{A}$.

A função $u : \tilde{V} \times \tilde{A} \rightarrow C(T, \mathbb{R}^n)$ é contínua e tem uma derivada contínua, que representamos por u' , em $\tilde{V} \times \tilde{A}$.

Para $\eta \in \tilde{V}$, $\Delta\eta \in \mathbb{R}^n$, $a \in \tilde{A}$, $\Delta a \in \mathbb{R}^k$, e

$$\begin{aligned} p(\cdot) &= u'(\eta, a) \cdot (\Delta\eta, \Delta a)(\cdot) \\ &= u'((\eta, a); (\Delta\eta, \Delta a))(\cdot) \end{aligned}$$

onde a última expressão representa a derivada de u no ponto (η, a) e segundo a direção $(\Delta\eta, \Delta a)$, tem-se

$$p(t) = \int_{t_0}^t f_{v,a}(\tau, u(\eta, a)(\tau), a)(p(\tau), \Delta a) d\tau + \Delta\eta \quad (t \in T). \quad (2.2)$$

As equações 2.1 e 2.2 definem $y(t)$ e $p(t)$ como soluções, respectivamente, das equações diferenciais:

$$\dot{y}(t) = f(t, y(t), a) \quad \text{a.e. } t \in T; \quad y(t_0) = \eta$$

e

$$\begin{aligned} \dot{p}(t) &= f_v(t, u(\eta, a)(t), a)p(t) + f_a(t, u(\eta, a)(t), a)\Delta a \quad \text{a.e. } t \in T \\ p(t_0) &= \Delta\eta \end{aligned}$$

onde *a.e.* (*almost everywhere*), significa que a equação é satisfeita no conjunto T a menos de um conjunto de medida nula, em relação à medida de Lebesgue.

Em particular, quando $f(t, v, a) = A(t)v$, sendo $A(t)$ uma função matricial, a solução para a equação em y é dada por:

$$y(t) = \Phi(t, t_0)\eta$$

onde $\Phi(t, t_0)$ é a matriz de transição de estado associada à equação diferencial, mais precisamente, é a única solução da equação matricial diferencial:

$$\begin{aligned} \frac{\partial}{\partial t} \Phi(t, t_0) &= A(t)\Phi(t, t_0) \\ \Phi(t_0, t_0) &= I \quad (\text{matriz identidade}). \end{aligned}$$

Neste caso, a solução da equação diferencial depende linearmente do estado inicial e vem naturalmente:

$$p(t) = u'(\eta) \cdot \Delta\eta(t) = u(\Delta\eta)(t) = \Phi(t, t_0)\Delta\eta.$$

No caso mais geral, $p(t)$ é a solução de uma equação diferencial obtida a partir da linearização de f em relação a $u(\eta, a)(t)$ e a . A sua expressão é agora:

$$p(t) = \tilde{\Phi}(t, t_0)\Delta\eta + \int_{t_0}^t \tilde{\Phi}(t, s) f_a(s, u(\eta, a)(s), a)\Delta a \, ds$$

onde $\tilde{\Phi}$ é a matriz de transição de estado associada a $\dot{x} = f_v(t, u(\eta, a)(t), a)x$.

2.4 Multifunções e Trajectórias

Uma multifunção $\Gamma : S \subset \mathbb{R}^m \Rightarrow \mathbb{R}^n$ é uma aplicação de S nos subconjuntos de $\mathbb{R}^n$. Para cada x em S , $\Gamma(x)$ é um conjunto em $\mathbb{R}^n$. O seu gráfico, que representamos

por $\text{Graf } \Gamma$, é o conjunto

$$\text{Graf } \Gamma := \{(x, y) : x \in S, y \in \Gamma(x)\}.$$

Γ será fechada, compacta ou não vazia em S desde que, para cada x em S , o conjunto $\Gamma(x)$ tenha essa propriedade.

Uma multifunção $\Gamma : S \subset \mathbb{R}^m \Rightarrow \mathbb{R}^n$ diz-se ser mensurável se, para qualquer conjunto aberto C em $\mathbb{R}^n$, o conjunto

$$\{x \in S : \Gamma(x) \cap C \neq \emptyset\}$$

é um conjunto mensurável de Lebesgue. Se nesta definição substituirmos conjunto aberto por conjunto fechado ainda se obtém uma definição equivalente para multifunção mensurável.

O teorema seguinte, conhecido como Teorema da Selecção Mensurável, afirma a possibilidade de seleccionarmos um elemento de cada conjunto $\Gamma(x)$ de forma a obtermos uma função mensurável (em relação à medida de Lebesgue). Esta função é designada por *selecção mensurável* da multifunção Γ .

Teorema 2.4.1 *Seja Γ mensurável, fechada e não vazia em S . Então, existe uma função mensurável $\gamma : S \rightarrow \mathbb{R}^n$, tal que $\gamma(x)$ pertence a $\Gamma(x)$ para todo o x em S .*

Quando a multifunção não é fechada é possível ainda garantir a existência de uma tal selecção mensurável, atendendo às propriedades do seu gráfico. O resultado seguinte, Auman [3], está neste contexto e pode ser encarado como uma generalização do teorema anterior. De facto, quando Γ é fechada, a sua mensurabilidade implica que o gráfico é $\mathcal{L} \times \mathcal{B}$ mensurável. $\mathcal{L} \times \mathcal{B}$ representa a σ -álgebra dos subconjuntos de $\mathbb{R}^m \times \mathbb{R}^n$, gerada pelos conjuntos da forma $M \times N$, onde M é um subconjunto de Lebesgue de $\mathbb{R}^m$ e N é um subconjunto de Borel de $\mathbb{R}^n$.

Teorema 2.4.2 *Seja Γ uma multifunção definida num conjunto mensurável $S \subset \mathbb{R}^m$ e tomando valores nos subconjuntos não vazios de $\mathbb{R}^n$. Suponha-se que o gráfico de Γ é $\mathcal{L} \times \mathcal{B}$ mensurável. Então, Γ admite uma selecção mensurável.*

Rockafellar [23], prova alguns resultados interessantes sobre mensurabilidade e selecções mensuráveis.

Uma multifunção Γ diz-se ser de Lipschitz, de característica K em S se, para quaisquer x_1 e x_2 em S e para qualquer y_1 em $\Gamma(x_1)$, existe um elemento y_2 em $\Gamma(x_2)$ tal que

$$|y_1 - y_2| \leq K |x_1 - x_2|,$$

onde $|\cdot|$ representa a norma Euclidiana no espaço respectivo. No caso particular da multifunção ter como imagens conjuntos singulares, esta definição reduz-se à definição usual de funções de Lipschitz, equação 2.3.

Dado $\Omega \subset \mathbb{R} \times \mathbb{R}^n$, sejam S e Ω_t os conjuntos definidos por:

$$\begin{aligned} S &= \{t : (t, x) \in \Omega \text{ para algum } x \in \mathbb{R}^n\} \\ \Omega_t &= \{x : (t, x) \in \Omega\}. \end{aligned}$$

O conjunto Ω é chamado um *tubo* se S é um intervalo $[a, b]$ e se existir uma função contínua $w(t)$ e uma função positiva e contínua $\epsilon(t)$ em $[a, b]$ tal que $\Omega_t = w(t) + \epsilon(t)B$, para t em $[a, b]$. (B representa a bola aberta, centrada na origem e de raio unitário.) Nesse caso, Ω diz-se um *tubo* em $[a, b]$.

Se $x(\cdot)$ é uma função contínua em $[a, b]$ e $\epsilon > 0$, o ϵ -tubo em redor de $x(\cdot)$, que representamos na forma $\mathcal{T}(x; \epsilon)$, é o tubo definido por

$$\{(t, \bar{x}) : a \leq t \leq b, \bar{x} \in x(t) + \epsilon B\}.$$

Uma função contínua $x(\cdot)$, definida em $[a, b]$ e tal que $(t, x(t))$ pertence a Ω para todo t em $[a, b]$, diz-se estar em Ω . Note-se que se uma função contínua, $x(\cdot)$, está num tubo Ω , então Ω também contém $\mathcal{T}(x; \epsilon)$, para algum $\epsilon > 0$.

Definição 2.4.1 *Seja Ω um tubo em $[a, b]$. A multifunção $F : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, diz-se ser mensurável de Lipschitz em Ω , desde que:*

- (i) *Para cada x em $\mathbb{R}^n$, a multifunção $t \rightarrow F(t, x)$ é mensurável em $[a, b]$;*
- (ii) *Existe uma função integrável $k(t)$ definida em $[a, b]$ tal que, para cada t em $[a, b]$, a multifunção $x \rightarrow F(t, x)$ é não vazia e de Lipschitz de característica $k(t)$ em Ω_t .*

Associada à multifunção F é definida uma função distância, $\rho : \Omega \times \mathbb{R}^n \rightarrow [0, \infty]$:

$$\rho(t, x, v) = \inf\{|v - y| : y \in F(t, x)\}.$$

No caso particular de $F(t, x)$ ser fechada e não vazia, um elemento $v \in \mathbb{R}^n$ pertencerá a $F(t, x)$ se e só se $\rho(t, x, v) = 0$.

Se $F(t, x)$ é mensurável de Lipschitz, a função $\rho(t, x, v)$, como função de t é mensurável, para cada x e v fixos, e satisfaz a seguinte relação:

$$|\rho(t, x_1, v_1) - \rho(t, x_2, v_2)| \leq k(t)|x_1 - x_2| + |v_1 - v_2| \quad \forall (t, x_1), (t, x_2) \in \Omega, \forall v_1, v_2 \in \mathbb{R}^n.$$

Uma função $x : [a, b] \rightarrow \mathbb{R}^n$, absolutamente contínua é designada por *arco*. Um *arco* é portanto uma função que tem derivada $\dot{x}(t)$, em *quase todos* os pontos do intervalo onde está definida e que é o integral da sua derivada. Representamos por $AC^1([a, b]; \mathbb{R}^n)$, o espaço das funções com estas propriedades.

Ao arco x associamos a norma do supremo, definida por

$$\|x(\cdot)\| = \max\{|x(t)| : a \leq t \leq b\}.$$

Uma *trajectória* para a multifunção $F(t, x)$ é um arco x tal que, para quase todo o $t \in [a, b]$, $\dot{x}(t)$ pertence ao conjunto $F(t, x(t))$.

O teorema seguinte estabelece a existência de trajectórias para uma multifunção F a partir de arcos suficientemente próximos num dado sentido. A constante M e a função $\rho_F(x)$ são definidas por

$$M = \exp\left\{\int_a^b k(t) dt\right\} \quad \rho_F(x) = \int_a^b \rho(t, x(t), \dot{x}(t)) dt$$

onde $k(t)$ é a constante de Lipschitz da multifunção.

Teorema 2.4.3 *Seja F uma multifunção mensurável de Lipschitz no tubo Ω .*

Se x é um arco tal que $\mathcal{T}(x; \epsilon)$ está contido em Ω , e se $\rho_F(x) < \frac{\epsilon}{M}$, então existe uma trajectória y para F que está no tubo $\mathcal{T}(x; \epsilon)$, com $y(a) = x(a)$ e

$$\|x - y\| \leq \int_a^b |\dot{x}(t) - \dot{y}(t)| dt \leq M\rho_F(x).$$

2.5 Cálculo não Diferenciável

Funções não diferenciáveis não admitem, por definição, derivada. Para estas funções, resultados clássicos que envolvem tal conceito, como por exemplo resultados de optimização, deixam naturalmente de ser válidos. É possível, no entanto, estabelecer generalizações da teoria clássica que permitem o tratamento de tais funções.

Seja X um espaço de Banach e Y um subconjunto de X . Uma função $f : Y \rightarrow \mathbb{R}$ diz-se ser de Lipschitz, ou Lipschitz contínua, de característica K , se satisfaz:

$$|f(y) - f(x)| \leq K\|y - x\| \quad \forall y, x \in Y \quad (2.3)$$

Uma função f é de Lipschitz de característica K numa vizinhança de $x \in X$ se, para algum $\epsilon > 0$, f satisfaz 2.3 com $Y = x + \epsilon B$.

Seja f de Lipschitz numa vizinhança de $x \in X$ e v um vector qualquer de X . A derivada direccional generalizada de f na direcção de v , que representamos por $f^0(x; v)$, é definida da seguinte forma:

$$f^0(x; v) = \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{f(y + tv) - f(y)}{t},$$

onde y é um vector de X e t é um escalar positivo.

Note-se que esta definição não obriga a existência de qualquer limite e difere da derivada direccional tradicional na medida em que agora, y pode também variar.

A função f^0 satisfaz algumas propriedades básicas que facilitam a sua manipulação.

Proposição 2.5.1 *Seja f uma função de Lipschitz de característica K , numa vizinhança de x . Então,*

- (a) *A função $v \rightarrow f^0(x; v)$ é finita, homogénea positiva e subaditiva em X . Satisfaz ainda a relação: $|f^0(x; v)| \leq K\|v\|$.*
- (b) *$f^0(x; v)$ é semicontínua superiormente como função de (x, v) e, como função de v apenas, é Lipschitz contínua de característica K em X .*
- (c) *$f^0(x; -v) = (-f)^0(x; v)$.*

Qualquer função subaditiva e homogénea positiva em X majora sempre uma funcional linear em X , [17]. Podemos assim garantir a existência de pelo menos uma funcional linear $\zeta : X \rightarrow \mathbb{R}$, tal que, para todo o v em X , $f^0(x; v) \geq \zeta(v)$. Das propriedades de f^0 resulta que tal funcional é limitada e portanto pertence ao espaço dual de X . O gradiente generalizado de f em x , denotado por $\partial f(x)$, é o subconjunto não vazio de X^* definido por

$$\partial f(x) = \left\{ \zeta \in X^* : f^0(x; v) \geq \langle \zeta, v \rangle \text{ para todo } v \in X \right\},$$

onde $\langle \zeta, v \rangle$ representa $\zeta(v)$.

A caracterização do gradiente generalizado dada pelo teorema seguinte revela-se de particular importância, na medida em que facilita bastante o cálculo de $\partial f(x)$ em espaços de dimensão finita. Quando $X = \mathbb{R}^n$, o seu dual pode ser identificado com o próprio espaço e o conjunto $\partial f(x)$ encarado como um subconjunto de $\mathbb{R}^n$. O Teorema de Rademacher estabelece que, uma função definida num conjunto aberto de $\mathbb{R}^n$ e Lipschitz contínua nesse conjunto, é diferenciável em *quase todos os pontos* (abreviadamente, *a.e.*), no sentido da medida de Lebesgue. O conjunto dos pontos onde f não é diferenciável é denotado por Ω_f e é portanto um conjunto de medida nula.

Teorema 2.5.1 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, Lipschitz contínua numa vizinhança de x . Seja S um subconjunto qualquer de $\mathbb{R}^n$ com medida de Lebesgue nula. Então,*

$$\partial f(x) = \text{co} \{ \lim \nabla f(x_i) : x_i \rightarrow x, x_i \notin S, x_i \notin \Omega_f \} \quad (2.4)$$

A equação 2.4 deve ser interpretada da seguinte forma: considere-se uma sucessão qualquer x_i , cujos elementos são pontos que não estão em S e onde f é diferenciável, convergente para x , e tal que a sucessão dos respectivos gradientes, $\nabla f(x_i)$, converge. A cobertura convexa destes limites é $\partial f(x)$.

O gradiente generalizado satisfaz algumas propriedades básicas:

Proposição 2.5.2 *Seja f de Lipschitz de característica K , numa vizinhança de $x \in X$. Então,*

- (a) $\partial f(x)$ é um subconjunto de X^* , não vazio, convexo, fracamente* compacto e $\|\zeta\|_* \leq K$ para todo $\zeta \in \partial f(x)$.
- (b) Para qualquer $v \in X$ tem-se

$$f^0(x; v) = \max \{ \langle \zeta, v \rangle : \zeta \in \partial f(x) \}.$$

A generalização de alguns resultados clássicos de optimização para funções de Lipschitz, surge agora em termos do gradiente generalizado.

Teorema 2.5.2 *Seja $f : X \rightarrow \mathbb{R}$ uma função Lipschitz contínua. Se f atinge um mínimo ou um máximo local em x , então $0 \in \partial f(x)$.*

Dado um subconjunto C de X , não vazio, vectores tangentes e vectores normais a C podem ser definidos à custa da derivada direcciona! generalizada e do gradiente generalizado de uma função. A função que permite a definição de tais conceitos geométricos é naturalmente a função distância, $d_C(.) : X \rightarrow \mathbb{R}$, definida por:

$$d_C(x) = \inf \{ \|x - c\| : c \in C \}.$$

A função d_C não sendo diferenciável, é uma função de Lipschitz de característica 1. Tem sentido então definir para um ponto $x \in X$:

Um vector $v \in X$ diz-se *tangente* a C em x , se $d_C^0(x; v) = 0$.

O conjunto de todos os vectores tangentes a C em x , denotado por $T_C(x)$, é um cone convexo fechado que contém sempre o elemento zero de X .

Por polaridade com $T_C(x)$, define-se o cone *normal* a C em x :

$$N_C(x) = \{ \zeta \in X^* : \langle \zeta, v \rangle \leq 0 \text{ para todo } v \in T_C(x) \}$$

No caso particular de C ser convexo, a análise convexa define um vector normal a C em x , como sendo um elemento $\zeta \in X^*$ tal que $\langle \zeta, x - c \rangle \geq 0$ para todo $c \in C$. Para estes conjuntos as duas noções de normalidade coincidem.

Proposição 2.5.3 *Se C é convexo, $N_C(x)$ coincide com o cone dos vectores normais no sentido da análise convexa.*

A generalização do resultado clássico sobre optimização restricta para funções de Lipschitz, vem agora na seguinte forma:

Proposição 2.5.4 *Seja f de Lipschitz numa vizinhança de x e suponha-se que f , quando restricta a C , atinge um mínimo em x . Então, $0 \in \partial f(x) + N_C(x)$.*

Capítulo 3

Problema de Controlo Óptimo com Restrições

Neste capítulo vai ser apresentado o problema de controlo óptimo que servirá de base ao estudo efectuado neste trabalho. Depois de algumas breves considerações sobre problemas de controlo óptimo em geral, é formulado o problema e são indicadas as hipóteses básicas a que estão sujeitos os elementos nele envolvidos. O Princípio do Máximo, fornecendo um conjunto de condições necessárias a serem satisfeitas pela solução óptima, resulta trivial para o problema em questão. A análise desta situação constitui a parte restante do capítulo.

3.1 Problema de Controlo Óptimo

Num problema de controlo óptimo a palavra *controlo* refere-se a um elemento $u(\cdot)$, de uma dada classe de funções $\mathcal{U}$, definidas num intervalo $[a, b]$ e tomando valores num conjunto dado, $\Omega \subset \mathbb{R}^m$. A cada controlo é associado um arco $x(\cdot)$, designado por *estado*, através da equação diferencial e condição inicial:

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(a) = x_0, \quad (3.1)$$

onde f é uma dada função e o ponto x_0 pode ser escolhido de entre os elementos de um dado conjunto, C_0 .

O problema de controlo óptimo consiste em determinar $u(\cdot)$ e x_0 , e portanto $x(\cdot)$, de forma a minimizar uma funcional do tipo

$$g(x(b)) + \int_a^b F(t, x(t), u(t)) dt$$

sujeito à condição de $x(b)$ pertencer a algum conjunto dado, C_1 . Outros tipos de restrições, envolvendo o estado e o controlo podem ainda surgir.

Uma larga classe de sistemas físicos é modelizada por equações diferenciais nas quais surgem parâmetros que tomam valores apenas num determinado domínio. É neste contexto que a equação 3.1 é tratada. A introdução do conjunto Ω , que pode ser muito genérico, explica, em grande parte, a larga aplicação dos problemas de controlo óptimo.

A formulação original do Princípio do Máximo na teoria do controlo óptimo, devida a Pontryagin e seus colaboradores [20], aplicava-se a problemas com restrições muito básicas e foi demonstrado mediante hipóteses que exigiam para as funções envolvidas, propriedades muito regulares.

Desenvolvimentos posteriores levaram ao enfraquecimento das hipóteses e à inclusão de uma classe mais larga de restrições, de forma a estender a aplicabilidade do mesmo princípio [7], [8], [9], [12], [24].

Os problemas que vão ser estudados, assentam essencialmente na seguinte formulação:

$$\begin{aligned}
 & \text{Minimizar } g(x(b)) \\
 & \text{tal que} \\
 & \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \\
 (P) \quad & x(a) = x_0 \\
 & u(t) \in \Omega \quad \text{a.e. } t \in [a, b] \\
 & h(t, x(t)) \leq 0 \quad \text{para todo } t \in [a, b] \\
 & d(x(b)) \leq 0.
 \end{aligned} \tag{3.2}$$

A função custo, g , a ser otimizada é função apenas do estado final, $x(b)$. Trata-se de um problema de Mayer, como habitualmente é designado na literatura. A inclusão de um termo integral na função custo não é, de uma maneira geral, restrictiva, na medida em que, definindo uma nova variável de estado, podemos reformular o problema como um problema de Mayer. Esta técnica será utilizada no Capítulo 5, quando tratamos um problema do Cálculo Variacional.

O estado final está sujeito a restrições do tipo " $\leq$ ". No Capítulo 6 é considerado o caso em que esta restrição se reduz a uma igualdade.

Os dados, a especificar, são:

Um intervalo $[a, b] \in \mathbb{R}$, funções $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$, $h : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}$ e $d : \mathbb{R}^n \rightarrow \mathbb{R}^k$, um subconjunto $\Omega \subset \mathbb{R}^m$ e um ponto $x_0 \in \mathbb{R}^n$.

O vector de restrições no ponto final, 3.3, é interpretado no sentido usual, de restrições sobre cada componente.

Defina-se o conjunto de funções de controlo,

$$\mathcal{U} := \{ \text{funções mensuráveis } u : [a, b] \rightarrow \mathbb{R}^m : u(t) \in \Omega \text{ a.e. } \}.$$

Ao par de funções (x, u) constituído por uma função $x \in AC^1([a, b]; \mathbb{R}^n)$ e $u \in \mathcal{U}$, para as quais a equação diferencial 3.2 é satisfeita chamamos *processo*. A componente x de um processo designa-se por *trajectória*. Um processo para o qual todas as restrições do problema (P) são satisfeitas é chamado um *processo admissível*.

Um *minimizante*, $(\bar{x}, \bar{u})$, é um processo admissível para o qual $g(x(b))$ atinge o seu mínimo sobre todos os processos admissíveis (x, u) .

Suponha-se que as seguintes hipóteses são satisfeitas :

(H1): $x \rightarrow f(t, x, u)$ é continuamente diferenciável para (t, u) fixo, $(t, u) \rightarrow f(t, x, u)$ é $\mathcal{L} \times \mathcal{B}$ - mensurável para x fixo e f e f_x são limitadas em conjuntos limitados .

(H2): g e d são funções continuamente diferenciáveis .

(H3): Ω é um conjunto limitado de Borel .

(H4): $x \rightarrow h(t, x)$ é diferenciável para t fixo e h e h_x são contínuas .

(H5): $f(t, x, \Omega)$ é convexa para cada t, x , fixos.

A hipótese (H1) sobre f está intimamente relacionada com a equação diferencial a que está ligada. Questões sobre existência e sensibilidade da equação a perturbações podem ser resolvidas mediante esta hipótese.

A hipótese (H5) surge usualmente ligada à questão de existência de solução para o problema de controlo óptimo. Trata-se de uma hipótese essencial nesse contexto.

A regularidade imposta às restantes funções envolvidas tem a ver com a abordagem que vai ser utilizada no tratamento do problema. Sendo verdade que para construir qualquer teoria é necessário impor uma regularidade mínima aos elementos envolvidos, existe a preocupação de tentar que a mesma teoria se possa aplicar ao maior número possível de problemas. Assim, problemas do tipo (P) têm sido estudados mediante hipóteses cada vez mais fracas. Em [8], Clarke deduziu um Princípio do Máximo para o problema (P) em que g é localmente Lipschitz contínua, f como função de x , é Lipschitz contínua de característica $k(t, u)$, para alguma função mensurável e de valores reais $k(t, u)$, h é semicontínua superiormente como função de (t, x) e como função de x é Lipschitz contínua. Nestas condições, foram utilizadas na dedução do Princípio do Máximo, elementos de cálculo não diferenciável.

Num ambiente de cálculo não diferenciável, insere-se também o Capítulo 7. Aqui, é deduzido um Princípio do Máximo não trivial quando g e d são funções Lipschitz contínuas.

3.2 Princípio do Máximo

Como já foi referido anteriormente, o Princípio do Máximo estabelece um conjunto de condições necessárias, a serem satisfeitas pela solução óptima do problema de controlo óptimo.

Diz-se que $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo com multiplicadores $(p, \mu, \alpha, \lambda)$ se $p \in AC^1([a, b]; \mathbb{R}^n)$, μ é uma medida representando um elemento de $C^*([a, b]) (= C^*([a, b]; \mathbb{R}))$, dual do espaço das funções contínuas definidas em $[a, b]$ e com valores em $\mathbb{R}$, $\alpha \in \mathbb{R}^k$ e $\lambda \in \mathbb{R}$ tais que

$$\mu \geq 0, \alpha \geq 0, \lambda \geq 0,$$

$$-\dot{p}(t) = \left(p(t) + \int_{[a, t)} h_x(s, \bar{x}(s)) \mu(ds) \right) \cdot f_x(t, \bar{x}(t), \bar{u}(t)), \quad \text{a.e. } t \in [a, b],$$

para quase todo $t \in [a, b]$, a função

$$u \rightarrow \left(p(t) + \int_{[a, t)} h_x(s, \bar{x}(s)) \mu(ds) \right) \cdot f(t, \bar{x}(t), u) \quad (3.4)$$

atinge o seu máximo em Ω quando $u = \bar{u}(t)$,

$$\text{supp}\{\mu\} \subset \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}, \quad (3.5)$$

$$\alpha_i = 0 \text{ se } d^i(\bar{x}(b)) < 0$$

e

$$-\left(p(b) + \int_{[a, b]} h_x(s, \bar{x}(s)) \mu(ds) \right) = \alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b)). \quad (3.6)$$

A expressão $\text{supp}\{\mu\}$ representa o suporte da medida μ . Assim, $\mu(B) = 0$ para qualquer conjunto mensurável B que não interseccione o conjunto descrito em 3.5.

O Princípio do Máximo, [8], estabelece o seguinte:

"Mediante as hipóteses H1-H5, o minimizante satisfaz as condições do Princípio do Máximo com multiplicadores $(p, \mu, \alpha, \lambda)$, para os quais

$$|\mu|_{C^*} + |\alpha| + \lambda > 0". \quad (3.7)$$

Estas condições podem considerar-se como uma generalização a dimensão infinita dos multiplicadores de Lagrange para problemas de dimensão finita. A cada restrição é associado um multiplicador de dimensão e características adequadas.

Sendo (P) um problema de otimização sobre um espaço de funções, a maximização descrita por 3.4 é efectuada sobre um subconjunto de $\mathbb{R}^m$, constituindo por isso, um problema de dimensão finita. Esta condição é essencial, fornecendo em muitos casos um ponto de partida na pesquisa do óptimo para (P) . É usualmente designada na literatura por condição de maximização do Hamiltoniano. Definindo a função Hamiltoniana (ou Hamiltoniano de Pontryagin), $H_p : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, por

$$H_p(t, x, p, u) := \langle p, f(t, x, u) \rangle$$

é a maximização desta função que está em questão, em relação à variável u e quando (t, x, p) é um determinado ponto relacionado com o óptimo.

As hipóteses consideradas sobre f e h garantem a existência de p como solução da equação diferencial a que está sujeito.

Condições necessárias, envolvendo um conjunto de parâmetros a determinar, e equações do próprio problema, constituem um conjunto "razoável" de condições, capazes de fornecer candidatos a óptimo. Supondo que a equação 3.4 consegue definir $u(t)$ como função de $p(t)$ e μ , passamos a ter duas equações diferenciais com as respectivas condições fronteira, as restrições do problema associadas a h e d e as condições sobre μ e α . Daqui resultarão os candidatos a $\bar{x}$, p , α , μ e λ .

É evidente que as condições do Princípio do Máximo são sempre satisfeitas quando os multiplicadores são todos nulos. No entanto, nenhuma informação para a solução óptima é obtida neste caso. O Princípio do Máximo garante a existência de multiplicadores para além destes triviais.

Em determinadas situações surge um outro tipo de trivialidade que não é considerada pelo Princípio do Máximo, tal como foi enunciado.

Considere-se o caso em que o estado inicial pertence à fronteira do conjunto de restrições de estado, ou seja

$$h(a, x_0) = 0.$$

É simples verificar que quando isso acontece o minimizante $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo com os multiplicadores triviais não nulos:

$$p(t) \equiv -h_x(a, x_0), \quad \mu = \delta_{\{a\}}, \quad \alpha = 0 \quad \text{e} \quad \lambda = 0, \quad (3.8)$$

onde $\delta_{\{a\}}$ representa a medida unitária concentrada no ponto a .

Nenhum dos autores mencionados apresenta condições necessárias para problemas com restrições que forneçam informação sobre minimizantes, quando esta situação ocorre.

Problemas deste tipo, com restrições de estado e ponto inicial fixo, sobre a fronteira do conjunto de restrições, surgem por exemplo pela aplicação da metodologia de *horizonte rolante*, no cálculo de controladores realimentados estabilizadores, para um sistema dinâmico, $\dot{x} = f(t, x, u)$, sem serem violadas restrições de estado e controlo do tipo, $u(t) \in \Omega$ e $h(t, x(t)) \leq 0$. De acordo com tal metodologia [18], a lei de controlo realimentado no instante $t = \tau$ e estado $x = \xi$, nomeadamente

$$u(t) = \Upsilon(\tau, \xi),$$

é obtida através da resolução de um problema de controlo óptimo da forma

$$\begin{aligned} \text{Minimizar} \quad & \int_{\tau}^{\tau+T} (|x(t)|^2 + k|u(t)|^2) dt \\ & \text{tal que} \\ & \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [\tau, \tau + T] \\ & u(t) \in \Omega \quad \text{a.e. } t \in [\tau, \tau + T] \\ & h(t, x(t)) \leq 0 \quad \forall t \in [\tau, \tau + T] \\ & x(\tau) = \xi \end{aligned}$$

Obtém-se um minimizante $(y(t), v(t))$ $\tau \leq t \leq \tau + T$. O *horizonte* T e a constante k são parâmetros fixos. O controlo realimentado é então escolhido como

$$\Upsilon(\tau, \xi) = v(t) |_{t=\tau}.$$

O problema de controlo óptimo envolvido pode ser reformulado como um problema do tipo (P). Se, como seria de esperar, a trajectória de estado gerada pela abordagem *horizonte rolante* atinge a fronteira da região dos estados admissíveis, surge então a necessidade de resolver um problema de controlo óptimo com ponto inicial fixo (τ, ξ) na fronteira da região de estados admissíveis.

Problemas deste tipo são ainda encontrados na resolução de um problema com restrições de estado, quando se constroi uma solução da equação de Hamilton-Jacobi via "campos de extremos". Neste caso, devem ser encontrados processos admissíveis que satisfazem o Princípio do Máximo de uma forma "normal" ou não trivial, para problemas com a restrição sobre o ponto inicial $x(\tau) = \xi$, e quando (τ, ξ) percorre a região dos estados admissíveis. Encontrar tais processos para pontos na fronteira desta região faz surgir o tipo de situações que são abordadas neste trabalho.

Capítulo 4

Princípio do Máximo não Trivial

No capítulo anterior foi formulado um problema de controlo óptimo, (P) , e apresentada uma situação em que o Princípio do Máximo não fornece qualquer informação útil na determinação da solução do problema. De facto, quando o estado inicial está na fronteira da região de admissibilidade, existe um conjunto de multiplicadores satisfazendo trivialmente as condições necessárias de optimalidade definidas pelo Princípio do Máximo. Neste capítulo é feito o estudo do problema (P) , quando essa situação está presente.

Começamos por apresentar uma série de resultados que asseguram a existência de outros multiplicadores para além dos triviais, desde que uma certa qualificação das restrições seja satisfeita. O Princípio do Máximo a que se faz alusão nestes resultados, toma a forma já enunciada na secção 3.2. Também as hipóteses H1-H5 que vão ser consideradas foram já expostas no capítulo anterior.

Na secção 4.2 é feita a demonstração desses resultados.

4.1 Condições Necessárias não triviais

Seja $(\bar{x}, \bar{u})$ um minimizante para o problema (P) e suponhamos que o estado inicial está sobre a fronteira da região dos estados admissíveis, ou seja

$$h(a, x_0) = 0. \quad (4.1)$$

Como já foi observado, quando 4.1 acontece $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo com os seguintes multiplicadores triviais (ou versões escalares dos mesmos):

$$p(t) \equiv -h_x(a, x_0), \quad \mu = \delta_{\{a\}}, \quad \alpha = 0 \text{ e } \lambda = 0.$$

Vejamos então como podemos garantir a existência de outros multiplicadores que possam dar informação sobre a solução do problema.

Dado um conjunto mensurável $A \subset \mathbb{R}$ e f uma função definida em A e tomando valores em $\mathbb{R}$, mensurável, o supremo essencial e ínfimo essencial de f em A são definidos da seguinte forma:

$$\operatorname{ess\,sup}_A f = \inf \left\{ \sup_{t \in A} g(t) : g(t) = f(t) \text{ a.e. } t \in A, g \text{ mensurável} \right\}$$

$$\operatorname{ess\,inf}_A f = - \operatorname{ess\,sup}_A (-f)$$

Teorema 4.1.1 *Seja $(\bar{x}, \bar{u})$ um minimizante para (P) e considere-se que as hipóteses H1-H5 são satisfeitas.*

Suponhamos que,

$(CQ)_1$: $h(a, x_0) = 0$ e existe $u' \in \mathcal{U}$ tal que

$$\lim_{\epsilon \rightarrow 0^+} \operatorname{ess\,sup}_{a \leq s < a + \epsilon} h_x(a, x_0) \cdot [f(s, x_0, u'(s)) - f(s, x_0, \bar{u}(s))] < 0.$$

Então, $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo, com multiplicadores $(p, \mu, \alpha, \lambda)$ para os quais

$$\int_{(a,b]} \mu(ds) + |\alpha| + \lambda \neq 0. \quad (4.2)$$

Teorema 4.1.2 *As conclusões do Teorema 4.1.1 permanecem válidas quando a qualificação de restrições $(CQ)_1$ é substituída por*

$$(CQ)_2 : \exists \bar{t} \in (a, b] \text{ tal que } h(t, \bar{x}(t)) < 0 \quad \forall t \in (a, \bar{t}].$$

Note-se que, a novidade na equação 4.2 está no facto de a medida μ ser integrada sobre o intervalo *semiaberto* $(a, b]$, excluindo desta forma os multiplicadores para os quais $\alpha = 0$, $\lambda = 0$ e μ está concentrada no ponto a .

Como interpretar $(CQ)_1$?

Se $t \rightarrow f(t, x_0, u)$ for uniformemente contínua em $t = a$, para $u \in \Omega$, esta condição é equivalente à condição $(CQ)_1'$ seguinte:

$(CQ)_1'$: Existe $v \in \Omega$ tal que

$$h_x(a, x_0) \cdot f(a, x_0, v) < \lim_{\epsilon \rightarrow 0^+} \operatorname{ess\,inf}_{a \leq s < a + \epsilon} h_x(a, x_0) \cdot f(s, x_0, \bar{u}(s)). \quad (4.3)$$

Consideremos o caso em que h não depende de t , isto é, $h(t, x) = \tilde{h}(x)$ para alguma função $\tilde{h}$, e $\tilde{h}(\bar{x}(t)) = 0$ numa vizinhança relativa de $t = a$ em $[a, b]$. Isto significa que $\bar{x}(t)$ permanece na fronteira da região admissível, próximo de $t = a$. Então, nessa vizinhança, $\frac{d}{dt} \tilde{h}(\bar{x}(t)) = 0$ a.e. e portanto

$$\lim_{\epsilon \rightarrow 0^+} \operatorname{ess\,inf}_{a \leq s < a + \epsilon} \frac{d}{ds} \tilde{h}(\bar{x}(s)) = 0,$$

vindo o segundo membro de 4.3 nulo.

$(CQ)'_1$ toma neste caso a forma: existe $v \in \Omega$ tal que

$$\tilde{h}_x(x_0) \cdot f(a, x_0, v) < 0.$$

A qualificação de restrições $(CQ)'_1$ estabelece assim a existência de um valor para o controlo que "puxa" o estado da fronteira, num ângulo não nulo, mantendo-o na região admissível.

Em contrapartida, $(CQ)_2$, quando $h(a, x_0) = 0$, considera o caso em que a trajectória minimizante $\bar{x}$ "abandona" a fronteira imediatamente.

Os teoremas seguintes definem condições mediante as quais é garantida *normalidade*, a propriedade de os multiplicadores serem tais que $\lambda \neq 0$. Quando $\lambda = 0$, é de esperar que o Princípio do Máximo não identifique minimizantes, uma vez que a função custo não intervém nas condições necessárias. Os teoremas 4.1.1 e 4.1.2 não cobrem este caso, permitindo que multiplicadores como $\lambda = 0$, $\alpha \neq 0$ ou, $\lambda = 0$ e $\int_{[a,b]} \mu(ds) \neq 0$ estejam associados à solução óptima.

As novas condições são descritas em termos das equações dinâmicas, linearizadas em relação às variáveis de estado. Para $v \in \Omega$, defina-se,

$$\Delta f_v(t) = f(t, \bar{x}(t), v) - f(t, \bar{x}(t), \bar{u}(t)). \quad (4.4)$$

Dado um controlo $u \in \mathcal{U}$, seja y_u a solução, única, absolutamente contínua, do sistema

$$\begin{aligned} \dot{y}_u(t) &= f_x(t, \bar{x}(t), \bar{u}(t)) \cdot y_u(t) + \Delta f_{u(t)}(t) \\ y_u(a) &= 0. \end{aligned} \quad (4.5)$$

Teorema 4.1.3 *Seja $(\bar{x}, \bar{u})$ um minimizante para (P) e considere-se que as hipóteses H1-H5 são satisfeitas.*

Suponhamos que

$(CQ)_3$: $h(a, x_0) = 0$ e existe $u' \in \mathcal{U}$ tal que

$$\lim_{\epsilon \rightarrow 0^+} \text{ess sup}_{a \leq s < a + \epsilon} h_x(a, x_0) \cdot [f(s, x_0, u'(s)) - f(s, x_0, \bar{u}(s))] < 0.$$

$$h_x(t, \bar{x}(t)) \cdot y_{u'}(t) < 0 \quad \forall t \in (a, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}$$

$$d_x^i(\bar{x}(b)) \cdot y_{u'}(b) < 0 \quad \text{se } d^i(\bar{x}(b)) = 0. \quad (4.6)$$

Então, $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo com multiplicadores $(p, \mu, \alpha, \lambda)$ para os quais $\lambda \neq 0$.

Teorema 4.1.4 Os resultados do Teorema 4.1.3 permanecem válidos quando substituimos a qualificação de restrições $(CQ)_3$ por:

$(CQ)_4$: Existe $\bar{t} \in (a, b]$ e $u' \in \mathcal{U}$ tal que

$$h(t, \bar{x}(t)) < 0 \quad \forall t \in (a, \bar{t}],$$

$$h_x(t, \bar{x}(t)) \cdot y_{u'}(t) < 0 \quad \forall t \in (a, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}$$

e

$$d_x^i(\bar{x}(b)) \cdot y_{u'}(b) < 0 \quad \text{se } d^i(\bar{x}(b)) = 0. \quad (4.7)$$

4.2 Demonstração dos Teoremas 4.1.1 - 4.1.4

4.2.1 Resultados Preliminares

Começamos por deduzir alguns resultados necessários para a demonstração destes teoremas.

Para cada $u \in \mathcal{U}$, y_u representa, tal como foi anteriormente definido, a solução da equação dinâmica linearizada. A proposição seguinte exprime a forma como funções do tipo $\bar{x} + \epsilon y_u$ são próximas de trajectórias para o problema (P) .

Proposição 4.2.1 *Seja $u \in \mathcal{U}$, qualquer. Então, existe $\epsilon' > 0$ e uma função crescente $w : [0, \epsilon'] \rightarrow [0, \infty)$, satisfazendo $\lim_{s \downarrow 0} w(s) = 0$, com a seguinte propriedade: A cada $\epsilon \in (0, \epsilon')$ é possível fazer corresponder um processo (x_ϵ, u_ϵ) para (P) , cuja trajectória satisfaz:*

$$\max_{s \in [a, b]} \|x_\epsilon(s) - [\bar{x}(s) + \epsilon y_u(s)]\| \leq (t - a)\epsilon w(\epsilon) \quad \forall t \in [a, b].$$

Demonstração.

Primeiro passo

Começamos por verificar a existência de x_ϵ nas condições descritas pela proposição.

Seja então $u \in \mathcal{U}$. Para $\epsilon' \in (0, 1]$ e $\epsilon \in [0, \epsilon']$, considere-se a equação

$$\begin{aligned}\dot{x}(t) &= f(t, x(t), \bar{u}(t)) + \epsilon[f(t, x(t), u(t)) - f(t, x(t), \bar{u}(t))] \\ x(a) &= x_0.\end{aligned}\tag{4.8}$$

A única solução desta equação é ainda solução da inclusão diferencial

$$\dot{x}(t) \in F(t, x(t))\tag{4.9}$$

onde $F(t, x)$ é a multifunção de valores simples,

$$F(t, x) = \{f(t, x, \bar{u}(t)) + \epsilon[f(t, x, u(t)) - f(t, x, \bar{u}(t))]\}\tag{4.10}$$

A solução da inclusão diferencial 4.9, que satisfaz $x(a) = x_0$, única neste caso, está nas condições especificadas pela proposição.

Para verificar esta afirmação, vejamos que estão criadas as condições de aplicabilidade do Teorema 2.4.3, sobre trajectórias aproximantes.

Fixe-se t e considere-se o intervalo $[a, t]$.

Designa-se por $\mathcal{T}(\bar{x}; \alpha)$, o tubo centrado em $\bar{x}$ e raio $\alpha > 0$, onde α é qualquer, escolhido de forma a que as hipóteses (H1)-(H5) sejam válidas nesse tubo.

Defina-se $\|x\|$ como o máximo de $|x(s)|$ em $[a, t]$ ou seja, considere-se a norma " ∞ " no espaço das funções contínuas definidas no intervalo $[a, t]$.

Da definição de y_u resulta,

$$\|\bar{x} + \epsilon y_u - \bar{x}\| = \epsilon \|y_u\| \leq 2\epsilon [\exp\{K(t-a)\} - 1]$$

onde

$$K = \max \left\{ \max_{[a, b] \times \mathcal{T}(\bar{x}; \alpha) \times \bar{\Omega}} f(t, x, u), \max_{[a, b] \times \mathcal{T}(\bar{x}; \alpha) \times \bar{\Omega}} f_x(t, x, u) \right\}.$$

Para $\nu > 0$,

$$\begin{aligned}z \in \mathcal{T}(\bar{x} + \epsilon y_u; \nu) &\Leftrightarrow \|z - \bar{x} - \epsilon y_u\| < \nu \\ &\Rightarrow \|z - \bar{x}\| < \nu + \epsilon \|y_u\| < \nu + 2\epsilon [\exp\{K(t-a)\} - 1].\end{aligned}$$

Assim, para

$$\epsilon < \epsilon' < \frac{\alpha}{4[\exp\{K(b-a)\} - 1]}, \quad \epsilon' < 1 \text{ e } \nu = \frac{\alpha}{2}$$

vem,

$$\begin{aligned}\bar{x} + \epsilon y_u &\in T(\bar{x}; \alpha) \\ T(\bar{x} + \epsilon y_u; \nu) &\subset T(\bar{x}; \alpha)\end{aligned}$$

A multifunção F , definida por 4.10, é mensurável em relação à variável s no intervalo $[a, t]$ e, como função de x , é Lipschitz contínua de característica K , no conjunto $T_s(\bar{x}; \alpha)$, $a \leq s \leq t$.

Para o intervalo $[a, t]$, considerem-se as funções, já definidas no Cap. 2,

$$\rho(s, x, v) = \inf\{|v - y| : y \in F(s, x)\} \text{ e } \rho_F(x) = \int_a^t \rho(s, x(s), \dot{x}(s)) ds.$$

Note-se que, como $F(s, x)$ é de valores simples, o ínfimo na expressão anterior é obtido calculando o único elemento do conjunto.

Diminuindo o valor de ϵ' , se necessário, de forma a que

$$\epsilon' < \frac{\alpha}{16K(b-a) \exp\{K(b-a)\}[\exp\{K(b-a)\} - 1]}$$

tem-se,

$$\begin{aligned}\rho_F(\bar{x} + \epsilon y_u) &= \int_a^t \|f(s, \bar{x}(s), \bar{u}(s)) + \epsilon f_x(s, \bar{x}(s), \bar{u}(s))y_u(s) + \epsilon[f(s, \bar{x}(s), u(s)) - \\ &\quad f(s, \bar{x}(s), \bar{u}(s))] - f(s, \bar{x}(s) + \epsilon y_u(s), \bar{u}(s)) - \\ &\quad \epsilon[f(s, \bar{x}(s) + \epsilon y_u(s), u(s)) - f(s, \bar{x}(s) + \epsilon y_u(s), \bar{u}(s))]\| ds \\ &\leq \int_a^t \|f(s, \bar{x}(s) + \epsilon y_u(s), \bar{u}(s)) - f(s, \bar{x}(s), \bar{u}(s)) - \\ &\quad \epsilon f_x(s, \bar{x}(s), \bar{u}(s))y_u(s)\| + \epsilon \max \|f_x\| \|\epsilon y_u(s)\| \\ &\quad + \epsilon \max \|f_x\| \|\epsilon y_u(s)\| ds \\ &\leq 4K\epsilon^2 [\exp\{K(b-a)\} - 1](t-a) + \\ &\quad \int_a^t \|\int_0^1 f_x(s, \bar{x}(s) + l\epsilon y_u(s), \bar{u}(s))\epsilon y_u(s) dl - \\ &\quad \epsilon f_x(s, \bar{x}(s), \bar{u}(s))y_u(s)\| ds \\ &\leq 4K\epsilon^2 [\exp\{K(b-a)\} - 1](t-a) + \epsilon \int_a^t g(s, \epsilon) ds \\ &< \frac{\nu}{\exp\{\int_a^t K dt\}}\end{aligned}$$

onde

$$g(s, \epsilon) = \|y_u(s)\| \int_0^1 \|f_x(s, \bar{x}(s) + l\epsilon y_u(s), \bar{u}(s)) - f_x(s, \bar{x}(s), \bar{u}(s))\| dl.$$

Finalmente, aplicando o Teorema 2.4.3, conclui-se que, para $\epsilon < \epsilon'$ e ϵ' menor que

$$\min \left\{ 1, \frac{\alpha}{4[(\exp \{K(b-a)\} - 1)]}, \frac{\alpha}{16K(b-a) \exp \{K(b-a)\} [\exp \{K(b-a)\} - 1]} \right\}$$

existe uma trajectória x_ϵ para F , contida no tubo $\mathcal{T}(\bar{x} + \epsilon y_u; \nu = \frac{\alpha}{2})$, com

$x_\epsilon(a) = \bar{x}(a)$ e

$$\max_{a \leq s \leq t} \|x_\epsilon(s) - (\bar{x}(s) + \epsilon y_u(s))\| \leq (t-a)\epsilon \bar{w}(\epsilon)$$

onde,

$$\bar{w}(\epsilon) = 4K\epsilon \exp\{K(b-a)\} [\exp\{K(b-a)\} - 1] + \\ \exp\{K(b-a)\} \sup_{t \in (a,b]} \frac{1}{t-a} \int_a^t g(s, \epsilon) ds.$$

A função $\bar{w}$ assim definida é uma função limitada e satisfaz: $\lim_{\epsilon \downarrow 0} \bar{w}(\epsilon) = 0$.

De facto,

$$\begin{aligned} 0 &\leq \sup_{t \in (a,b]} \frac{1}{t-a} \int_a^t g(s, \epsilon) ds \\ &= \sup_{t \in (a,b]} \frac{1}{t-a} \int_a^t \|y_u(s)\| \int_0^1 \|f_x(s, \bar{x}(s) + l\epsilon y_u(s), \bar{u}(s)) - \\ &\quad f_x(s, \bar{x}(s), \bar{u}(s))\| dl ds \\ &\leq \sup_{t \in (a,b]} \frac{2}{t-a} [\exp\{K(t-a)\} - 1] \int_a^b \int_0^1 \|f_x(s, \bar{x}(s) + l\epsilon y_u(s), \bar{u}(s)) - \\ &\quad f_x(s, \bar{x}(s), \bar{u}(s))\| dl ds \end{aligned}$$

A aplicação do Teorema da Convergência Dominante de Lebesgue, Rudin [25], permite concluir,

$$\lim_{\epsilon \downarrow 0} \int_a^b \int_0^1 \|f_x(s, \bar{x}(s) + l\epsilon y_u(s), \bar{u}(s)) - f_x(s, \bar{x}(s), \bar{u}(s))\| dl ds = 0.$$

Por outro lado, a função $t \rightarrow \frac{2}{t-a} [\exp\{K(t-a)\} - 1]$ é limitada em $(a, b]$. É fácil concluir agora que $\bar{w}$ tem as propriedades proclamadas anteriormente.

Se definirmos $w(\epsilon) = \sup_{\lambda \in [0, \epsilon]} \bar{w}(\lambda)$, obtém-se uma função nas condições especificadas pela proposição.

Note-se que o resultado foi obtido actuando sobre o intervalo $[a, t]$. Das propriedades das funções envolvidas resulta facilmente que ao tomarmos outro intervalo $[a, t']$, com $t' > t$, a trajectória aproximante obtida seria igual a x_ϵ no intervalo comum $[a, t]$. A função x_ϵ está assim bem definida em todo o intervalo $[a, b]$.

Segundo passo

Provemos que x_ϵ é uma trajectória associada a um processo para o problema (P).

Para isso basta verificar que existe $u_\epsilon \in \mathcal{U}$ tal que

$$\dot{x}_\epsilon(t) = f(t, x_\epsilon(t), u_\epsilon(t)) \quad a.e. \ t \in [a, b].$$

Das propriedades de x_ϵ e do facto de $f(t, x, \Omega)$ ser convexo, resulta, para $t \in [a, b]$ a.e.,

$$\dot{x}_\epsilon(t) = (1 - \epsilon)f(t, x_\epsilon(t), \bar{u}(t)) + \epsilon f(t, x_\epsilon(t), u(t)) \in f(t, x_\epsilon(t), \Omega).$$

Considere-se a função seguinte, definida em $[a, b] \times \mathbb{R}^m$:

$$g(t, v) = f(t, x_\epsilon(t), v) - [(1 - \epsilon)f(t, x_\epsilon(t), \bar{u}(t)) + \epsilon f(t, x_\epsilon(t), u(t))]$$

e as multifunções

$$\Gamma : [a, b] \Rightarrow \mathbb{R}^m \quad \Gamma(t) = \{v \in \Omega : g(t, v) = 0\}$$

$$\Gamma' : [a, b] \Rightarrow \mathbb{R}^m \quad \Gamma'(t) = \{v \in \mathbb{R}^m : g(t, v) = 0\}.$$

Γ e Γ' são aplicações do subconjunto mensurável $[a, b]$ nos subconjuntos não vazios de $\mathbb{R}^m$. Os seus gráficos estão relacionados da seguinte forma:

$$Graf \ \Gamma = \{[a, b] \times \Omega\} \cap Graph \ \Gamma'. \quad (4.11)$$

A verificação da existência de uma selecção mensurável para o gráfico de Γ conduz ao resultado desejado. A continuidade de $x_\epsilon(\cdot)$ e o facto de f ser $\mathcal{L} \times \mathcal{B}$ -mensurável, como função de (t, u) , implicam que as funções

$$(t, v) \rightarrow f(t, x_\epsilon(t), v) \quad (t, v) \rightarrow (1 - \epsilon)f(t, x_\epsilon(t), \bar{u}(t)) + \epsilon f(t, x_\epsilon(t), u(t))$$

sejam $\mathcal{L} \times \mathcal{B}$ -mensuráveis ([7]).

Daqui resulta a $\mathcal{L} \times \mathcal{B}$ -mensurabilidade de $g(t, v)$. Assim, por definição de função mensurável, imagens recíprocas de conjuntos fechados são conjuntos mensuráveis. Destes factos se conclui,

$$Graph \ \Gamma' = g^{-1}(0) \in \mathcal{L} \times \mathcal{B}.$$

Mas $[a, b] \times \Omega$ é um conjunto $\mathcal{L} \times \mathcal{B}$ -mensurável. Da equação 4.11 deduz-se então:

$$Graph \ \Gamma \text{ é } \mathcal{L} \times \mathcal{B}\text{-mensurável.}$$

Resultados sobre mensurabilidade, Teor. 2.4.2, permitem concluir que Γ admite uma selecção mensurável, ou seja, existe $u_\epsilon(\cdot)$ tal que $u_\epsilon(t) \in \Omega$ e $g(t, u_\epsilon(t)) = 0$.

O par (x_ϵ, u_ϵ) constitui um processo.

A demonstração está concluída. ■

São estabelecidas em seguida algumas propriedades que vão conduzir à forma final dos multiplicadores, tal como se apresentam nas condições necessárias de optimalidade para o problema (P).

Recorde-se a definição 4.4 de $\Delta f_v(t)$. De forma a não tornar pesada a notação, representamos algumas vezes $\Delta f_{u(t)}(t)$, quando $u \in \mathcal{U}$, simplesmente por $\Delta f_u(t)$. Φ designa a matriz de transição de estado associada à equação diferencial linear, $\dot{z}(t) = f_x(t, \bar{x}(t), \bar{u}(t)) \cdot z(t)$.

Proposição 4.2.2 *Seja $\mu \in C^*([a, b])$, $\alpha \in \mathbb{R}^k$, $\lambda \in \mathbb{R}$ e $u \in \mathcal{U}$. Então*

$$\begin{aligned} \int_{[a,b]} h_x(t, \bar{x}(t)) \cdot y_u(t) \mu(dt) + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \cdot y_u(b) \\ = \int_{[a,b]} q(t) \cdot \Delta f_u(t) dt. \end{aligned} \quad (4.12)$$

Nesta equação, $q(t)$ é a função definida por

$$q(t) := \int_{[t,b]} h_x(s, \bar{x}(s)) \Phi(s, t) \mu(ds) + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \Phi(b, t)$$

em $[a, b]$ e pela condição fronteira:

$$q(b) = [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))].$$

$q(t)$ pode ser expresso como,

$$q(t) = \begin{cases} -[p(t) + \int_{[a,t]} h_x(s, \bar{x}(s)) \mu(ds)] & t \in [a, b) \\ -[p(b) + \int_{[a,b]} h_x(s, \bar{x}(s)) \mu(ds)] & t = b, \end{cases}$$

onde p é a solução da equação diferencial

$$-\dot{p}(t) = \left(p(t) + \int_{[a,t]} h_x(s, \bar{x}(s)) \mu(ds) \right) \cdot f_x(t, \bar{x}(t), \bar{u}(t))$$

com condição fronteira,

$$-\left(p(b) + \int_{[a,b]} h_x(s, \bar{x}(s)) \mu(ds) \right) = \alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b)).$$

Demonstração.

Começemos por considerar μ absolutamente contínua em relação à medida de Lebesgue, isto é $d\mu = m dt$ para alguma função $m \in L^1([a, b])$, onde $L^1([a, b])$ representa o espaço das funções definidas em $[a, b]$ e tomando valores em $\mathbb{R}$, mensuráveis e integravelmente limitadas. Neste caso a equação 4.12 resulta facilmente da definição de $y_u(t)$ e das propriedades do integral, como vamos verificar.

Represente-se por π o primeiro membro da equação 4.12, ou seja,

$$\pi := \int_{[a,b]} h_x(t, \bar{x}(t)) \cdot y_u(t) m(t) dt + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \cdot y_u(b). \quad (4.13)$$

Aplicando o método de variação das constantes para a resolução de equações diferenciais, obtém-se para y_u a expressão

$$y_u(t) = \int_{[a,t]} \Phi(t, s) \Delta f_u(s) ds.$$

Substituindo na expressão de π e efectuando uma mudança de ordem nas variáveis de integração, vem

$$\pi = \int_{[a,b]} q(t) \Delta f_{u(t)}(t) dt \quad (4.14)$$

onde,

$$q(t) = \int_{[t,b]} h_x(s, \bar{x}(s)) \Phi(s, t) m(s) ds + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \Phi(b, t). \quad (4.15)$$

Definindo $p(t)$ como sendo a função absolutamente contínua,

$$p(t) := - \left(q(t) + \int_{[a,t]} h_x(s, \bar{x}(s)) m(s) ds \right) \quad (4.16)$$

a sua derivada, tendo em conta que $\frac{\partial}{\partial t} \Phi(s, t) = -\Phi(s, t) f_x(t, \bar{x}(t), \bar{u}(t))$ a.e., toma a forma

$$\begin{aligned} \dot{p}(t) &= h_x m(t) - \left[\int_{[t,b]} h_x \Phi m ds + (\alpha^T d_x + \lambda g_x) \Phi(b, t) \right] \cdot f_x(t, \bar{x}(t), \bar{u}(t)) \\ &\quad - h_x m(t) \\ &= - \left[p(t) + \int_{[a,t]} h_x m ds \right] \cdot f_x(t, \bar{x}(t), \bar{u}(t)). \end{aligned} \quad (4.17)$$

Ainda da definição de $p(t)$ resulta,

$$- \left(p(b) + \int_{[a,b]} h_x(s, \bar{x}(s)) m(s) ds \right) = \alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b)). \quad (4.18)$$

As equações 4.15-4.18 e as expressões equivalentes para π , 4.13 e 4.14, provam a proposição quando μ é absolutamente contínua. Para provar o caso geral, basta considerar uma sucessão de funções $\{m_i\}$, em $L^1([a, b])$, tal que $m_i dt$ converge fracamente* para $d\mu$ em $C^*([a, b])$. Para cada i tome-se q_i e p_i definidas respectivamente por 4.15 e 4.16, mas com m substituído por m_i . As propriedades destas funções garantem a existência de uma subsucessão de $\{p_i\}$ que converge uniformemente para alguma função absolutamente contínua p . De forma análoga é possível concluir a existência de uma medida dq , podendo ser representada como um elemento de $C^*([a, b])$, que é o limite fraco* da sucessão de medidas $q_i dt$. Resultados, que se podem considerar já clássicos, sobre convergência [26], permitem aplicar limites às expressões correspondentes a 4.13-4.18 concluindo desta forma a proposição.

■

A proposição seguinte analisa certas implicações das qualificações de restrições apresentadas na secção anterior.

Proposição 4.2.3 (a) *Suponhamos que a qualificação de restrições do Teorema 4.1.1 é satisfeita. Então existe $\tilde{u} \in \mathcal{U}$ tal que*

$$\lim_{t \rightarrow a^+} \frac{1}{(t-a)} h_x(a, x_0) \cdot y_{\tilde{u}}(t) < 0. \quad (4.19)$$

(b) *Suponhamos que a qualificação de restrições do Teorema 4.1.3 é satisfeita. Então $\tilde{u}$ pode ser escolhido de forma a satisfazer 4.19 e ainda*

$$h_x(t, \bar{x}(t)) \cdot y_{\tilde{u}}(t) < 0 \quad \forall t \in (a, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\} \quad (4.20)$$

e

$$d_x^i(\bar{x}(b)) \cdot y_{\tilde{u}}(b) < 0 \quad \text{se } d^i(\bar{x}(b)) = 0. \quad (4.21)$$

Uma contribuição essencial desta proposição é a possibilidade de escolher $\tilde{u}$ de forma a que o limite na desigualdade 4.19 existe. Em geral, $\tilde{u}$ não poderia ser tomado igual a u' das qualificações das restrições, uma vez que este apenas satisfaz

$$\limsup_{t \rightarrow a^+} \frac{1}{(t-a)} \cdot y_{u'}(t) < 0.$$

A propriedade de convexidade associada a $f(t, x, \Omega)$ é essencial na demonstração da proposição.

Demonstração.

(a) Seja u' a função controlo de $(CQ)_1$. Esta condição implica a existência de $\bar{t} \in (a, b]$ e $\delta > 0$ tal que

$$h_x(a, x_0) \cdot \Delta f_{u'}(t) < -\delta, \quad a.e. \ t \in [a, \bar{t}]. \quad (4.22)$$

Tome-se $\tau \in (a, \bar{t}]$. Defina-se, *a.e.*, $\beta : [a, \tau] \rightarrow \mathbb{R}$

$$\beta(t) := -[h_x(a, x_0) \cdot \Delta f_{u'}(t)]^{-1} \frac{\delta}{2} \quad (4.23)$$

A função $\beta(t)$ toma valores, inferiores a $1/2$, *a.e.* $t \in [a, \tau]$.

Tendo em conta a convexidade de $f(t, \bar{x}(t), \Omega)$ e aplicando uma argumentação do tipo efectuado na parte final da demonstração da Proposição 4.2.1, deduz-se a existência de $u_\tau \in \mathcal{U}$, tal que

$$f(t, \bar{x}(t), u_\tau(t)) = f(t, \bar{x}(t), \bar{u}(t)) + \beta(t) \Delta f_{u'}(t) \quad a.e. \ t \in [a, \tau].$$

Na parte restante do intervalo $[a, b]$, defina-se

$$u_\tau(t) = u'(t) \quad a.e. \text{ em } (\tau, b].$$

Por 4.23

$$h_x(a, x_0) \cdot \Delta f_{u_\tau}(t) = \beta(t) h_x(a, x_0) \cdot \Delta f_{u'}(t) = -\frac{\delta}{2} \quad a.e. \ t \in [a, \tau]. \quad (4.24)$$

Da definição de $u_\tau(t)$ em $(\tau, \bar{t}]$ e de 4.22 vem ainda,

$$h_x(a, x_0) \cdot \Delta f_{u_\tau}(t) < -\frac{\delta}{2} \quad a.e. \ t \in (\tau, \bar{t}]. \quad (4.25)$$

Por 4.24,

$$\frac{1}{t-a} h_x(a, x_0) \int_a^t \Delta f_{u_\tau}(s) ds = -\frac{\delta}{2}, \quad \text{para todo } t \in (a, \tau).$$

Contudo, pelas propriedades da matriz de transição Φ ,

$$\begin{aligned} -\frac{\delta}{2} &= \lim_{t \rightarrow a^+} (t-a)^{-1} h_x(a, x_0) \int_a^t \Delta f_{u_\tau}(s) ds \\ &= \lim_{t \rightarrow a^+} (t-a)^{-1} h_x(a, x_0) \int_a^t \Phi(t, s) \Delta f_{u_\tau}(s) ds \\ &= \lim_{t \rightarrow a^+} (t-a)^{-1} h_x(a, x_0) \cdot y_{u_\tau}. \end{aligned}$$

A parte (a) da proposição fica então demonstrada, com $\tilde{u} = u_\tau$.

(b) Seja u' a função controlo de $(CQ)_3$. Fixe-se novamente $\bar{t} \in (a, b]$ e $\delta > 0$ de forma a serem satisfeitas, além da condição 4.22, a condição seguinte:

$$\left| (t-a)^{-1} h_x(a, x_0) \int_a^t \Delta f_{u'}(s) ds - (t-a)^{-1} h_x(t, \bar{x}(t)) \int_a^t \Phi(t, s) \Delta f_{u'}(s) ds \right| < \frac{\delta}{4} \quad \text{para todo } t \in (a, \bar{t}]. \quad (4.26)$$

As propriedades de u' implicam a existência de $\beta > 0$ tal que

$$(t-a)^{-1} h_x(t, \bar{x}(t)) y_{u'}(t) < -\beta \quad \forall t \in [\bar{t}, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}.$$

Para $\tau \in (a, \bar{t}]$, seja u_τ , como na parte (a). A aplicação do Lema de Gronwell [27], a y_{u_τ} permite concluir:

$$y_{u_\tau} \rightarrow y_{u'} \quad \text{uniformemente quando } \tau \rightarrow a^+.$$

Fixe-se então um valor de τ em $(a, \bar{t}]$ tal que

$$(t-a)^{-1} h_x(t, \bar{x}(t)) y_{u_\tau}(t) < -\frac{\beta}{2} \quad \forall t \in [\bar{t}, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\} \quad (4.27)$$

e

$$d_x^i(\bar{x}(b)) \cdot y_{u_\tau}(b) < 0 \quad \text{para } i \in I_0 := \{i : d^i(\bar{x}(b)) = 0\}. \quad (4.28)$$

Por 4.24 e 4.25

$$(t-a)^{-1} h_x(a, x_0) \int_a^t \Delta f_{u_\tau}(s) ds \leq -\frac{\delta}{2}, \quad \text{para todo } t \in (a, \bar{t}].$$

Por 4.26 vem então

$$(t-a)^{-1} h_x(t, \bar{x}(t)) \cdot y_{u_\tau}(t) < -\frac{\delta}{4} \quad \forall t \in (a, \bar{t}]. \quad (4.29)$$

Tome-se agora $\tilde{u} = u_\tau$. Como se viu em (a), $\tilde{u}$ satisfaz a primeira condição em $(CQ)_3$. De 4.27, 4.28 e 4.29 também se conclui que satisfaz a segunda. ■

4.2.2 Demonstração dos Teoremas 4.1.1 e 4.1.3

Suponhamos que a condição $(CQ)_1$ se verifica. Defina-se:

$$\mathcal{U}^* := \{u \in \mathcal{U} : t = a \text{ é um ponto de Lebesgue de } t \rightarrow h_x(a, x_0) \cdot \Delta f_u(t)\}$$

Por definição de ponto de Lebesgue [25],

$$\lim_{t \rightarrow a^+} \frac{1}{t-a} \int_a^t h_x(a, x_0) \cdot \Delta f_u(s) ds = h_x(a, x_0) \cdot \Delta f_u(a).$$

Tome-se $u \in \mathcal{U}^*$. É fácil verificar que,

$$(t-a)^{-1} h_x(t, \bar{x}(t)) \cdot y_u(t) \rightarrow h_x(a, x_0) \cdot \Delta f_u(a) \text{ quando } t \downarrow a.$$

Represente-se por $\eta_u : [a, b] \rightarrow \mathfrak{R}$, a extensão contínua de $(t-a)^{-1} h_x(t, \bar{x}(t)) \cdot y_u(t)$ definida em $(a, b]$, ao intervalo $[a, b]$, ou seja

$$\eta_u(t) := \begin{cases} (t-a)^{-1} h_x(t, \bar{x}(t)) \cdot y_u(t) & t \in (a, b] \\ h_x(a, x_0) \cdot \Delta f_u(a) & t = a, \end{cases}$$

e defina-se,

$$G_0 = \{t \in [a, b] : h(t, \bar{x}(t)) = 0\} \quad I_0 = \{i : d^i(\bar{x}(b)) = 0\}.$$

A condição $(CQ)_1$ implica que o conjunto G_0 é não vazio. Contudo, I_0 pode ser vazio.

Proposição 4.2.4 *Para qualquer $u \in \mathcal{U}^*$, tem-se*

$$\left(\max_{t \in G_0} \{\eta_u(t)\} \right) \vee \left(\bigvee_{i \in I_0} d_x^i(\bar{x}(b)) \cdot y_u(b) \right) \vee (g_x(\bar{x}(b)) \cdot y_u(b)) \geq 0$$

onde $e \vee \dots \vee f$ representa o máximo de $e, \dots, f$. No caso de $I_0 = \emptyset$, o respectivo termo na inequação deve ser tomado como $-\infty$.

Demonstração.

Fixemos $u \in \mathcal{U}^*$ e seja $\epsilon' > 0$ como na Proposição 4.2.1. Represente-se por (x_ϵ, u_ϵ) o processo associado a u , com $\epsilon \in [0, \epsilon']$, tal como foi definido pela mesma proposição.

Dado $\delta > 0$, defina-se

$$G_\delta := \{t \in [a, b] : h(t, \bar{x}(t)) \geq -\delta\}.$$

Como $(\bar{x}, \bar{u})$ é um processo admissível, $h(t, \bar{x}(t)) \leq 0 \forall t \in [a, b]$. O conjunto G_δ selecciona os instantes nos quais a trajectória se aproxima, a menos de δ , da fronteira deste conjunto de admissibilidade. Para $t \in [a, b] \setminus G_\delta$, a trajectória afasta-se dessa

mesma fronteira, mais precisamente, $h(t, \bar{x}(t)) < -\delta < 0$.

A continuidade da função h permite determinar uma vizinhança em torno do gráfico de $\bar{x}$ de forma a que, qualquer trajectória cujo gráfico caia nessa vizinhança, é transformada por h em valores que distam dos correspondentes valores para $\bar{x}$ menos de $\delta/2$, por exemplo. Neste caso, trajectórias deste tipo mantêm-se afastadas da fronteira mencionada atrás, quando $\bar{x}$ de alguma forma também o está.

Reduzindo, se necessário, o tamanho de ϵ' , as relações seguintes resultam facilmente da Proposição 4.2.1 e da continuidade das funções envolvidas.

$$h(t, x_\epsilon(t)) < 0 \quad \forall t \in [a, b] \setminus G_\delta, \quad \epsilon \in [0, \epsilon'],$$

$$d^i(x_\epsilon(b)) < 0 \quad \text{para } i \notin I_0, \quad \epsilon \in [0, \epsilon']$$

Para $\epsilon \in (0, \epsilon']$, tem-se então,

$$\left(\sup_{t \in G_\delta \setminus \{a\}} \left\{ (t-a)^{-1} [h(t, x_\epsilon(t)) - h(t, \bar{x}(t))] \right\} \right) \vee \left(\bigvee_{i \in I_0} [d^i(x_\epsilon(b)) - d^i(\bar{x}(b))] \right) \\ \vee (g(x_\epsilon(b)) - g(\bar{x}(b))) \geq 0, \quad (4.30)$$

De facto, se tal não acontecesse, (x_ϵ, u_ϵ) seria um processo admissível para (P) com um custo inferior a $(\bar{x}, \bar{u})$, o processo óptimo.

Quando x_ϵ é substituído por $\bar{x} + \epsilon y_u$, a continuidade uniforme e a diferenciabilidade das funções envolvidas permitem concluir que

$$\left(\sup_{t \in G_\delta \setminus \{a\}} \left\{ (t-a)^{-1} [h(t, \bar{x}(t) + \epsilon y_u(t)) - h(t, \bar{x}(t))] \right\} \right) \vee \\ \left(\bigvee_{i \in I_0} [d^i(\bar{x}(b) + \epsilon y_u(b)) - d^i(\bar{x}(b))] \right) \vee (g(\bar{x}(b) + \epsilon y_u(b)) - g(\bar{x}(b))) \geq -\epsilon \tilde{w}(\epsilon).$$

onde $\tilde{w}(\epsilon)$ é uma função positiva, crescente e satisfaz $\tilde{w}(\epsilon) \downarrow 0$ quando $\epsilon \downarrow 0$.

Desenvolvendo isoladamente cada uma das expressões desta desigualdade,

$$\frac{1}{t-a} [h(t, \bar{x}(t) + \epsilon y_u(t)) - h(t, \bar{x}(t))] = \frac{1}{t-a} h_x(t, \bar{x}(t) + \theta(t)\epsilon y_u(t)) \cdot \epsilon y_u(t) \\ \text{com } 0 < \theta(t) < 1 \\ = \epsilon \left[\frac{1}{t-a} h_x(t, \bar{x}(t)) \cdot y_u(t) + R(t, \epsilon) \right]$$

com,

$$R(t, \epsilon) = \frac{1}{t-a} [h_x(t, \bar{x}(t) + \theta(t)\epsilon y_u(t)) - h_x(t, \bar{x}(t))] \cdot y_u(t).$$

Das propriedades de h e y_u resulta que $\lim_{\epsilon \rightarrow 0} \sup_{t \in [a, b]} R(t, \epsilon) = 0$.

De forma análoga podemos escrever,

$$g(\bar{x}(b) + \epsilon y_u(b)) - g(\bar{x}(b)) = \epsilon g_x(\bar{x}(b)) \cdot y_u(b) + \epsilon G(\epsilon) \quad \text{com } \lim_{\epsilon \rightarrow 0} G(\epsilon) = 0$$

e

$$d^i(\bar{x}(b) + \epsilon y_u(b)) - d^i(\bar{x}(b)) = \epsilon d_x^i(\bar{x}(b)) \cdot y_u(b) + \epsilon D^i(\epsilon) \quad \text{com } \lim_{\epsilon \rightarrow 0} D^i(\epsilon) = 0$$

Destas conclusões e da anterior desigualdade para o supremo, vem

$$\left(\sup_{t \in G_\delta \setminus \{a\}} \left\{ (t-a)^{-1} [h_x(t, \bar{x}(t)) \cdot y_u(t)] \right\} \right) \vee \left(\bigvee_{i \in I_0} [d_x^i(\bar{x}(b)) \cdot y_u(b)] \right) \\ \vee (g_x(\bar{x}(b)) \cdot y_u(b)) \geq -\tilde{w}(\epsilon) - \max \left\{ \sup_{t \in [a, b]} R(t, \epsilon), G(\epsilon), D^i(\epsilon), i \in I_0 \right\}. \quad (4.31)$$

Aplicando limites, $\epsilon \rightarrow 0^+$, a ambos os termos da desigualdade, obtém-se,

$$\left(\sup_{t \in G_\delta \setminus \{a\}} \left\{ (t-a)^{-1} [h_x(t, \bar{x}(t)) \cdot y_u(t)] \right\} \right) \vee \left(\bigvee_{i \in I_0} [d_x^i(\bar{x}(b)) \cdot y_u(b)] \right) \\ \vee (g_x(\bar{x}(b)) \cdot y_u(b)) \geq 0. \quad (4.32)$$

Recorde-se que η_u é a extensão contínua da função com domínio $(a, b]$, definida por $t \rightarrow (t-a)^{-1} h_x(t, \bar{x}(t)) \cdot y_u(t)$, a todo o intervalo $[a, b]$. Como $[a, a+\tau] \subset G_\delta$ para $\tau > 0$, suficientemente pequeno, e $\eta_u(\cdot)$ é contínua, vem

$$\sup_{t \in G_\delta \setminus \{a\}} \left\{ (t-a)^{-1} [h_x(t, \bar{x}(t)) \cdot y_u(t)] \right\} = \max_{t \in G_\delta} \{\eta_u(t)\}. \quad (4.33)$$

Até agora δ foi considerado fixo. Considerando $\delta \downarrow 0$ e utilizando argumentos de compactidade tem-se

$$\max_{t \in G_\delta} \{\eta_u(t)\} \rightarrow \max_{t \in G_0} \{\eta_u(t)\} \quad (4.34)$$

A proposição pode então ser concluída de 4.32, 4.33 e 4.34. ■

Seja k_0 o número de elementos no conjunto de índices, I_0 .

Defina-se S como sendo o conjunto:

$$S := \begin{cases} \{(\eta_u, \eta_u^0, \eta_u^1) \in C(G_0) \times \mathbb{R}^{k_0} \times \mathbb{R} : u \in \mathcal{U}^*\} & \text{se } k_0 \geq 1 \\ \{(\eta_u, \eta_u^1) \in C(G_0) \times \mathbb{R} : u \in \mathcal{U}^*\} & \text{se } k_0 = 0 \end{cases}$$

para η_u tal como foi definida atrás e

$$\eta_u^0 := \{d_x^i(\bar{x}(b)) \cdot y_u(b) : i \in I_0\} \quad \eta_u^1 := g_x(\bar{x}(b)) \cdot y_u(b).$$

$C(G_0)$ representa o espaço das funções contínuas definidas em G_0 e tomando valores em $\mathbb{R}$.

A proposição seguinte além de afirmar a convexidade de S , exprime a desigualdade da proposição anterior em termos da disjunção de conjuntos.

Proposição 4.2.5 *S é um conjunto convexo e, denotando por K^- o cone aberto negativo em $C(G_0) \times \mathbb{R}^{k_0} \times \mathbb{R}$ (ou $C(G_0) \times \mathbb{R}$ se $I_0 = \emptyset$), ou seja*

$$K^- := \begin{cases} \{(\phi, \phi^0, \phi^1) \in C(G_0) \times \mathbb{R}^{k_0} \times \mathbb{R} : \phi(t) < 0 \text{ em } G_0, \\ \quad \phi^0 < 0, \phi^1 < 0\}, & \text{se } k_0 \geq 1 \\ \{(\phi, \phi^1) \in C(G_0) \times \mathbb{R} : \phi(t) < 0 \text{ em } G_0, \phi^1 < 0\}, & \text{se } k_0 = 0, \end{cases}$$

tem-se

$$S \cap K^- = \emptyset.$$

Demonstração.

O facto de S e K^- terem intersecção vazia resulta imediatamente da Prop. 4.2.4. Resta provar a convexidade de S . Para isso, tome-se $u', u'' \in \mathcal{U}^*$ e $\epsilon \in [0, 1]$. A convexidade de $f(t, \bar{x}(t), \Omega)$ e argumentos já utilizados neste trabalho sobre selecções mensuráveis, permitem afirmar a existência de $u \in \mathcal{U}$ tal que

$$f(t, \bar{x}(t), u(t)) = \epsilon f(t, \bar{x}(t), u'(t)) + (1 - \epsilon) f(t, \bar{x}(t), u''(t)).$$

Que o ponto $t = a$ é um ponto de Lebesgue de $h_x(a, x_0) \cdot \Delta f_u(t)$ resulta facilmente do facto de $u', u'' \in \mathcal{U}^*$ e da identidade

$$\begin{aligned} (t - a)^{-1} \int_a^t h_x(a, x_0) \cdot \Delta f_u(s) \, ds &= \epsilon \left\{ (t - a)^{-1} \int_a^t h_x(a, x_0) \cdot \Delta f_{u'}(s) \, ds \right\} \\ &+ (1 - \epsilon) \left\{ (t - a)^{-1} \int_a^t h_x(a, x_0) \cdot \Delta f_{u''}(s) \, ds \right\} \end{aligned}$$

que é verdadeira para qualquer $t \in (a, b]$. É fácil concluir agora que $(\eta_u, \eta_u^0, \eta_u^1) = \epsilon(\eta_{u'}, \eta_{u'}^0, \eta_{u'}^1) + (1 - \epsilon)(\eta_{u''}, \eta_{u''}^0, \eta_{u''}^1)$ e daqui a convexidade de S . ■

Proposição 4.2.6 *Existe $(\nu, \alpha, \lambda) \in C^*([a, b]) \times \mathbb{R}^k \times \mathbb{R}$ tal que*

$$\nu \geq 0, \alpha \geq 0, \lambda \geq 0$$

$$\text{supp}\{\nu\} \subset \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}$$

$$\nu\{(a, b]\} + |\alpha| + \lambda \neq 0,$$

$$\alpha_i = 0 \text{ se } d^i(\bar{x}(b)) < 0$$

e

$$\int_{[a, b]} \eta_u(t) \nu(dt) + \alpha^T d_x(\bar{x}(b)) \cdot y_u(b) + \lambda g_x(\bar{x}(b)) \cdot y_u(b) \geq 0 \quad (4.35)$$

para todo $u \in \mathcal{U}^*$.

Demonstração.

Os conjuntos convexos S e K^- são disjuntos, pela Prop. 4.2.5. Uma vez que K^- tem interior não vazio, os conjuntos podem ser separados por um hiperplano fechado, cuja normal é representada por um elemento não nulo $(\tilde{\nu}, \alpha, \lambda) \in C^*(G_0) \times \mathbb{R}^k \times \mathbb{R}$, com $\alpha_i = 0$ para $i \notin I_0$. Tem-se então que,

$$\int_{G_0} \phi(t) \tilde{\nu}(dt) + \alpha^T \phi^0 + \lambda \phi^1 \geq 0 \quad \forall (\phi(\cdot), \phi^0, \phi^1) \in S \quad (4.36)$$

$$\int_{G_0} \phi(t) \tilde{\nu}(dt) + \alpha^T \phi^0 + \lambda \phi^1 \leq 0 \quad \forall (\phi(\cdot), \phi^0, \phi^1) \in K^-. \quad (4.37)$$

Estendendo $\tilde{\nu}$ como uma medida de Borel regular ν aos subconjuntos de Borel de $[a, b] : \nu(B) = \tilde{\nu}(B \cap G_0)$, vem

$$\text{supp}\{\nu\} \subset G_0 \text{ e } \int_{G_0} \eta_u(t) \tilde{\nu}(dt) = \int_{[a, b]} \eta_u(t) \nu(dt)$$

para todo $u \in \mathcal{U}^*$.

Estes factos e as equações 4.36, 4.37 estabelecem a desigualdade 4.35 e ainda que $\nu \geq 0, \alpha \geq 0, \lambda \geq 0$.

Resta verificar que $\nu\{(a, b]\} + |\alpha| + \lambda \neq 0$. Como $\nu\{(a, b]\} + |\alpha| + \lambda \neq 0$, basta verificar que

$$(\nu, \alpha, \lambda) \neq (\gamma \delta_{\{a\}}, 0, 0),$$

para algum $\gamma > 0$.

Suponhamos, por redução ao absurdo, que $(\nu, \alpha, \lambda) = (\gamma \delta_{\{a\}}, 0, 0)$, com $\gamma > 0$. Considere-se a função controlo $\tilde{u}$ da Prop. 4.2.3. Esta função controlo satisfaz, $\tilde{u} \in \mathcal{U}^*$ e $\eta_{\tilde{u}}(a) < 0$. Então, para tais multiplicadores,

$$\int_{[a,b]} \eta_{\tilde{u}}(t) \nu(dt) + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \cdot y_{\tilde{u}}(b) = \gamma \eta_{\tilde{u}}(a) < 0.$$

Mas, por 4.35, isto é um absurdo.

A proposição está provada. ■

Proposição 4.2.7 *Seja $q : (a, b] \rightarrow \mathbb{R}^n$ a função mensurável definida por*

$$q(t) = \int_{[t,b]} h_x(s, \bar{x}(s)) \Phi(s, t) (s - a)^{-1} \nu(ds) + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \Phi(b, t).$$

Então,

$$q(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) = \min_{u \in \Omega} q(t) \cdot f(t, \bar{x}(t), u) \quad a.e. \ t \in (a, b]. \quad (4.38)$$

Demonstração.

Tome-se uma sucessão $\{a_i\}$ em $(a, b]$ com $a_i \downarrow a$. Defina-se

$$\mathcal{U}_i := \{ \text{funções mensuráveis } u : [a_i, b] \rightarrow \mathbb{R} : u(t) \in \Omega \text{ a.e.} \}.$$

Fixemos i e considere-se $u \in \mathcal{U}_i$. Tome-se o elemento de $\mathcal{U}^*$ que é igual a $\bar{u}(t)$ para $t \in [a, a_i]$ e igual a $u(t)$ para $t \in [a_i, b]$. A desigualdade 4.35 da Prop. 4.2.6 fica então

$$\int_{[a_i,b]} h_x(t, \bar{x}(t)) \cdot y_u^i(t) (t - a)^{-1} \nu(dt) + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \cdot y_u^i(b) \geq 0, \quad (4.39)$$

onde y_u^i representa o elemento em $AC^1([a_i, b]; \mathbb{R}^n)$ que satisfaz

$$\begin{cases} \dot{y}_u^i(t) &= f_x(t, \bar{x}(t), \bar{u}(t)) \cdot y_u^i(t) + \Delta f_u(t) \quad a.e. \ t \in [a_i, b] \\ y_u^i(a_i) &= 0 \end{cases}$$

Consideremos agora a Prop. 4.2.2, adaptada de forma a ser aplicada ao intervalo $[a_i, b]$ e quando a medida em questão é $(t - a)^{-1} \nu(dt)$, com domínio nos subconjuntos de Borel de $[a_i, b]$. Desta proposição e de 4.39 resulta

$$\int_{[a_i,b]} q(t) \cdot [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))] dt \geq 0$$

para todo $u \in \mathcal{U}_i$. Esta desigualdade implica 4.38 num subconjunto de $[a_i, b]$ de medida total. Como isto é válido para todo o i e $a_i \downarrow a$, podemos concluir que é válido num subconjunto de $(a, b]$ de medida total. ■

Proposição 4.2.8

$$\lim_{t \downarrow a} \int_{[t,b]} (s-a)^{-1} \nu(ds) < \infty$$

Demonstração.

Pela hipótese $(CQ)_1$ e Prop. 4.2.7, podemos escolher $\delta > 0$, $t' \in (a, b]$ e uma sucessão $\{t_i\} \subset (a, t']$, com $t_i \downarrow a$, tal que

$$h_x(s, \bar{x}(s))\Phi(s, t_i) \cdot \Delta f_{u'}(t_i) < -\delta \quad (4.40)$$

$$q(t_i) \cdot \Delta f_{u'}(t_i) \geq 0.$$

para todo $s \in [a, t']$, para $i = 1, 2, \dots$

Nestas expressões, u' é a função controlo de $(CQ)_1$ e q está definida como na Prop. 4.2.7. A última desigualdade pode ser escrita de forma equivalente,

$$a_i + b_i \geq 0 \quad \text{para } i = 1, 2, \dots \quad (4.41)$$

onde

$$a_i := \int_{[t_i, t']} h_x(s, \bar{x}(s))\Phi(s, t_i) \Delta f_{u'}(t_i) (s-a)^{-1} \nu(ds)$$

$$b_i := \left[\int_{(t', b]} h_x(s, \bar{x}(s))\Phi(s, t_i) (s-a)^{-1} \nu(ds) + (\alpha^T d_x + \lambda g_x)\Phi(b, t) \right] \cdot \Delta f_{u'}(t_i).$$

Escolha-se uma constante K tal que

$$|b_i| < K \left(\int_{(t', b]} (s-a)^{-1} \nu(ds) + |\alpha| + \lambda \right) \quad i = 1, 2, \dots$$

Por 4.40,

$$a_i < -\delta \int_{[t_i, t']} (s-a)^{-1} \nu(ds) \quad i = 1, 2, \dots$$

Estas desigualdades e 4.41 dão

$$\int_{[t_i, b]} (s-a)^{-1} \nu(ds) < \delta^{-1} K \left(\int_{(t', b]} (s-a)^{-1} \nu(ds) + |\alpha| + \lambda \right).$$

Deduz-se então,

$$\begin{aligned} \int_{[t_i, b]} (s-a)^{-1} \nu(ds) &= \int_{[t_i, t']} (s-a)^{-1} \nu(ds) + \int_{(t', b]} (s-a)^{-1} \nu(ds) \\ &\leq (1 + \delta^{-1} K) \int_{(t', b]} (s-a)^{-1} \nu(ds) + \delta^{-1} K (|\alpha| + \lambda) \end{aligned}$$

Uma vez que o segundo membro é independente de i , e a função definida por $t \rightarrow \int_{[t,b]} (s-a)^{-1} \nu(ds)$ é monótona decrescente, vem

$$\lim_{t_i \downarrow a} \int_{[t_i,b]} (s-a)^{-1} \nu(ds) < \infty.$$

A proposição fica assim demonstrada. ■

Temos agora tudo o que é necessário para a demonstração do Teor. 4.1.1.

Pela Prop. 4.2.8 a função $s \rightarrow (s-a)^{-1}$ é integrável em $(a, b]$, em relação à medida regular de Borel ν . A medida que a seguir é apresentada está bem definida e é ainda uma medida regular de Borel:

$$\mu(A) := \int_{(a,b] \cap A} (s-a)^{-1} \nu(ds).$$

Como $\nu \geq 0$, tem-se também

$$\mu \geq 0. \quad (4.42)$$

Uma vez que μ é absolutamente contínua em relação a ν , vem

$$\text{supp}\{\mu\} \subset \text{supp}\{\nu\} \subset \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}. \quad (4.43)$$

Da definição de μ resulta ainda que $\mu\{(a, b]\} \geq (b-a)^{-1} \nu\{(a, b]\}$. Da Prop. 4.2.6 conclui-se então,

$$\mu\{(a, b]\} + |\alpha| + \lambda \neq 0. \quad (4.44)$$

A função $q : (a, b] \rightarrow \mathbb{R}^n$ definida na Prop. 4.2.7 pode ser expressa em termos de μ da seguinte forma:

$$q(t) = \int_{[t,b]} h_x(s, \bar{x}(s)) \Phi(s, t) \mu(ds) + [\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \Phi(b, t).$$

Como o segundo membro desta igualdade está definido para todo $t \in [a, b]$, podemos prolongar a função a todo o intervalo e q passará a designar esta extensão. Designe-se ainda por $p \in AC^1([a, b]; \mathbb{R}^n)$, a solução da equação diferencial

$$\begin{aligned} -\dot{p}(t) &= \left[p(t) + \int_{[a,t]} h_x(s, \bar{x}(s)) \mu(ds) \right] f_x(t, \bar{x}(t), \bar{u}(t)) \quad \text{a.e. } t \in [a, b] \\ &\quad - \left(p(b) + \int_{[a,b]} h_x(s, \bar{x}(s)) \mu(ds) \right) = \alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b)). \end{aligned} \quad (4.45)$$

De acordo com a Prop. 4.2.2 vem,

$$q(t) := - \left[p(t) + \int_{[a,t]} h_x(s, \bar{x}(s)) \mu(ds) \right] \quad \text{para todo } t \in [a, b]. \quad (4.46)$$

As equações 4.38 e 4.42 - 4.46 permitem concluir o Teorema 4.1.1.

Para demonstrar o Teorema 4.1.3 considere-se agora que $(CQ)_3$ se verifica. É necessário provar que os multiplicadores (μ, α, λ) podem ser escolhidos com $\lambda \neq 0$. Como $(CQ)_3$ é mais forte que $(CQ)_1$, todas as conclusões obtidas quando consideramos a primeira condição continuam a ser válidas e é possível construir multiplicadores (μ, α, λ) como anteriormente. Contudo, com $(CQ)_3$, a função controlo $\bar{u}$ da Prop. 4.2.3(b) pertence a $\mathcal{U}^*$ e satisfaz

$$\eta_{\bar{u}}(t) < 0 \quad \forall t \in [a, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}$$

$$d_x^i(\bar{x}(t)) \cdot y_{\bar{u}}(b) < 0, \quad \text{para } i \in I_0.$$

Estas condições contradizem 4.35 a menos que $\lambda \neq 0$.

A demonstração do Teorema 4.1.3 está terminada.

4.2.3 Demonstração dos Teoremas 4.1.2 e 4.1.4

Comecemos por demonstrar o Teorema 4.1.2 e para isso consideremos a condição $(CQ)_2$ satisfeita.

Tome-se uma sucessão $\{t_i\}$ em $(a, \bar{t}]$ com $t_i \downarrow a$, onde $\bar{t}$ é o ponto em $(a, b]$ referido em $(CQ)_2$. Para $i = 1, 2, \dots$ defina-se (P_i) o problema semelhante ao problema (P) mas com o intervalo sendo agora $[t_i, b]$ e com condição inicial,

$$x(t_i) = \bar{x}(t_i).$$

Para cada i , o processo para (P_i) composto pelas restrições de $\bar{u}$ e $\bar{x}$ ao intervalo $[t_i, b]$, é um processo minimizante para (P_i) . Consideremos multiplicadores $(\tilde{p}_i, \tilde{\mu}_i, \tilde{\alpha}_i, \tilde{\lambda}_i)$ associados a este processo, cuja existência é garantida pelo Princípio do Máximo convencional (ver §3). Da condição $(CQ)_2$ resulta,

$$\text{supp}\{\tilde{\mu}_i\} \subset [\bar{t}, b], \quad \forall i. \quad (4.47)$$

Para cada i , seja μ_i a extensão de $\tilde{\mu}_i$ aos subconjuntos de Borel do intervalo $[a, b]$,

$$\mu_i(A) := \tilde{\mu}_i(A \cap [t_i, b]), \quad (4.48)$$

e p_i a extensão de $\tilde{p}_i$ ao intervalo $[a, b]$,

$$p_i(t) := \begin{cases} \tilde{p}_i(t_i) & \text{para } a \leq t \leq t_i \\ \tilde{p}_i(t) & \text{para } t_i \leq t \leq b. \end{cases} \quad (4.49)$$

Escalando os multiplicadores, se necessário, podemos considerar

$$\int_{[a,b]} \mu_i ds + \sum_{j \in I_0} (\alpha^j)_i + \lambda_i = 1 \quad (4.50)$$

onde $(a^j)_i$ é a j -ésima componente de a_i . Extraindo subsucessões, podemos afirmar que

$$p_i \rightarrow p \text{ uniformemente}, \quad \mu_i \rightarrow \mu \text{ fracamente}^*, \quad \alpha_i \rightarrow \alpha \text{ e } \lambda_i \rightarrow \lambda$$

para um certo $(p, \mu, \alpha, \lambda)$, que serão multiplicadores associados a $(\bar{x}, \bar{u})$. (Esta análise de convergência é bem conhecida da literatura, ver por exemplo [26].)

Consideremos a função contínua

$$\phi(t) := \begin{cases} \frac{t-a}{\bar{t}-a} & a \leq t \leq \bar{t} \\ 1 & \bar{t} \leq t \leq b. \end{cases}$$

De acordo com 4.47 e 4.48,

$$\int_{[a,b]} \mu_i(ds) = \int_{[a,b]} \phi(s) \mu_i(ds).$$

Da convergência *fraca*^{*} da sucessão μ_i e de 4.50, resulta

$$\int_{[a,b]} \phi(s) \mu(ds) + \sum_{j \in I_0} (\alpha^j) + \lambda = 1.$$

Uma vez que $\phi(a) = 0$, podemos concluir

$$\int_{(a,b]} \mu(ds) + \sum_{j \in I_0} (\alpha^j) + \lambda \neq 0.$$

O Teorema 4.1.2 fica assim demonstrado.

Suponhamos agora que a condição $(CQ)_2$ é substituída por $(CQ)_4$. Seja u' a função controlo que aparece nesta condição. Então, existe $\epsilon > 0$ tal que

$$h_x(t, \bar{x}(t)) \cdot y_{u'}(t) < -\epsilon \quad \text{em } [\bar{t}, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\} \quad (4.51)$$

$$d_x^i(\bar{x}(b)) \cdot y_{u'}(b) < -\epsilon \quad \text{para } i \in I_0. \quad (4.52)$$

Sejam $(\tilde{p}_i, \tilde{\mu}_i, \alpha_i, \lambda_i)$ os multiplicadores associados ao problema (P_i) , como anteriormente, e sejam p_i, μ_i as extensões tal como foram definidos em 4.48 e 4.49, para $i = 1, 2, \dots$. Como foi já mencionado, $(p_i, \mu_i, \alpha_i, \lambda_i)$ converge, no sentido referido, para $(p, \mu, \alpha, \lambda)$, que são multiplicadores associados a $(\bar{x}, \bar{u})$.

A Prop. 4.2.2 permite concluir, pelo facto de $(\tilde{p}_i, \tilde{\mu}_i, \alpha_i, \lambda_i)$ serem multiplicadores para um minimizante de (P_i) ,

$$\int_{[t_i, b]} h_x \cdot y_{u'}(s) \tilde{\mu}_i(ds) + (\alpha_i^T d_x + \lambda_i g_x) \cdot y_{u'}(b) \geq 0. \quad (4.53)$$

Contudo, por 4.47, 4.48 e 4.51

$$\int_{[t_i, b]} h_x(s, \tilde{x}(s)) \cdot y_{u'}(s) \mu_i(ds) \leq -\epsilon \int_{[a, b]} \mu_i(ds). \quad (4.54)$$

Resulta então, de 4.53 e 4.54

$$-\epsilon \int_{[a, b]} \mu_i(ds) - \epsilon \sum_{j \in I_0} (\alpha^j)_i + \lambda_i g_x \cdot y_{u'}(b) \geq 0,$$

para cada i . No limite obtém-se,

$$-\epsilon \int_{[a, b]} \mu(ds) - \epsilon \sum_{j \in I_0} (\alpha^j) + \lambda g_x \cdot y_{u'}(b) \geq 0.$$

Esta desigualdade implica que $\lambda > 0$, pois de outra forma α e μ deveriam ser ambos nulos, o que é um absurdo.

Capítulo 5

Condições Necessárias no Cálculo Variacional

As condições de optimalidade desenvolvidas no Capítulo 4, aplicam-se em situações onde $h(a, x_0) = 0$ e, ou o estado associado à trajectória minimizante abandona imediatamente a fronteira do conjunto de restrições de estado ou, em certo sentido, permanece nessa fronteira. De uma forma geral quando deparamos com um problema de controlo óptimo não temos à priori uma informação deste tipo acerca do minimizante. Contudo, para certas classes de problemas, estes resultados permitem estabelecer a existência de multiplicadores não triviais a partir de hipóteses facilmente verificáveis.

Vamos considerar neste capítulo um problema de uma tal classe. Mediante uma certa hipótese, fácil de testar a partir dos dados do problema, é possível estabelecer condições necessárias não triviais, no sentido que temos vindo a tratar. O problema insere-se no Cálculo das Variações, mas pode ser reformulado como um problema de controlo óptimo. Feita a reformulação, os resultados do capítulo anterior são aplicáveis.

Começamos por introduzir o problema clássico do Cálculo das Variações e possíveis variantes deste problema. É apresentado um resultado sobre existência de solução e são descritas condições necessárias sobre a forma de uma equação de Euler.

Existe uma vasta literatura sobre este assunto, da qual destacamos [4], [15], [5], [16]. Em particular o livro de Cesari [5], apresenta uma grande quantidade de exemplos destes problemas.

Na segunda parte do capítulo formulamos um problema particular que vai ser estudado e ao qual vão ser aplicados os resultados do capítulo anterior.

5.1 Problema do Cálculo Variacional

O problema mais simples do Cálculo das Variações consiste em minimizar o integral,

$$I[x] = \int_a^b L(t, x(t), \dot{x}(t)) dt \quad (5.1)$$

na classe de todas as funções absolutamente contínuas,

$$x(t) \quad a \leq t \leq b$$

que satisfazem as condições fronteira,

$$x(a) = x_0 \quad x(b) = x_1, \quad (5.2)$$

onde $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ e x_0, x_1 são pontos dados em $\mathbb{R}^n$.

Uma função absolutamente contínua, $x(t)$, $a \leq t \leq b$, tal que $L(\cdot, x(\cdot), \dot{x}(\cdot))$ é integrável de Lebesgue no intervalo $[a, b]$, satisfazendo as restrições do problema, neste caso a restrição 5.2, diz-se uma trajectória admissível. O problema terá solução se existir uma trajectória admissível $\bar{x}(t)$, $a \leq t \leq b$, tal que $I[\bar{x}] \leq I[x]$ para qualquer outra trajectória admissível $x(t)$, $a \leq t \leq b$. Neste caso, $\bar{x}$ dir-se-á um mínimo (ou um minimizante) para o problema.

Problemas mais complexos podem ainda incluir restrições sobre o estado,

$$(t, x(t)) \in A, \quad a \leq t \leq b \quad (A \subset \mathbb{R}^{1+n})$$

restrições designadas por *isoperimétricas*,

$$\int_a^b f_j(t, x(t), \dot{x}(t)) dt = C_j \quad j = 1, \dots, N$$

restrições com funcionais de comparação,

$$\int_a^b H(t, x(t), \dot{x}(t)) dt \leq C.$$

Os estados, $x(a)$ e $x(b)$, podem ainda estar sujeitos a restrições mais genéricas, por exemplo pertencer a conjuntos inicialmente dados.

Uma propriedade característica do problema clássico do Cálculo das Variações consiste em supor que todas as funções nele intervenientes são continuamente diferenciáveis, pelo menos.

Se a função $L(t, x, v)$, da expressão 5.1, é limitada inferiormente e convexa em v , o problema variacional diz-se *regular*. Esta propriedade permite, com uma hipótese adicional, estabelecer o resultado seguinte sobre existência de solução.

Teorema 5.1.1 *Se o problema,*

$$\text{Minimizar } \int_a^b L(t, x(t), \dot{x}(t)) dt, \quad x(a) = x_0, \quad x(b) = x_1,$$

é regular e satisfaz

$$\lim_{|v| \rightarrow \infty} \frac{|v|}{L(t, x, v)} = 0,$$

onde t e x são arbitrários mas fixos, então tem uma solução $x(t)$, $a \leq t \leq b$.

Considere-se agora o problema de minimizar $I[x]$, como em 5.1, restrito às condições, $x(a) = x_0$ e $(t, x(t)) \in A$, $A \subset \mathbb{R}^{1+n}$.

Resultados clássicos, ver por exemplo [5], estabelecem a seguinte condição necessária a ser satisfeita pela solução óptima do problema:

Teorema 5.1.2 *Seja A um subconjunto fechado de $\mathbb{R}^{n+1}$ que é o fecho dos seus pontos interiores e $L(t, x, v)$ continuamente diferenciável em $A \times \mathbb{R}^n$. Considere-se $I[x]$ definida na classe de todas as funções absolutamente contínuas, $x(t)$ $a \leq t \leq b$, com derivada, $\dot{x}(t)$, essencialmente limitada, satisfazendo $(t, x(t)) \in A$, $x(a) = x_0$ e tal que $L(\cdot, x(\cdot), \dot{x}(\cdot))$ é integrável de Lebesgue.*

Seja $\bar{x}$ uma trajectória que minimiza $I[x]$ nesta classe de funções e suponha-se que $\bar{x}(t)$ está no interior de A , para todo $t \in [a, b]$.

Então,

$$\frac{d}{dt} L_{\dot{x}}(t, \bar{x}(t), \dot{\bar{x}}(t)) = L_x(t, \bar{x}(t), \dot{\bar{x}}(t)) \quad \text{a.e. } t \in [a, b]. \quad (5.3)$$

A equação 5.3 é reconhecida na literatura como a equação de Euler.

Note-se que uma das hipóteses do teorema obriga a que a trajectória esteja no interior dos estados admissíveis.

Do ponto de vista matemático, o problema de controlo óptimo inclui o problema clássico do Cálculo das Variações. Este pode ser reformulado de forma a incluir uma variável de controlo, $u(t) \in \mathbb{R}^n$, que identificamos com $\dot{x}(t)$. Tal reformulação vai ser exemplificada nas secções seguintes.

Inversamente, muitos problemas de controlo óptimo podem ser reformulados como problemas clássicos do Cálculo das Variações. Tal reformulação faz surgir em muitos casos funções não diferenciáveis, não admissíveis na teoria clássica do Cálculo das Variações.

5.2 Condições Necessárias para um Problema do Cálculo Variacional

Considere-se o seguinte problema, com uma variável independente:

$$\text{Minimizar } \int_a^b L(x(t), \dot{x}(t)) dt \quad \text{sobre } x \in AC^1([a, b]; \mathbb{R}^n)$$

$$\begin{aligned}
& \text{tal que} \\
(CV) \quad & x(a) = x_0 \\
& \tilde{h}(x(t)) \leq 0
\end{aligned} \tag{5.4}$$

onde $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $\tilde{h} : \mathbb{R}^n \rightarrow \mathbb{R}$ e $x_0 \in \mathbb{R}^n$.

Os dados satisfazem as seguintes hipóteses:

($\tilde{H}1$) L é continuamente diferenciável.

($\tilde{H}2$) Para cada $x \in \mathbb{R}^n$, $L(x, \cdot)$ é convexa.

($\tilde{H}3$) Existe $\theta : [0, \infty) \rightarrow [0, \infty)$, crescente, com $\lim_{\alpha \rightarrow \infty} \frac{\theta(\alpha)}{\alpha} = +\infty$, e existem constantes β, γ tais que

$$L(x, v) > \theta(|v|) - \beta|v| - \gamma \quad \text{para todo } x \in \mathbb{R}^n, v \in \mathbb{R}^n.$$

($\tilde{H}4$) $\tilde{h}$ é duas vezes continuamente diferenciável.

($\tilde{H}5$) Existe $x \in AC^1([a, b]; \mathbb{R}^n)$ tal que $\tilde{h}(x(t)) \leq 0$, para todo $t \in [a, b]$.

O Teorema 5.1.1 continua a ser válido quando excluimos do problema nele descrito a restrição $x(b) = x_1$, ou seja o estado final passa a ser livre, e incluimos uma restrição de estado do tipo 5.4, desde que $\tilde{h}$ seja contínua [5]. A hipótese ($\tilde{H}3$) e a continuidade da função $L(x, v)$ garantem que as hipóteses do teorema são verificadas quando o problema em questão é (CV) e portanto, este tem um minimizante.

Embora se estejam a considerar trajectórias na classe $AC^1([a, b]; \mathbb{R}^n)$, é possível concluir que os arcos minimizantes para o problema (CV), são de facto funções Lipschitz contínuas.

Lema 5.2.1 *Se $\bar{x}$ é um arco minimizante para (CV) então $\bar{x}$ é uma função Lipschitz contínua.*

Demonstração.

Quando a restrição de estado 5.4 não existe, este lema é um caso particular dos teoremas de regularidade geral provados por Clarke e Vinter [10] e Ambrosio, Ascenzi e Butazzo [1]. Com o acréscimo da restrição de estado, o lema pode ser demonstrado com base nas técnicas de reparametrização usadas pelos últimos autores mencionados. Uma vez que estes resultados não são muito conhecidos, apresentamos a demonstração.

Considere-se o seguinte problema auxiliar (A):

$$\begin{aligned}
 & \text{Minimizar } \int_a^b l(s, v(s)) \, ds \\
 & \text{tal que} \\
 (A) \quad & \dot{y}(s) = v(s) \quad s \in [a, b] \quad \text{a.e.} \\
 & v(s) \in [0.5, 1.5] \\
 & y(a) = 0 \\
 & y(b) = b - a
 \end{aligned}$$

onde,

$$l(s, v) := L(\bar{x}(s), \dot{\bar{x}}(s)/v) v.$$

Provemos que $(\bar{y}(s) \equiv s - a, \bar{v}(s) \equiv 1)$ é um processo minimizante.

Dado qualquer processo admissível (y, v) para o problema (A), a função $\tau : [a, b] \rightarrow [a, b]$ definida por $\tau(s) = a + \int_a^s v(\sigma) \, d\sigma$ é Lipschitz contínua, estritamente crescente e tem inversa contínua. Efectuando a mudança de variável $t = \tau(s)$, a função custo toma a forma:

$$\int_a^b l(s, v(s)) \, ds = \int_a^b L\left(\bar{x}(s), \frac{\dot{\bar{x}}(s)}{v(s)}\right) v(s) \, ds = \int_a^b L(z(t), \dot{z}(t)) \, dt$$

onde $z(t) = (\bar{x} \circ \tau^{-1})(t)$. Para $v(s) = \bar{v}(s)$, vem

$$\int_a^b l(s, \bar{v}(s)) \, ds = \int_a^b L(\bar{x}(t), \dot{\bar{x}}(t)) \, dt.$$

A função z é absolutamente contínua e satisfaz, $h(z(t)) \leq 0$ para todo $t \in [a, b]$ e $z(a) = x_0$. Da optimalidade de $\bar{x}$ para o problema (CV) resulta

$$\int_a^b l(s, v(s)) \, ds \geq \int_a^b l(s, \bar{v}(s)) \, ds$$

e portanto $(\bar{y}, \bar{v})$ é um minimizante.

Para cada s fixo, $l(s, \cdot)$ é convexa em $(0, \infty)$. De facto, tomem-se $v_1, v_2 \in (0, \infty)$. Tendo em conta a Prop. 2.2.1 do Cap. 2, bastará provar que

$$l_v(s, v_1)(v_2 - v_1) \leq l(s, v_2) - l(s, v_1). \quad (5.5)$$

A convexidade de $L(\bar{x}(s), \cdot)$ permite escrever uma desigualdade análoga para esta função. (De forma a simplificar a notação, não escrevemos a componente x e o argumento s em $\dot{\bar{x}}(s)$.)

$$L_v\left(\frac{\dot{\bar{x}}}{v_1}\right)\left(\frac{\dot{\bar{x}}}{v_2} - \frac{\dot{\bar{x}}}{v_1}\right) \leq L\left(\frac{\dot{\bar{x}}}{v_2}\right) - L\left(\frac{\dot{\bar{x}}}{v_1}\right).$$

Subtraindo $L(\dot{\bar{x}}/v_1)v_1v_2^{-1}$ a ambos os termos da desigualdade e multiplicando pelo número positivo v_2 , obtém-se

$$\left[-L_v \left(\frac{\dot{\bar{x}}}{v_1} \right) \dot{\bar{x}}v_1^{-1} + L \left(\frac{\dot{\bar{x}}}{v_1} \right) \right] (v_2 - v_1) \leq L \left(\frac{\dot{\bar{x}}}{v_2} \right) v_2 - L \left(\frac{\dot{\bar{x}}}{v_1} \right) v_1.$$

Mas esta expressão é equivalente a 5.5. Fica assim provada a convexidade de $l(s, \cdot)$.

Para concluir a demonstração, consideremos o Princípio do Máximo aplicado ao problema (A). Daqui resulta a existência de um número p tal que a função

$$v \rightarrow pv - l(s, v)$$

definida em $[0.5, 1.5]$, é maximizada quando $v = 1$, para todos os pontos s num subconjunto $S \subset [a, b]$ de medida total. Uma vez que $l(s, \cdot)$ é convexa, maximização local implica maximização global e tem-se, para *quase todo* s ,

$$p - l(s, 1) \geq pv - l(s, v), \quad \text{para todo } v \in (0, \infty).$$

Seja $v = 1 + |\dot{\bar{x}}(s)|$. Substituindo na expressão anterior e considerando a definição de l , obtém-se

$$L(\bar{x}(s), \dot{\bar{x}}(s)) \leq p - p(1 + |\dot{\bar{x}}(s)|) + L(\bar{x}(s), \dot{\bar{x}}(s)/(1 + |\dot{\bar{x}}(s)|))(1 + |\dot{\bar{x}}(s)|) \quad a.e. .$$

Da hipótese ($\tilde{H}3$) sobre L resulta

$$\theta(|\dot{\bar{x}}(s)|) - \beta|\dot{\bar{x}}(s)| - \gamma \leq L(\bar{x}(s), \dot{\bar{x}}(s)) \quad a.e. .$$

Defina-se $k = \|L(\bar{x}, \dot{\bar{x}}/(1 + |\dot{\bar{x}}|))\|_{L^\infty}$, $k_1 = k - p + \beta$ e $k_2 = k + \gamma$.

L^∞ representa o espaço das funções $f : [a, b] \rightarrow \mathbb{R}^n$, mensuráveis e essencialmente limitadas. A norma neste espaço, $\|f\|_{L^\infty}$, é dada pelo supremo essencial do conjunto $\{|f(t)| : a \leq t \leq b\}$.

A relação seguinte é implicada pelas duas últimas desigualdades:

$$\theta(|\dot{\bar{x}}(s)|) \leq k_1|\dot{\bar{x}}(s)| + k_2 \quad a.e. . \quad (5.6)$$

De acordo com as propriedades de θ , podemos escolher $K > k_1$ e $\bar{\alpha} > 0$ tal que $\theta(\alpha) \geq K\alpha$ se $\alpha \geq \bar{\alpha}$ e $(K - k_1)\bar{\alpha} > k_2$. Seja $t \in [a, b]$ um ponto para o qual $\dot{\bar{x}}(t)$ existe e $|\dot{\bar{x}}(t)| \geq \bar{\alpha}$. Então, $(K - k_1)|\dot{\bar{x}}(t)| \geq (K - k_1)\bar{\alpha} > k_2$ e portanto

$$K|\dot{\bar{x}}(t)| > k_1|\dot{\bar{x}}(t)| + k_2 .$$

Mas, ainda pela definição de K e equação 5.6, se $|\dot{\bar{x}}(s)| \geq \bar{\alpha}$ vem

$$K|\dot{\bar{x}}(s)| \leq \theta(|\dot{\bar{x}}(s)|) \leq k_1|\dot{\bar{x}}(s)| + k_2 \quad a.e. .$$

Tomamos t tal que $|\dot{\bar{x}}(t)| \geq \bar{\alpha}$ e chegamos a um absurdo, *quase sempre*. Assim, $|\dot{\bar{x}}(t)| \geq \bar{\alpha}$ num conjunto de medida nula. Daqui resulta que $\bar{x}$ é Lipschitz contínua.

■

Este lema e as condições de optimalidade de §4.1 dão origem a uma qualificação de restrições, simples e verificável, mediante a qual uma regra de multiplicadores não triviais para o problema (CV) pode ser aplicada. Esta regra será definida através de um sistema de equações na forma de Euler-Lagrange, adaptada de forma a incluir a restrição de estado (cf. [8]). Contém como novidade a possibilidade de considerar $\tilde{h}(x_0) = 0$. Note-se que as condições de optimalidade são dadas na forma normalizada, isto é, o multiplicador associado ao custo toma o valor 1.

Proposição 5.2.1 *Seja $\bar{x}$ um minimizante para (CV). Suponhamos que*

$$\tilde{h}_x(\bar{x}(t)) \neq 0 \quad \text{para } t \in \{s \in [a, b] : \tilde{h}(\bar{x}(s)) = 0\}. \quad (5.7)$$

Então, existe uma função $p : [a, b] \rightarrow \mathbb{R}^n$ e uma medida representando um ponto em $C^([a, b])$, tais que*

$$\begin{aligned} \mu &\geq 0, \quad \text{supp } \mu \subset \{t \in [a, b] : \tilde{h}(\bar{x}(t)) = 0\} \\ \left(\dot{p}(t), p(t) + \int_{[a, t)} \tilde{h}_x(\bar{x}(s)) \mu(ds) \right) &= \nabla L(\bar{x}(t), \dot{\bar{x}}(t)) \quad a.e. \ t \in [a, b] \end{aligned}$$

e

$$p(b) + \int_{[a, b]} \tilde{h}_x(\bar{x}(s)) \mu(ds) = 0.$$

A qualificação de restrições simples e verificável que foi referida anteriormente é a condição

$$|\tilde{h}_x(x)| \neq 0, \quad \text{para todo } x \text{ tal que } \tilde{h}_x(x) = 0,$$

que implica 5.7.

Demonstração.

De acordo com o Lema 5.2.1, o minimizante $\bar{x}$ para (CV) é uma função Lipschitz contínua. Assim, é ainda um minimizante quando acrescentamos a restrição:

$$|\dot{x}(t)| \leq K \quad a.e. \ t \in [a, b] \quad (5.8)$$

onde K é qualquer número estritamente maior que $\|\dot{x}\|_{L^\infty}$.

Considere-se uma reformulação de (CV) como um exemplo de (P) , de forma a aplicar os resultados de §4.1. Em virtude da exigência de convexidade nos dados do problema (P) , considere-se a variável de estado como sendo $(z(t), x(t)) \in \mathbb{R} \times \mathbb{R}^n$ e a variável de controlo, $((u_1(t), \dots, u_{n+1}(t)), (\alpha_1(t), \dots, \alpha_{n+1}(t))) \in (\mathbb{R}^n \times \dots \times \mathbb{R}^n) \times \mathbb{R}^{n+1}$. A reformulação de (CV) toma a forma:

$$\begin{aligned}
 & \text{Minimizar } z(b) \\
 & \text{tal que} \\
 & \dot{z}(t) = \sum_i \alpha_i(t) L(x(t), u_i(t)) \\
 (\bar{A}) \quad & \dot{x}(t) = \sum_i \alpha_i(t) u_i(t) \\
 & u_i(t) \in K \bar{B} \\
 & \alpha(t) \in \{(\sigma_1, \dots, \sigma_{n+1}) \in \mathbb{R}^{n+1} : \sigma_i \geq 0 \forall i, \sum \sigma_i = 1\} \\
 & z(a) = 0, \quad x(a) = x_0 \\
 & \tilde{h}(x(t)) \leq 0, \quad \forall t \in [a, b].
 \end{aligned}$$

Para termos um problema exactamente com a forma de (P) podemos considerar uma restrição simbólica sobre o estado final: $\tilde{d}(z, x) \equiv -1$.

Note-se que o processo $\bar{\nu}$ definido por

$$\begin{aligned}
 & \left(\left(\bar{z}(t) \equiv \int_a^t L(\bar{x}(s), \dot{\bar{x}}(s)) ds, \bar{x}(t) \right), ((\bar{u}_1 = \dot{\bar{x}}, \bar{u}_2 = 0, \dots, \bar{u}_{n+1} = 0), \right. \\
 & \left. (\bar{\alpha}_1 = 1, \bar{\alpha}_2 = 0, \dots, \bar{\alpha}_{n+1} = 0) \right)
 \end{aligned}$$

satisfaz as restrições de $(\bar{A})$ e tem custo

$$J(\bar{\nu}) = \int_a^b L(\bar{x}(s), \dot{\bar{x}}(s)) ds. \quad (5.9)$$

O custo $J(\nu)$ de um processo arbitrário,

$$\nu = ((z, x), ((u_1, \dots, u_{n+1}), (\alpha_1, \dots, \alpha_{n+1}))),$$

que satisfaz as restrições de $(\bar{A})$ é dado por:

$$J(\nu) = \int_a^b \sum \alpha_i(t) L(x(t), u_i(t)) dt.$$

As restrições de (CV) são satisfeitas pela componente x de tais processos e da convexidade de $L(x, \cdot)$ resulta

$$J(\nu) \geq \int_a^b L(x(t), \dot{x}(t)) dt. \quad (5.10)$$

De 5.9 e 5.10 conclui-se que $\bar{v}$ é um processo minimizante para $(\bar{A})$. Os dados deste problema satisfazem as hipóteses (H1)-(H5) de §4.1 e é imediato verificar que

(a) Se $\tilde{h}(x_0) = 0$, $(CQ)_3$ é satisfeita.

(b) Se $\tilde{h}(x_0) < 0$, $(CQ)_4$ é satisfeita.

Em cada caso o controlo u' em $(CQ)_3$ e $(CQ)_4$, é definido por:

$$((u'_1 = w, u'_2 = 0, \dots, u'_{n+1} = 0), (\alpha'_1 = 1, \alpha'_2 = 0, \dots, \alpha'_{n+1} = 0))$$

onde w é a função

$$w(t) = \dot{\bar{x}}(t) - \tau \frac{d}{dt} \left\{ \tilde{h}_x(\bar{x}(t)) \cdot (t - a) \right\}.$$

Na expressão, τ é um número positivo e suficientemente pequeno de forma a que $w(t) \in K\bar{B}$ a.e., onde $\bar{B}$ representa o fecho de B .

Aplicando a Prop. 4.1.3 (ou 4.1.4) ao problema $(\bar{A})$, tomando como minimizante o processo $\bar{v}$, obtemos uma função absolutamente contínua p e uma medida μ tal que

$$\begin{aligned} \mu &\geq 0, \quad \text{supp}\{\mu\} \subset \{t \in [a, b] : \tilde{h}(\bar{x}(t)) = 0\}, \\ \max_{|u| \leq K} \left\{ -L(\bar{x}(t), u) + \left(p(t) + \int_{[a, t)} \tilde{h}_x \mu(ds) \right) u \right\} &= \\ = -L(\bar{x}(t), \dot{\bar{x}}(t)) + \left(p(t) + \int_{[a, t)} \tilde{h}_x \mu(ds) \right) \dot{\bar{x}}(t) &\quad a.e. \\ -\dot{p}(t) &= -L_x(\bar{x}(t), \dot{\bar{x}}(t)) \end{aligned}$$

e

$$p(b) + \int_{[a, b]} \tilde{h}_x(\bar{x}(s)) \mu(ds) = 0.$$

Como $|\dot{\bar{x}}| < K$ a.e., a condição de maximização do Hamiltoniano implica

$$p(t) + \int_{[a, t)} \tilde{h}_x \mu(ds) = L_{\dot{x}}(\bar{x}(t), \dot{\bar{x}}(t)) \quad a.e..$$

Estas condições permitem concluir os resultados estabelecidos pela Proposição. ■

Capítulo 6

Restrições de Igualdade no Estado Final

O problema (P) que temos vindo a estudar contém uma restrição no estado final do tipo $d(x(b)) \leq 0$. Substituindo esta desigualdade por uma igualdade estrita, as soluções admissíveis para o novo problema serão um subconjunto do conjunto de soluções admissíveis para o problema (P) e um novo óptimo, de uma maneira geral, surgirá. A igualdade poderia ser substituída, de forma equivalente, por duas restrições de desigualdade, $d(x(b)) \leq 0$ e $-d(x(b)) \leq 0$, e novamente um problema do tipo (P) estaria em questão. Uma análise das condições necessárias neste caso, leva-nos a concluir da pouca eficiência desta decomposição para a determinação da solução. Teríamos mais um multiplicador, variável a determinar, mantendo-se o número de equações que são dados do problema. Surgiriam dois multiplicadores associados à restrição que poderiam ser tomados com valores simétricos, dando origem a um conjunto de multiplicadores triviais.

A determinação de novas condições necessárias através de um desenvolvimento semelhante ao efectuado para (P) deixa de ser válida, como iremos exemplificar, e uma outra abordagem será considerada. Outras técnicas de cálculo e um resultado diferente sobre separação de conjuntos, serão utilizados.

6.1 Condições Necessárias não triviais

Considere-se o problema (P_e) com a seguinte formulação:

$$\begin{aligned}
 & \text{Minimizar } g(x(b)) \\
 & \text{tal que} \\
 & \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \tag{6.1} \\
 (P_e) \quad & x(a) = x_0 \\
 & u(t) \in \Omega \quad \text{a.e. } t \in [a, b] \\
 & h(t, x(t)) \leq 0 \quad \text{para todo } t \in [a, b] \\
 & e(x(b)) = 0 \tag{6.2}
 \end{aligned}$$

A função e satisfaz a hipótese adicional:

(H6): $e : \mathbb{R}^n \rightarrow \mathbb{R}^l$ é uma função continuamente diferenciável.

Todas as outras funções que intervêm no problema assim como o conjunto Ω , satisfazem as hipóteses H1-H5 apresentadas no Capítulo 3.

O Princípio do Máximo para o problema (P_e) toma uma forma semelhante à apresentada para (P) , no Capítulo 3, [8]. Neste caso, um minimizante $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo com multiplicadores $(p, \mu, \sigma, \lambda)$ se:

$\sigma \in \mathbb{R}^l$, $p \in AC^1([a, b]; \mathbb{R}^n)$, $\mu \in C^*([a, b])$, $\lambda \in \mathbb{R}$ são tais que,

$$\mu \geq 0, \lambda \geq 0,$$

$$-\dot{p}(t) = \left(p(t) + \int_{[a, t)} h_x(s, \bar{x}(s)) \mu(ds) \right) \cdot f_x(t, \bar{x}(t), \bar{u}(t)), \quad \text{a.e. } t \in [a, b],$$

para quase todo $t \in [a, b]$, a função

$$u \rightarrow \left(p(t) + \int_{[a, t)} h_x(s, \bar{x}(s)) \mu(ds) \right) \cdot f(t, \bar{x}(t), u)$$

atinge o seu máximo em Ω quando $u = \bar{u}(t)$,

$$\text{supp}\{\mu\} \subset \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}$$

e

$$-\left(p(b) + \int_{[a, b]} h_x(s, \bar{x}(s)) \mu(ds) \right) = \sigma^T e_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b)).$$

O Princípio do Máximo estabelece o seguinte para este problema:

"Mediante as hipóteses H1-H6, o minimizante satisfaz as condições do Princípio do Máximo com multiplicadores $(p, \mu, \sigma, \lambda)$, para os quais

$$|\mu|_{C^*} + |\sigma| + \lambda > 0''.$$

Também para o problema (P_e) , e no caso de $h(a, x_0) = 0$, os multiplicadores

$$p(t) \equiv -h_x(a, x_0), \quad \mu = \delta_{\{a\}}, \quad \sigma = 0 \text{ e } \lambda = 0,$$

são triviais, no sentido de não fornecerem qualquer informação sobre o minimizante. É possível, no entanto, concluir resultados semelhantes aos do Cap. 4, sobre a existência de outros multiplicadores associados à solução ótima. Os teoremas seguintes estabelecem esses resultados.

Teorema 6.1.1 *Seja $(\bar{x}, \bar{u})$ um minimizante para o problema (P_e) e considere-se que as hipóteses H1-H6 são satisfeitas no que se refere aos dados deste problema.*

Suponhamos que,

$(CQ)_1$: $h(a, x_0) = 0$ e existe $u' \in U$ tal que

$$\lim_{\epsilon \rightarrow 0^+} \text{ess sup}_{a \leq s < a + \epsilon} h_x(a, x_0) \cdot [f(s, x_0, u'(s)) - f(s, x_0, \bar{u}(s))] < 0 .$$

Então, $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo com multiplicadores $(p, \mu, \sigma, \lambda)$ para os quais,

$$\int_{(a,b)} \mu(ds) + |\sigma| + \lambda \neq 0 . \quad (6.3)$$

Teorema 6.1.2 *As conclusões do Teorema 6.1.1 permanecem válidas quando a qualificação de restrições $(CQ)_1$ é substituída por*

$$(CQ)_2 : \exists \bar{t} \in (a, b] \text{ tal que } h(t, \bar{x}(t)) < 0 \quad \forall t \in (a, \bar{t}] .$$

A abordagem considerada no Cap. 4 para demonstrar os resultados correspondentes para o problema (P) , deixa de ser válida para o problema (P_e) . A desigualdade 4.30 deixa de ter uma correspondente e a substituição de d por $(e, -e)$, não conduz a nenhum resultado útil para o novo problema em questão. De facto, as condições necessárias, quando efectuamos tal substituição, fornecem de imediato multiplicadores triviais definidos por

$$(p, \mu, \alpha = (\alpha_1, \alpha_2), \lambda) = (0, 0, (\alpha_1, -\alpha_1), 0) \quad \alpha_1 \in \mathbb{R}, \quad \alpha_1 \neq 0.$$

Estaríamos perante um outro tipo de trivialidade, nenhuma informação seria obtida destes multiplicadores.

Antes de iniciarmos a demonstração propriamente dita dos teoremas anteriores apresentamos o desenvolvimento de alguns resultados cuja contribuição é significativa para essa demonstração. Pretende-se desta forma, simplificar formalmente a exposição dos argumentos de prova.

6.2 Demonstração dos Teoremas 6.1.1 e 6.1.2

6.2.1 Resultados Preliminares

Dado um número natural k , defina-se o conjunto convexo,

$$\mathcal{T}_k \equiv \{\theta \in \mathbb{R}^k : \theta_j \geq 0 \quad \forall j = 1, \dots, k, \quad \sum_{j=1}^k \theta_j \leq 1\}.$$

Para $u \in \mathcal{U}$ e $v \in \Omega$, sejam y_u e $\Delta f_v(t)$ tal como foram definidos em 4.5 e 4.4.

Proposição 6.2.1 *Seja $(\bar{x}, \bar{u})$ um processo admissível para o problema (P_e) .*

Tomem-se k funções de controlo arbitrárias $\{u_1(\cdot), \dots, u_k(\cdot)\} \subset \mathcal{U}$.

Então, existe uma vizinhança T' de $\theta = 0$, $T' \subset T_k$, e uma função $w : T' \rightarrow [0, \infty)$, satisfazendo $\lim_{\theta \rightarrow 0} w(\theta) = 0$, com a seguinte propriedade:

Para cada $\theta \in T'$ existe um processo $(x_\theta(\cdot), u_\theta(\cdot))$ para (P_e) tal que, $\forall t \in [a, b]$,

$$\left| h(t, x_\theta(t)) - h(t, \bar{x}(t)) - \sum_{j=1}^k \theta_j h_x(t, \bar{x}(t)) y_{u_j}(t) \right| \leq (t - a) \|\theta\| w(\theta).$$

Demonstração.

Fixem-se k e $u_1(\cdot), \dots, u_k(\cdot)$ como na Proposição. Escolha-se $L \in \mathbb{R}^+$ de forma a que $\bar{x}(t)$ seja, para todo o $t \in [a, b]$, interior à bola aberta de raio L e centro na origem, que representamos por B_L . Defina-se,

$$\begin{aligned} F : [a, b] \times B_L \times T_k &\rightarrow \mathbb{R}^n \\ (t, x, \theta) &\rightsquigarrow f(t, x, \bar{u}(t)) + \sum_{j=1}^k \theta_j [f(t, x, u_j(t)) - f(t, x, \bar{u}(t))] \end{aligned}$$

Em termos de F , a função $\bar{x}(\cdot) \in AC^1([a, b], B_L)$ pode escrever-se:

$$\bar{x}(t) = x_0 + \int_a^t F(\tau, \bar{x}(\tau), 0) d\tau.$$

A função F é continuamente diferenciável em (x, θ) e mensurável em t . As funções f e f_x são limitadas em conjuntos limitados e o conjunto T_k é convexo e limitado. O Teorema 2.3.1 do Capítulo 2 permite deduzir a existência de um conjunto T' , vizinhança de $\theta = 0$, $T' \in T_k$, tal que a equação

$$x(t) = x_0 + \int_a^t F(\tau, x(\tau), \theta) d\tau$$

tem, para todo $\theta \in T'$, uma solução única $x(\cdot, \theta)$ em $C([a, b], B_L)$; a função

$$\begin{aligned} p : T' &\rightarrow C([a, b], B_L) \\ \theta &\rightsquigarrow x(\cdot, \theta) \end{aligned}$$

é continuamente diferenciável, e

$$[p'(0)\theta](t) = \int_a^t \{f_x(\tau, \bar{x}(\tau), \bar{u}(\tau))[p'(0)\theta](\tau) + \sum_{j=1}^k \theta_j [f(\tau, \bar{x}(\tau), u_j(\tau)) - f(\tau, \bar{x}(\tau), \bar{u}(\tau))] \} d\tau.$$

Em particular, podemos concluir

$$\frac{\partial p}{\partial \theta_j}(0) = y_{u_j}(\cdot).$$

A definição de derivada permite escrever,

$$x(\cdot, \theta) = \bar{x}(\cdot) + \sum_{j=1}^k \theta_j y_{u_j}(\cdot) + R(\cdot, \theta) \quad \text{com} \quad \lim_{\theta \rightarrow 0} \frac{\|R(t, \theta)\|_{max}}{\|\theta\|} = 0.$$

Considere-se a aplicação

$$\begin{aligned} q(\theta) : T' &\rightarrow C([a, b]; \mathbb{R}^n) \\ \theta &\leadsto q(\theta)(\cdot) \end{aligned}$$

definida por,

$$\theta \neq 0 \quad q(\theta)(t) = \begin{cases} (t-a)^{-1} \|\theta\|^{-1} [(x(t, \theta) - \bar{x}(t)) - \sum_{j=1}^k \theta_j y_{u_j}(t)] & t \neq a \\ 0 & t = a \end{cases}$$

$$\theta = 0 \quad q(0)(t) = 0 \quad \forall t \in [a, b]$$

O lema seguinte justifica o conjunto de chegada que atribuímos à partida para a aplicação q .

Lema 6.2.1 $\forall \theta \in T' \quad q(\theta)(\cdot) \in C([a, b]; \mathbb{R}^n).$

Demonstração.

Se $\theta = 0$, o resultado é trivial.

Quando $\theta \neq 0$, a continuidade das funções $x(\cdot, \theta)$, $\bar{x}(\cdot)$ e $y_{u_j}(\cdot)$ implica a continuidade de $q(\theta)(\cdot)$ para $t \neq a$.

Resta verificar que $\lim_{t \rightarrow a} q(\theta)(t) = q(\theta)(a) = 0$. Neste caso, a definição de $q(\theta)$

permite escrever, para $t \neq a$,

$$\begin{aligned}
\|q(\theta)(t)\| &\leq \frac{1}{t-a} \frac{1}{\|\theta\|} \left[\int_a^t \|f(\tau, x(\tau, \theta), \bar{u}(\tau)) - f(\tau, \bar{x}(\tau), \bar{u}(\tau))\| d\tau \right. \\
&\quad + \int_a^t \sum_{j=1}^k \theta_j \|f(\tau, x(\tau, \theta), u_j(\tau)) - f(\tau, \bar{x}(\tau), u_j(\tau))\| d\tau \\
&\quad + \int_a^t \sum_{j=1}^k \theta_j \|f(\tau, x(\tau, \theta), \bar{u}(\tau)) - f(\tau, \bar{x}(\tau), \bar{u}(\tau))\| d\tau \\
&\quad \left. + \sum_{j=1}^k \theta_j \int_a^t \|f_x(\tau, \bar{x}(\tau), \bar{u}(\tau)) y_{u_j}(\tau)\| d\tau \right] \\
&\leq \frac{1}{t-a} \frac{1}{\|\theta\|} \left[(1 + 2 \sum_{j=1}^k \theta_j) \int_a^t K \|x(\tau, \theta) - \bar{x}(\tau)\| d\tau + \right. \\
&\quad \left. K \sum_{j=1}^k \theta_j \int_a^t \|y_{u_j}(\tau)\| d\tau \right]
\end{aligned}$$

onde K é uma constante definida por,

$$K = \max \left\{ \sup_{[a,b] \times B_L \times \Omega} \|f\|, \sup_{[a,b] \times B_L \times \Omega} \|f_x\| \right\}.$$

As funções $t \rightarrow \|x(t, \theta) - \bar{x}(t)\|$ e $t \rightarrow y_{u_j}(t)$ são contínuas e iguais a zero para $t = a$, $\forall j = 1, \dots, k$. Sendo um ponto de continuidade, $t = a$ é um ponto de Lebesgue das mesmas funções. $\|q(\theta)(t)\|$ é então majorada por uma função que tende para zero quando $t \rightarrow a$, dando o resultado desejado. ■

Lema 6.2.2 $\lim_{\theta \rightarrow 0} q(\theta)(t) = q(0)(t)$ uniformemente em t .

Demonstração.

De forma a simplificar a notação, represente-se

$$\begin{aligned}
f_x(\tau) &= f_x(\tau, \bar{x}(\tau), \bar{u}(\tau)) \\
\Delta f(\tau, \theta) &= f(\tau, x(\tau, \theta), \bar{u}(\tau)) - f(\tau, \bar{x}(\tau), \bar{u}(\tau))
\end{aligned}$$

Seja $\theta \neq 0$. Para $t \neq a$ tem-se

$$\begin{aligned}
\|q(\theta)(t)\| &\leq \frac{1}{t-a} \frac{1}{\|\theta\|} \left[\int_a^t \|\Delta f(\tau, \theta) - f_x(\tau) \sum_{j=1}^k \theta_j y_{u_j}(\tau)\| d\tau \right. \\
&\quad \left. + \sum_{j=1}^k \theta_j \int_a^t \sup_{[a,b] \times B_L \times \Omega} \|f_x(\tau, x, v)\| \|x(\tau, \theta) - \bar{x}(\tau)\| d\tau \right]
\end{aligned}$$

$$\begin{aligned}
& + \sum_{j=1}^k \theta_j \int_a^t \sup_{[a,b] \times B_L \times \Omega} \|f_x(\tau, x, v)\| \|x(\tau, \theta) - \bar{x}(\tau)\| d\tau \Big] \\
\leq & \frac{1}{t-a} \frac{1}{\|\theta\|} \left[2K \sum_{j=1}^k \theta_j \int_a^t \|x(\tau, \theta) - \bar{x}(\tau)\| d\tau \right. \\
& + K \int_a^t \|x(\tau, \theta) - \bar{x}(\tau) - \sum_{j=1}^k \theta_j y_{u_j}(\tau)\| d\tau \\
& + \int_a^t \int_0^1 \|f_x(\tau, \bar{x}(\tau) + l(x(\tau, \theta) - \bar{x}(\tau)), \bar{u}(\tau)) - f_x(\tau)\| \\
& \quad \left. \left\| \sum_{j=1}^k \theta_j y_{u_j}(\tau) \right\| dl d\tau \right]
\end{aligned}$$

Analisando as expressões do último membro da sucessão de desigualdades verifica-se facilmente que:

$$\frac{1}{\|\theta\|} \|x(\tau, \theta) - \bar{x}(\tau) - \sum_{j=1}^k \theta_j y_{u_j}(\tau)\| \leq \frac{1}{\|\theta\|} \|R(t, \theta)\|_{\max} = w_1(\theta), \quad \forall \tau \in [a, b]$$

$$\text{com } \lim_{\theta \rightarrow 0} w_1(\theta) = 0,$$

$$x(\cdot, \theta) \xrightarrow{\theta} \bar{x}(\cdot) \Rightarrow \|x(\tau, \theta) - \bar{x}(\tau)\| \leq \max_{\tau \in [a, b]} \|x(\tau, \theta) - \bar{x}(\tau)\| = w_2(\theta),$$

$$\text{com } \lim_{\theta \rightarrow 0} w_2(\theta) = 0,$$

$$\|y_{u_j}(\tau)\| \leq 2[\exp\{K(t-a)\} - 1], \quad \forall \tau \in [a, t].$$

A função $(t-a)^{-1}[\exp\{K(t-a)\} - 1]$, definida em $(a, b]$, pode ser continuamente prolongada ao intervalo $[a, b]$ e portanto é limitada, ou seja

$$\exists L_y \geq 0 : \frac{1}{(t-a)} \|y_{u_j}(\tau)\| \leq L_y \quad \forall \tau \in [a, t] \quad \forall t \in (a, b] \quad \forall j = 1, \dots, k.$$

A função $f_x(\cdot, \cdot, \cdot)$ é limitada em $[a, b] \times B_L \times \Omega$ e para qualquer sucessão θ_n convergente para 0 tem-se,

$$\lim_{n \rightarrow \infty} \|f_x(\tau, \bar{x}(\tau) + l(x(\tau, \theta_n) - \bar{x}(\tau)), \bar{u}(\tau)) - f_x(\tau)\| = 0.$$

O Teorema da Convergência Dominante de Lebesgue [25], permite então concluir que

$$\int_a^b \int_0^1 \|f_x(\tau, \bar{x}(\tau) + l(x(\tau, \theta_n) - \bar{x}(\tau)), \bar{u}(\tau)) - f_x(\tau)\| dl d\tau \rightarrow 0.$$

Como a convergência é válida para qualquer sucessão,

$$\int_a^b \int_0^1 \|f_x(\tau, \bar{x}(\tau) + l(x(\tau, \theta) - \bar{x}(\tau)), \bar{u}(\tau)) - f_x(\tau)\| dl d\tau = w_3(\theta),$$

com, $\lim_{\theta \rightarrow 0} w_3(\theta) = 0$.

Finalmente, reunindo as conclusões obtidas, podemos afirmar

$$\|q(\theta)(t)\| \leq K w_1(\theta) + 2K w_2(\theta) + k L_y w_3(\theta) = w_4(\theta) \quad \forall t \in (a, b]$$

com $\lim_{\theta \rightarrow 0} w_4(\theta) = 0$.

Como $q(\theta)(a) = 0$, a última desigualdade é válida para todo o t em $[a, b]$.

Desta desigualdade resulta o lema. ■

Obtidas as propriedades estabelecidas pelos Lemas 6.2.1 e 6.2.2, voltemos à proposição a demonstrar.

Para $t \neq a$ e $\theta \neq 0$,

$$\begin{aligned} & \frac{1}{t-a} \frac{1}{\|\theta\|} \left| h(t, x(t, \theta)) - h(t, \bar{x}(t)) - \sum_{j=1}^k \theta_j h_x(t, \bar{x}(t)) y_{u_j}(t) \right| \\ & \leq L_h \max_{t \in [a, b]} \|q(\theta)(t)\| + \frac{1}{t-a} \frac{1}{\|\theta\|} \|\tilde{h}_x(t, \theta) - h_x(t, \bar{x}(t))\| \|x(t, \theta) - \bar{x}(t)\| \\ & \leq L_h w_4(\theta) + \frac{1}{t-a} \frac{1}{\|\theta\|} \|\tilde{h}_x(t, \theta) - h_x(t, \bar{x}(t))\| \|x(t, \theta) - \bar{x}(t)\| \end{aligned} \quad (6.4)$$

onde,

$$L_h = \max_{t \in [a, b]} \|h_x(t, \bar{x}(t))\|$$

$$\tilde{h}_x(t, \theta) = h_x(t, \bar{x}(t) + \alpha_h(t, \theta)(x(t, \theta) - \bar{x}(t))), \quad \|\alpha_h(t, \theta)\| \leq 1.$$

Uma vez que h_x é contínua, é também uniformemente contínua no conjunto fechado e limitado, $[a, b] \times \bar{B}_L$, onde $\bar{B}_L$ é o fecho de B_L . Quando θ tende para zero, $x(\cdot, \theta)$ converge para a função $x(\cdot, 0) = \bar{x}(\cdot)$. Deste modo,

$$\|\tilde{h}_x(t, \theta) - h_x(t, \bar{x}(t))\| \leq \max_{t \in [a, b]} \|\tilde{h}_x(t, \theta) - h_x(t, \bar{x}(t))\| = w_5(\theta), \quad (6.5)$$

com $\lim_{\theta \rightarrow 0} w_5(\theta) = 0$.

Da definição de $x(t, \theta)$ resulta,

$$\|x(t, \theta) - \bar{x}(t)\| \leq K \int_a^t \|x(\tau, \theta) - \bar{x}(\tau)\| d\tau + 2kK\|\theta\|(t-a).$$

Aplicando o Lema de Gronwall [27],

$$\frac{1}{t-a} \frac{1}{\|\theta\|} \|x(t, \theta) - \bar{x}(t)\| \leq 2kK[1 + K(b-a)\exp\{K(b-a)\}] = L \quad (6.6)$$

As desigualdades 6.5 e 6.6 aplicadas em 6.4 permitem concluir,

$$\left| h(t, x(t, \theta)) - h(t, \bar{x}(t)) - \sum_{j=1}^k \theta_j h_x(t, \bar{x}(t)) y_{u_j}(t) \right| \leq (t - a) \|\theta\| w(\theta)$$

$$\text{com } w(\theta) = L_h w_4(\theta) + L w_5(\theta), \quad \text{e } \lim_{\theta \rightarrow 0} w(\theta) = 0.$$

Definindo $w(0) = 0$, por exemplo, a desigualdade é trivialmente satisfeita para $t = a$ ou $\theta = 0$.

Resta provar que $x(t, \theta)$ pode ser considerada como uma trajectória para o problema (P_e) .

Para *quase todo* t em $[a, b]$, tem-se

$$\begin{aligned} \dot{x}(t, \theta) &= \left(1 - \sum_{j=1}^k \theta_j \right) f(t, x(t, \theta), \bar{u}(t)) + \sum_{j=1}^k \theta_j f(t, x(t, \theta), u_j(t)) \\ &\in f(t, x(t, \theta), \Omega), \end{aligned}$$

uma vez que $f(t, x, \Omega)$ é convexo para qualquer (t, x) . Novamente, argumentos sobre selecções mensuráveis já considerados anteriormente, Cap. 4, permitem deduzir a existência de uma função controlo $u_\theta \in \mathcal{U}$ tal que $(x_\theta(t) = x(t, \theta), u_\theta)$ é um processo. Este processo tem as propriedades desejadas para a conclusão da proposição. ■

6.2.2 Demonstração do Teorema 6.1.1

Vai ser construído um conjunto $\mathcal{W}$ que corresponde na sua estrutura ao conjunto S da demonstração do resultado análogo, Teor. 4.1.1. O conjunto S era então disjunto do conjunto K^- e a aplicação de um teorema de separação fazia surgir os multiplicadores na sua forma quase final. O problema em questão agora, não permite garantir tal separação entre $\mathcal{W}$ e K^- , a Prop. 4.2.4 deixa de ter uma análoga.

Um outro teorema de separação, Teor. 2.2.3, vai resolver a questão. A proposição 2) deste teorema não se verifica em virtude da optimalidade de $(\bar{x}, \bar{u})$ e da Prop. 6.2.2 que vamos apresentar. Da proposição 1) vão surgir os multiplicadores.

Suponhamos então que a qualificação de restrições $(CQ)_1$ é satisfeita.

Tal como foram definidos no Capítulo 4 seja $\mathcal{U}^*$ o conjunto das funções controlo para as quais $t = a$ é um ponto de Lebesgue da função

$$t \rightsquigarrow h_x(a, x_0) \cdot \Delta f_u(t)$$

e seja G_0 o conjunto, fechado, dos pontos t em $[a, b]$ para os quais $h(t, \bar{x}(t)) = 0$.

Para $u \in \mathcal{U}^*$, recorde-se que η_u representa a extensão contínua ao intervalo $[a, b]$, da função $t \rightarrow (t - a)^{-1} h_x(t, \bar{x}(t)) \cdot y_u(t)$, definida em $(a, b]$.

Como $(CQ)_1$ é satisfeita, G_0 é não vazio.

Sem alterar a notação, considere-se a restrição de η_u ao conjunto G_0 e defina-se o conjunto $\mathcal{W}$ da seguinte forma:

$$\mathcal{W} = \{(\eta_u(\cdot), e_x(\bar{x}(b)) \cdot y_u(b), g_x(\bar{x}(b)) \cdot y_u(b)) : u \in \mathcal{U}^*\} \subset C([G_0]) \times \mathbb{R}^l \times \mathbb{R} \quad (6.7)$$

Da convexidade de $f(t, x, \Omega)$ resulta fácilmente que o conjunto $\mathcal{W}$ é convexo. Não é agora possível garantir a separação entre este conjunto e o quadrante negativo do espaço $C([G_0]) \times \mathbb{R}^l \times \mathbb{R}$ como foi feito anteriormente.

A proposição seguinte resulta de uma adaptação do Teorema V.2.2 [27], ao problema em questão. O conjunto $\mathcal{C}$ é definido por

$$\mathcal{C} = \{c(t) \in C([G_0]) : c(t) < 0 \ \forall t \in G_0\}$$

Proposição 6.2.2 *Suponha-se que existem pontos $\xi^i = (\xi_0^i, \xi_1^i, \xi_2^i) \in \mathcal{W}$ e números $\beta^i > 0$, $i = 0, \dots, l$, tais que $\sum_{i=0}^l \beta^i = 1$, $\xi_0^i \in \mathcal{C}$, o conjunto $\{(\xi_1^0, \xi_2^0), \dots, (\xi_1^l, \xi_2^l)\}$ é linearmente independente, $\xi_2^i < 0$ para todo o i e $\sum_{i=0}^l \beta^i \xi_1^i = 0$.*

Seja $\mathcal{T}$, vizinhança de 0 em $\mathcal{T}_{l+1}$.

Então, existe $\tilde{\gamma} > 0$ e uma função $\tilde{\theta} : (0, \tilde{\gamma}] \rightarrow \mathcal{T}$ tais que, para cada $\gamma \in (0, \tilde{\gamma}]$, um processo admissível para P_e , $(x_{\tilde{\theta}(\gamma)}, u_{\tilde{\theta}(\gamma)})$, pode ser determinado de forma a satisfazer

$$g(x_{\tilde{\theta}(\gamma)}(b)) < g(\bar{x}(b)).$$

Demonstração.

Tomem-se $\xi^i = (\xi_0^i, \xi_1^i, \xi_2^i) \in \mathcal{W}$ e números $\beta^i > 0$, $i = 0, \dots, l$, satisfazendo as propriedades referidas na proposição.

Então, existem $\{\tilde{u}_0, \dots, \tilde{u}_l\} \subset \mathcal{U}^*$ tais que

$$\xi^i = (\eta_{\tilde{u}_i}(\cdot), e_x(\bar{x}(b)) \cdot y_{\tilde{u}_i}(b), g_x(\bar{x}(b)) \cdot y_{\tilde{u}_i}(b))$$

$$\eta_{\tilde{u}_i}(t) < 0 \quad \forall t \in G_0, \quad \forall i = 0, \dots, l$$

$$\{(e_x(\bar{x}(b)) \cdot y_{\tilde{u}_0}(b), g_x(\bar{x}(b)) \cdot y_{\tilde{u}_0}(b)), \dots, (e_x(\bar{x}(b)) \cdot y_{\tilde{u}_l}(b), g_x(\bar{x}(b)) \cdot y_{\tilde{u}_l}(b))\}$$

é um conjunto linearmente independente

$$g_x(\bar{x}(b)) \cdot y_{\tilde{u}_i}(b) \leq 0, \quad i = 0, \dots, l \quad \text{e} \quad \sum_{i=0}^l \beta^i e_x(\bar{x}(b)) \cdot y_{\tilde{u}_i}(b) = 0.$$

Para o conjunto de funções controlo $\{\tilde{u}_0, \dots, \tilde{u}_l\}$ tomem-se

$$T' \subset T_{l+1}, \quad w(\theta) \quad \text{e} \quad (x(\cdot, \theta) = x_\theta(\cdot), u_\theta)$$

de acordo com a Prop. 6.2.1 tomando k , que aí era qualquer número natural, com o valor $l + 1$.

Primeiro passo

Vamos começar por verificar que $h(t, x(t, \theta)) \leq 0, \forall t \in [a, b], \forall \theta \in T''$, onde T'' é um conjunto a determinar.

A função $\eta_{\tilde{u}_i}(\cdot)$, tal como foi inicialmente definida em todo o intervalo $[a, b]$, é contínua e portanto uniformemente contínua no mesmo intervalo. Por outro lado, em G_0 é negativa, por hipótese. Para $L_\eta = -\max_{t \in G_0; i=0, \dots, l} (\eta_{\tilde{u}_i}(t)) (> 0)$, existe então $\delta_i > 0$ tal que, $|\eta_{\tilde{u}_i}(t_1) - \eta_{\tilde{u}_i}(t_2)| < \frac{L_\eta}{2}$ desde que $|t_1 - t_2| < \delta_i, t_1, t_2 \in [a, b]$.
Sejam,

$$\delta = \min\{\delta_i : i = 0, \dots, l\} \quad \text{e} \quad V = [a, b] \cap \left(\bigcup_{t \in G_0} B_\delta(t) \right).$$

O conjunto V é não vazio ($G_0 \neq \emptyset$), e

$$t \in V \Rightarrow \exists t_0 \in G_0 : |t - t_0| < \delta \leq \delta_i \Rightarrow \eta_{\tilde{u}_i}(t) < \eta_{\tilde{u}_i}(t_0) + \frac{L_\eta}{2} \leq -\frac{L_\eta}{2} < 0.$$

Defina-se,

$$\bar{V} = [a, b] \setminus V$$

$$-\bar{L}_h = \max_{t \in \bar{V}} h(t, \bar{x}(t)) < 0$$

$$\bar{L}_\eta = \max_{t \in \bar{V}; i=0, \dots, l} |\eta_{\tilde{u}_i}(t)|$$

$$T'' = \left\{ \theta \in T' : w(\theta) < \frac{L_\eta}{4}, \quad \sum_{i=0}^l \theta_i < \frac{2\bar{L}_h}{(b-a)(4\bar{L}_\eta + L_\eta)} \right\}.$$

No caso de $\bar{V}$ ser vazio, a demonstração segue a mesma argumentação, considerando apenas o conjunto V .

Para $\theta \in \mathcal{T}''$, a Prop. 6.2.1 permite afirmar que, se $t \in V$ e $t \neq a$,

$$\begin{aligned} \frac{1}{t-a} [h(t, x(t, \theta)) - h(t, \bar{x}(t))] &\leq \sum_{i=0}^l \theta_i \eta_{\bar{u}_i}(t) + \|\theta\| w(\theta) \\ &\leq \left[-\frac{L_\eta}{2} + w(\theta) \right] \sum_{i=0}^l \theta_i \\ &\leq -\frac{L_\eta}{4} \sum_{i=0}^l \theta_i \leq 0. \end{aligned}$$

De forma análoga, para $\theta \in \mathcal{T}''$, mas agora $t \in \bar{V}$, podemos escrever

$$\begin{aligned} h(t, x(t, \theta)) &\leq h(t, \bar{x}(t)) + \sum_{i=0}^l \theta_i h_x(t, \bar{x}(t)) y_{\bar{u}_i}(t) + (t-a) \|\theta\| w(\theta) \\ &\leq h(t, \bar{x}(t)) + (t-a) \left[\sum_{i=0}^l \theta_i |\eta_{\bar{u}_i}(t)| + \|\theta\| w(\theta) \right] \\ &\leq -\bar{L}_h + (b-a) \sum_{i=0}^l \theta_i \left(\bar{L}_\eta + \frac{L_\eta}{4} \right) \leq -\frac{\bar{L}_h}{2} < 0. \end{aligned}$$

Concluindo, se $t = a$ fica trivial, podemos afirmar que:

$$\theta \in \mathcal{T}'' \Rightarrow h(t, x(t, \theta)) \leq 0 \quad \forall t \in [a, b].$$

Segundo passo

Para $\theta \in \mathcal{T}'$, defina-se $H(\theta) = (e(x(b, \theta)), g(x(b, \theta)))$. H é contínua e tem derivada contínua em $\mathcal{T}'$. Seja $\delta > 0$ tal que $[0, \delta]^{l+1}$ está contido em $\mathcal{T}'$ e denote-se por β o vector $(\beta^0, \dots, \beta^l)$.

Definam-se

$$\begin{aligned} \bar{H} : \mathbb{R}^{l+1} &\rightarrow \mathbb{R}^{l+1} & \psi : \mathcal{T}' &\rightarrow \mathbb{R}^{l+1} \\ \theta &\rightsquigarrow \sum_{i=0}^l (\xi_1^i, \xi_2^i) \theta_i & \theta &\rightsquigarrow H(\theta) - H(0) - H'(0)\theta. \end{aligned}$$

H e $\bar{H}$ (linear), estão relacionadas da seguinte forma:

$$\begin{aligned} H'(0)(\theta) &= \left(e_x(\bar{x}(b)) \cdot \sum_{i=0}^l \theta_i y_{\bar{u}_i}(b), g_x(\bar{x}(b)) \cdot \sum_{i=0}^l \theta_i y_{\bar{u}_i}(b) \right) \\ &= \sum_{i=0}^l (\xi_1^i, \xi_2^i) \theta_i = \bar{H}(\theta). \end{aligned}$$

Uma vez que o conjunto $\{(\xi_1^0, \xi_2^0), \dots, (\xi_1^l, \xi_2^l)\}$ é, por hipótese, linearmente independente, $\bar{H}$ tem uma inversa. Assim, para $\gamma > 0$, fixo, as equações seguintes são equivalentes.

$$H(\theta) - H(0) - \gamma H'(0)\beta = 0 \tag{6.8}$$

$$\theta = \gamma\beta - \bar{H}^{-1}(\psi(\theta)). \quad (6.9)$$

Como H é diferenciável em $\theta = 0$, tem-se $\lim_{\theta \rightarrow 0} \frac{\|\psi(\theta)\|}{\|\theta\|} = 0$.

Assim, para $\beta_{\min} = \min_i \{\beta^i\}$, $\beta_{\max} = \max_i \{\beta^i\}$ e $\theta_{\max} = \max_i \{\theta_i\}$ podemos garantir a existência de $\delta' \in (0, \delta]$ tal que,

$$\|\bar{H}^{-1}(\psi(\theta))\| \leq \frac{1}{3} \frac{\beta_{\min}}{\beta_{\max}} \theta_{\max} \quad \text{se } \theta : 0 \leq \theta_i \leq \delta' \quad i = 0, \dots, l$$

As funções H e ψ são contínuas no conjunto

$$\mathcal{T}'_\gamma = \left\{ \theta \in \mathbb{R}^{l+1} : |\theta_i - \gamma\beta^i| \leq \frac{1}{2} \gamma \beta_{\min} \quad i = 0, \dots, l \right\} \subset [0, \delta]^{l+1}$$

onde

$$\gamma \in (0, \bar{\gamma}], \quad \text{e } \bar{\gamma} = \frac{2}{3} \frac{\delta'}{\beta_{\max}}.$$

O conjunto $\mathcal{T}'_\gamma$ é convexo e limitado. Pelo Teorema 2.2.1, Teorema II.2.3 de Brouwer [27], tem a propriedade de ponto fixo.

A função $\theta \mapsto \gamma\beta - \bar{H}^{-1}(\psi(\theta))$ transforma $\mathcal{T}'_\gamma$ em si mesmo e é contínua neste conjunto.

Assim, a equação 6.9, e portanto também a equação 6.8, admite uma solução $\tilde{\theta}(\gamma) \in \mathcal{T}'_\gamma$, para cada $\gamma \in (0, \bar{\gamma}]$. Para estes valores de γ vamos decompor a equação 6.8 nas suas componentes:

$$\begin{aligned} g(x(b, \tilde{\theta}(\gamma))) &= g(\bar{x}(b)) + \gamma H'_2(0)\beta = g(\bar{x}(b)) + \gamma \sum_{i=0}^l \beta^i \xi_2^i \\ &< g(\bar{x}(b)) \\ e(x(b, \tilde{\theta}(\gamma))) &= e(\bar{x}(b)) + \gamma H'_1(0)\beta = e(\bar{x}(b)) + \gamma \sum_{i=0}^l \beta^i \xi_1^i \\ &= e(\bar{x}(b)) \end{aligned}$$

Observe-se que $\lim_{\gamma \rightarrow 0^+} \sup \{\|\theta\| : \theta \in \mathcal{T}'_\gamma\} = 0$, o que implica

$$\lim_{\gamma \rightarrow 0^+} \tilde{\theta}(\gamma) = 0. \quad (6.10)$$

Para concluir, 6.10 implica a existência de $\tilde{\gamma} < \bar{\gamma}$, tal que para $\gamma \in (0, \tilde{\gamma}]$, $\tilde{\theta}(\gamma) \in \mathcal{T}''$ e $\tilde{\theta}(\gamma) \in \mathcal{T}$, vizinhança da proposição. As trajectórias resultantes destes valores de γ são admissíveis e têm custo inferior ao óptimo.

A proposição está provada. ■

Proposição 6.2.3 Existe $(\nu, \sigma, \lambda) \in C^*([a, b]) \times \mathbb{R}^l \times \mathbb{R}$ tal que

$$\nu \geq 0, \lambda \geq 0$$

$$\text{supp}\{\nu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$$

$$\nu\{(a, b]\} + |\sigma| + \lambda \neq 0,$$

e

$$\int_{[a, b]} \eta_u(t) \nu(dt) + \sigma^T e_x(\bar{x}(b)) \cdot y_u(b) + \lambda g_x(\bar{x}(b)) \cdot y_u(b) \geq 0 \quad (6.11)$$

para todo $u \in \mathcal{U}^*$.

Demonstração.

O conjunto $\mathcal{C} = \{c(t) \in C([G_0]) : c(t) < 0 \forall t\}$ é um conjunto aberto e convexo. A função constante e igual a zero, é aderente ao conjunto. Por outro lado, o conjunto W , definido pela equação 6.7, contém o zero do espaço onde se insere.

A aplicação do Teor. 2.2.3 de separação, permite afirmar que uma das seguintes situações estará sempre presente:

$$\exists \varphi = (\varphi_0, \sigma, \lambda) \in C^*([G_0]) \times \mathbb{R}^l \times [0, \infty] : \quad \varphi \neq 0 ;$$

$$\varphi(z) = \varphi_0(z_0) + \sigma^T z_1 + \lambda z_2 \geq 0 \quad \forall z = (z_0, z_1, z_2) \in W$$

$$\varphi_0(c) \leq 0 \quad \forall c \in \bar{\mathcal{C}}. \quad (6.12)$$

ou

$$\text{Existem } \xi^i = (\xi_0^i, \xi_1^i, \xi_2^i) \in W$$

e números $\beta^i > 0, i = 0, \dots, l$ tais que

$$\sum_{i=0}^l \beta^i = 1, \quad \xi_0^i \in \mathcal{C}$$

$\{(\xi_1^0, \xi_2^0), \dots, (\xi_1^l, \xi_2^l)\}$ é linearmente independente.

$$\xi_2^i < 0 \quad \text{e} \quad \sum_{i=0}^l \beta^i \xi_1^i = 0.$$

Da optimalidade de $(\bar{x}, \bar{u})$ e da Proposição 6.2.2 anterior, podemos concluir que o segundo caso não acontece. Teríamos um processo admissível com custo inferior ao ótimo.

Assim, a primeira alternativa é a válida e daí resulta a existência de uma medida $\tilde{\nu}$ que representa φ_0 , tal que

$$\int_{G_0} \eta_u(t) \tilde{\nu}(dt) + \sigma^T e_x(\bar{x}(b)) \cdot y_u(b) + \lambda g_x(\bar{x}(b)) \cdot y_u(b) \geq 0 \quad \forall u \in \mathcal{U}^*. \quad (6.13)$$

Extendendo $\tilde{\nu}$ como uma medida de Borel regular ν aos subconjuntos de Borel de $[a, b]$ (tal como foi feito no Cap. 4): $\nu(B) = \tilde{\nu}(B \cap G_0)$, vem

$$\text{supp}\{\nu\} \subset G_0 \text{ e } \int_{G_0} \eta_u(t) \tilde{\nu}(dt) = \int_{[a,b]} \eta_u(t) \nu(dt)$$

para todo $u \in \mathcal{U}^*$. Estes factos combinados com a equação 6.12, implicam $\nu \geq 0$ e a equação 6.11.

Resta provar que $\nu\{(a, b]\} + |\sigma| + \lambda \neq 0$. Como $\nu\{[a, b]\} + |\sigma| + \lambda \neq 0$, basta verificar que $(\nu, \sigma, \lambda) \neq (\rho\delta_{\{a\}}, 0, 0)$ para qualquer $\rho > 0$.

Por redução ao absurdo, suponhamos que $(\nu, \sigma, \lambda) = (\rho\delta_{\{a\}}, 0, 0)$ para algum $\rho > 0$. Considere-se a função controlo $\tilde{u}$ da Prop. 4.2.3, que naturalmente continua a ser válida. Esta função é um elemento de $\mathcal{U}^*$, verifica $\eta_{\tilde{u}}(a) < 0$ e

$$\int_{[a,b]} \eta_{\tilde{u}}(t) \nu(dt) + \sigma^T e_x(\bar{x}(b)) \cdot y_{\tilde{u}}(b) + \lambda g_x(\bar{x}(b)) \cdot y_{\tilde{u}}(b) = \rho \eta_{\tilde{u}}(a) < 0$$

o que contradiz 6.13.

A proposição está provada. ■

As Proposições 4.2.2, 4.2.7 e 4.2.8 do Cap. 4, onde α e d são agora respectivamente σ e e , dão os resultados necessários para a conclusão da demonstração do Teorema 6.1.1.

6.2.3 Demonstração do Teorema 6.1.2

A demonstração do Teorema 6.1.2 pode ser feita usando exactamente os mesmos argumentos da demonstração do correspondente teorema, Teor. 4.1.2, para o problema (P) no Cap. 4. O multiplicador α , associada à restrição $d(x(b)) \leq 0$, deverá ser substituído pelo novo multiplicador σ . As componentes de σ podem tomar valores positivos ou negativos. É esta essencialmente a diferença que deve ser tida em consideração.

Assim, a equação 4.50 na demonstração do Teor. 4.1.2 deverá agora tomar a forma:

$$\int_{[a,b]} \mu_i \, ds + \sum_{j \in I_0} (|\sigma^j|)_i + \lambda_i = 1 \quad (6.14)$$

onde $(|a|^j)_i$ é o valor absoluto da j -ésima componente de a_i .

Todo o desenvolvimento posterior é válido com esta substituição natural.

O teorema ficará desta forma demonstrado.

Capítulo 7

Problema de Controlo Óptimo Não Regular

O problema (P) apresentado no capítulo 3, foi estudado no capítulo 4 mediante as hipóteses H1-H5 que exigiam, entre outras condições de regularidade, diferenciabilidade da função custo e da função restrição sobre o estado final.

Problemas de Controlo Óptimo em que a função custo a otimizar não é uma função diferenciável, surgem frequentemente associados a problemas da vida real. Comprimentos e distâncias são vulgarmente elementos a otimizar e as funções que os exprimem são não diferenciáveis. Neste caso, o problema necessita de ser abordado com técnicas de cálculo não diferenciável.

Tem sido desenvolvida nos últimos anos toda uma teoria envolvendo funções não diferenciáveis. Resultados elementares do cálculo clássico foram generalizados e a sua aplicação a vários tipos de problemas de controlo óptimo tem sido objecto de estudo de diversos autores [6], [7], [8], [11], [12], [24], [26]. Condições necessárias sob a forma de um princípio do máximo foram também já derivadas para esses problemas. O tipo de trivialidade que temos vindo a considerar neste trabalho pode no entanto ainda lhes estar associada.

Neste capítulo vai ser desenvolvido um Princípio do Máximo não trivial, para um problema de controlo óptimo em que a função custo e a função que define restrições sobre o estado final são não diferenciáveis. Mais precisamente, vão ser generalizados os Teoremas 4.1.1 e 4.1.2, para este caso. A demonstração dos resultados generalizados segue uma sequência semelhante à dos anteriores, embora os elementos envolvidos sejam diferentes e o seu tratamento um pouco mais elaborado. Termos como gradiente generalizado e derivadas generalizadas aparecem como substitutos das anteriores derivadas das funções.

7.1 Princípio do Máximo

Formalizando, consideremos o problema (P) , tal como foi apresentado no Cap. 3, mas em que g e d satisfazem agora propriedades mais fracas. Considere-se então a hipótese (H2) substituída pela hipótese $\bar{H}2$ seguinte:

$\bar{H}2$: g e d são funções Lipschitz contínuas.

Nestas condições, um minimizante $(\bar{x}, \bar{u})$ para (P) satisfaz as condições do Princípio do Máximo, se existirem $p \in AC^1([a, b]; \mathbb{R}^n)$, $\mu \in C^*([a, b])$, $\alpha \in \mathbb{R}^k$, $\lambda \in \mathbb{R}$, $\gamma_i \in \mathbb{R}^n$,

$i = 1, \dots, k$, com $\gamma_i \in \partial d^i(\bar{x}(b))$, e $\xi \in \partial g(\bar{x}(b))$ tais que:

$$\mu \geq 0, \alpha \geq 0, \lambda \geq 0,$$

$$-\dot{p}(t) = \left(p(t) + \int_{[a,t)} h_x(s, \bar{x}(s)) \mu(ds) \right) \cdot f_x(t, \bar{x}(t), \bar{u}(t)), \quad \text{a.e. } t \in [a, b],$$

para quase todo $t \in [a, b]$, a função

$$u \rightarrow \left(p(t) + \int_{[a,t)} h_x(s, \bar{x}(s)) \mu(ds) \right) \cdot f(t, \bar{x}(t), u)$$

atinge o seu máximo em Ω quando $u = \bar{u}(t)$,

$$\text{supp}\{\mu\} \subset \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}, \quad \alpha_i = 0 \text{ se } d^i(\bar{x}(b)) < 0$$

e

$$-\left(p(b) + \int_{[a,b]} h_x(s, \bar{x}(s)) \mu(ds) \right) = \sum_{i=1}^k \alpha_i \gamma_i + \lambda \xi. \quad (7.1)$$

Por analogia vamos designar $(p, \mu, \alpha, \gamma_i, \lambda, \xi)$ por multiplicadores.

Note-se que estas condições correspondem às do Cap. 3 com a equação 3.6 substituída por 7.1.

De forma análoga, o Princípio do Máximo estabelece o seguinte:

"Mediante as hipóteses H1, $\bar{H}2$, H3-H5, um minimizante para (P) satisfaz as condições do Princípio do Máximo com $p, \mu, \alpha, \lambda, \gamma_i, i = 1, \dots, k$ e ξ , para os quais

$$|\mu|_{C^*} + |\alpha| + \lambda > 0''. \quad (7.2)$$

É simples verificar que quando $h(a, x_0) = 0$, o minimizante $(\bar{x}, \bar{u})$ continua a satisfazer as condições do Princípio do Máximo com:

$$p(t) \equiv -h_x(a, x_0), \quad \mu = \delta_{\{a\}}, \quad \alpha = 0 \quad \text{e} \quad \lambda = 0. \quad (7.3)$$

Estamos perante o mesmo tipo de trivialidade das condições e da mesma forma é possível garantir a existência de outros multiplicadores, mediante a satisfação de uma certa qualificação das restrições.

Os Teoremas 4.1.1 e 4.1.2 continuam a ser válidos na íntegra, com as novas condições de regularidade sobre as funções g e d .

Os Teoremas 4.1.3 e 4.1.4 permanecem válidos com uma natural substituição das equação 4.6 e 4.7. Derivadas direccionais generalizadas deverão substituir as tradicionais.

Teorema 7.1.1 *Considere-se o problema (P) e as hipóteses H1, $\bar{H}2$, H3-H5. Seja $(\bar{x}, \bar{u})$ um minimizante.*

Suponhamos que

$(CQF)_3$: $h(a, x_0) = 0$ e existe $u' \in \mathcal{U}$ tal que

$$\lim_{\epsilon \rightarrow 0^+} \text{ess sup}_{a \leq s < a + \epsilon} h_x(a, x_0) \cdot [f(s, x_0, u'(s)) - f(s, x_0, \bar{u}(s))] < 0.$$

$$h_x(t, \bar{x}(t)) \cdot y_{u'}(t) < 0 \quad \forall t \in (a, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}$$

$$d_0^i(\bar{x}(b); y_{u'}(b)) < 0 \quad \text{se } d^i(\bar{x}(b)) = 0. \quad (7.4)$$

Então $(\bar{x}, \bar{u})$ satisfaz as condições do Princípio do Máximo com multiplicadores $(p, \mu, \alpha, \gamma_i, \lambda, \xi)$ para os quais $\lambda \neq 0$.

Teorema 7.1.2 *Os resultados do Teorema 7.1.1 permanecem válidos quando se substitui a qualificação de restrições $(CQF)_3$ por:*

$(CQF)_4$: Existe $\bar{t} \in (a, b]$ e $u' \in \mathcal{U}$ tal que

$$h(t, \bar{x}(t)) < 0 \quad \forall t \in (a, \bar{t}],$$

$$h_x(t, \bar{x}(t)) \cdot y_{u'}(t) < 0 \quad \forall t \in (a, b] \cap \{t \in [a, b] : h(t, \bar{x}(t)) = 0\}$$

e

$$d_0^i(\bar{x}(b); y_{u'}(b)) < 0 \quad \text{se } d^i(\bar{x}(b)) = 0. \quad (7.5)$$

7.2 Demonstração dos Resultados Principais

De acordo com os dados do problema, a Proposição 4.2.1 continua a ser válida. Este resultado envolve apenas processos para o problema (P) e portanto o que está em questão é a dinâmica do sistema associado. Como a função f que define essa dinâmica mantém as condições de regularidade segundo as quais o resultado foi derivado, a proposição continua a ser válida.

Se substituirmos na Prop. 4.2.2, a expressão $[\alpha^T d_x(\bar{x}(b)) + \lambda g_x(\bar{x}(b))] \cdot y_u(b)$ por $\sum_{i=1}^k \alpha_i \gamma_i + \lambda \xi$, onde γ_i e ξ são vectores em $\mathbb{R}^n$, a proposição continua a ser verdadeira. O elemento substituído funciona como uma constante cujas características não interferem directamente na demonstração dessa proposição.

A qualificação de restrições $(CQ)_1$ (na Proposição 4.1.1) continua a implicar a Proposição 4.2.3 com os novos dados do problema, desde que substituamos a equação 4.21 pela sua correspondente:

$$d_0^i(\bar{x}(b); y_u(t)) < 0 \quad \text{se } d^i(\bar{x}(b)) = 0. \quad (7.6)$$

Uma vez que a derivada direccional generalizada é contínua em relação ao segundo argumento, os sucessivos passos da demonstração continuam a ser válidos.

A grande diferença entre a demonstração dos Teoremas 4.1.1, 4.1.2, caso não regular, e Teoremas 7.1.1, 7.1.2 e a demonstração dos resultados correspondentes do Cap. 4 assenta na seguinte proposição que corresponde ainda, no seu conteúdo, à Prop. 4.2.4.

Proposição 7.2.1 *Existe $\gamma_i \in \partial d^i(\bar{x}(b))$, $i = 1, \dots, k$, e $\xi \in \partial g(\bar{x}(b))$ tal que, para todo o elemento $u \in \mathcal{U}^*$, verifica-se:*

$$\left(\max_{t \in G_0} \{\eta_u(t)\} \right) \vee \left(\bigvee_{i \in I_0} \gamma_i \cdot y_u(b) \right) \vee (\xi \cdot y_u(b)) \geq 0. \quad (7.7)$$

Demonstração.

Começamos por fixar $u \in \mathcal{U}^*$, $\epsilon' > 0$ como na Prop. 4.2.1, $\epsilon \in (0, \epsilon']$ e $\delta > 0$.

Por argumentos análogos aos apresentados na demonstração da Prop. 4.2.4, tendo em conta que g e d são contínuas de Lipschitz, é possível concluir:

$$\left(\sup_{t \in G_\delta \setminus \{a\}} \left\{ (t-a)^{-1} [h(t, \bar{x}(t) + \epsilon y_u(t)) - h(t, \bar{x}(t))] \right\} \right) \vee \left(\bigvee_{i \in I_0} [d^i(\bar{x}(b) + \epsilon y_u(b)) - d^i(\bar{x}(b))] \right) \vee (g(\bar{x}(b) + \epsilon y_u(b)) - g(\bar{x}(b))) \geq -\epsilon \tilde{w}(\epsilon). \quad (7.8)$$

onde $\tilde{w}(\epsilon)$ é uma função positiva, crescente e satisfaz $\tilde{w}(\epsilon) \downarrow 0$ quando $\epsilon \downarrow 0$.

Desenvolvendo separadamente as expressões anteriores, obtém-se:

$$\frac{1}{t-a} [h(t, \bar{x}(t) + \epsilon y_u(t)) - h(t, \bar{x}(t))] = \epsilon \left[\frac{1}{t-a} h_x(t, \bar{x}(t)) \cdot y_u(t) + \bar{R}(t, \epsilon) \right]$$

com,

$$\lim_{\epsilon \rightarrow 0} \sup_{t \in [a, b]} |\bar{R}(t, \epsilon)| = 0.$$

Daqui vem a desigualdade:

$$\sup_{t \in G_\delta \setminus \{a\}} \left\{ \frac{1}{t-a} [h(t, \bar{x}(t) + \epsilon y_u(t)) - h(t, \bar{x}(t))] \right\} \leq$$

$$\epsilon \sup_{t \in G_\delta \setminus \{a\}} \left\{ \frac{1}{t-a} h_x(t, \bar{x}(t)) \cdot y_u(t) \right\} + \epsilon \sup_{t \in [a, b]} |\bar{R}(t, \epsilon)|.$$

•

$$g(\bar{x}(b) + \epsilon y_u(b)) - g(\bar{x}(b)) = \epsilon g_0(\bar{x}(b); y_u(b)) + \epsilon G(\epsilon)$$

onde,

$$G(\epsilon) = \frac{g(\bar{x}(b) + \epsilon y_u(b)) - g(\bar{x}(b))}{\epsilon} - g_0(\bar{x}(b); y_u(b)).$$

Da definição de derivada direccional resulta, $\limsup_{\epsilon \downarrow 0} G(\epsilon) \leq 0$.

De forma análoga podemos escrever,

•

$$d^i(\bar{x}(b) + \epsilon y_u(b)) - d^i(\bar{x}(b)) = \epsilon d_0^i(\bar{x}(b); y_u(b)) + \epsilon D_i(\epsilon)$$

com $\limsup_{\epsilon \downarrow 0} D_i(\epsilon) \leq 0$.

Destes desenvolvimentos e da equação 7.8 vem

$$\left(\sup_{t \in G_\delta \setminus \{a\}} \left\{ \frac{1}{t-a} h_x(t, \bar{x}(t)) \cdot y_u(t) \right\} \right) \vee \left(\bigvee_{i \in I_0} d_0^i(\bar{x}(b); y_u(b)) \right) \vee (g_0(\bar{x}(b); y_u(b)))$$

$$\geq -\tilde{w}(\epsilon) - \max \left\{ \sup_{t \in [a, b]} |\bar{R}(t, \epsilon)|, G(\epsilon), \max_{i \in I_0} \{D_i(\epsilon)\} \right\}$$

Aplicando $\liminf_{\epsilon \downarrow 0}$ a ambos os termos da desigualdade, obtém-se

$$\left(\sup_{t \in G_\delta \setminus \{a\}} \left\{ \frac{1}{t-a} h_x(t, \bar{x}(t)) \cdot y_u(t) \right\} \right) \vee \left(\bigvee_{i \in I_0} d_0^i(\bar{x}(b); y_u(b)) \right)$$

$$\vee (g_0(\bar{x}(b); y_u(b))) \geq 0.$$

Ou, utilizando as equações 4.33 e 4.34 que naturalmente continuam a ser válidas, e fazendo $\delta \rightarrow 0$,

$$\max_{t \in G_0} \eta_u(t) \vee \left(\bigvee_{i \in I_0} d_0^i(\bar{x}(b); y_u(b)) \right) \vee (g_0(\bar{x}(b); y_u(b))) \geq 0. \quad (7.9)$$

Considere-se agora o conjunto:

$$Y = \{ y_u(b) : u \in \mathcal{U}^*, \dot{y}_u(t) = f_x(t, \bar{x}(t), \bar{u}(t)) y_u(t) + \Delta f_u(t), y_u(0) = 0, \\ \max_{t \in G_0} \eta_u(t) < 0 \text{ e } d_0^i(\bar{x}(b); y_u(b)) < 0, \forall i \in I_0 \}$$

Se $Y = \emptyset$ a equação 7.11 é imediatamente satisfeita e neste caso podemos tomar para ξ um elemento qualquer do conjunto $\partial g(\bar{x}(b))$. Suponhamos então que $Y \neq \emptyset$.

O conjunto $f(t, x, \Omega)$ é convexo e a derivada direccional generalizada é uma função homogénea positiva e semiaditiva em relação ao segundo argumento. Destas propriedades, das propriedades do supremo e por argumentos envolvendo selecções mensuráveis, já utilizados nos capítulos anteriores, é possível concluir que o conjunto Y é convexo. Por outro lado, embora 0 não pertença ao conjunto, qualquer ponto da forma ρv com $v \in Y$ e $0 < \rho \leq 1$, está em Y .

Considere-se a função,

$$\begin{aligned} G : \bar{Y} &\rightarrow \mathbb{R} \\ v &\rightarrow g_0(\bar{x}(b); v) \end{aligned}$$

onde $\bar{Y}$ representa a aderência de Y .

Como função de v , $g_0(\bar{x}(b); v)$ é convexa, Lipschitz contínua e, em virtude da eq. 7.9, é não negativa em Y . A função G é portanto não negativa e como ainda satisfaz $G(0) = 0$, podemos afirmar que $v = 0$ é um ponto onde a função toma o seu valor mínimo. Aplicando o resultado de optimização restricta para funções de Lipschitz [8], Prop. 2.5.4 do Cap. 2, podemos concluir

$$0 \in \partial g_0(\bar{x}(b); 0) + N_{\bar{Y}}(0). \quad (7.10)$$

Mas $\partial g_0(\bar{x}(b); 0) = \partial g(\bar{x}(b))$ e como Y é um conjunto convexo, a sua aderência também o é. A equação 7.10 equivale deste modo a afirmar a existência de um elemento ξ do conjunto $\partial g(\bar{x}(b))$ tal que $\langle \xi, y_u(b) \rangle \geq 0, \forall y_u(b) \in Y$.

Com este elemento ξ , a equação 7.9 pode ser transformada em:

$$\max_{t \in G_0} \eta_u(t) \vee \left(\bigvee_{i \in I_0} d_0^i(\bar{x}(b); y_u(b)) \right) \vee (\xi \cdot y_u(b)) \geq 0. \quad (7.11)$$

Repetindo o processo sucessivamente, para os conjuntos:

$$Y_j = \{y_u(b) : u \in \mathcal{U}^* \mid \dot{y}_u(t) = f_x(t, \bar{x}(t), \bar{u}(t)) y_u(t) + \Delta f_u(t), y_u(0) = 0, \\ \max_{t \in G_0} \eta_u(t) < 0, \quad d_0^i(\bar{x}(b); y_u(b)) < 0 \quad i \in I_0 \setminus \{j\} \text{ e } \xi \cdot y_u(b) < 0\}$$

com $j \in I_0$, podemos obter a Proposição 7.2.1. ■

A demonstração dos Teoremas 4.1.1, 4.1.2, caso não regular, 7.1.1 e 7.1.2 poderá ser concluída utilizando uma sequência de argumentos, análoga à utilizada no Cap. 4:

1) Construção do conjunto S :

$$S := \begin{cases} \{(\eta_u, \eta_u^0, \eta_u^1) \in C(G_0) \times \mathbb{R}^{k_0} \times \mathbb{R} : u \in \mathcal{U}^*\} & \text{se } k_0 \geq 1 \\ \{(\eta_u, \eta_u^1) \in C(G_0) \times \mathbb{R} : u \in \mathcal{U}^*\} & \text{se } k_0 = 0 \end{cases}$$

onde, agora

$$\eta_u^0 := \{\gamma_i \cdot y_u(b) : i \in I_0\} \quad \eta_u^1 := \xi \cdot y_u(b).$$

S é um conjunto convexo e, denotando por K^- o cone aberto negativo em $C(G_0) \times \mathbb{R}^{k_0} \times \mathbb{R}$ (ou $C(G_0) \times \mathbb{R}$ se $I_0 = \emptyset$), tem-se $S \cap K^- = \emptyset$.

2) Estabelecimento de proposições correspondentes a 4.2.6, 4.2.7 e 4.2.8, cujas demonstrações são semelhantes:

Proposição 7.2.2 *Existe $(\nu, \alpha, \lambda) \in C^*([a, b]) \times \mathbb{R}^k \times \mathbb{R}$ tal que*

$$\nu \geq 0, \quad \alpha \geq 0, \quad \lambda \geq 0$$

$$\text{supp}\{\nu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$$

$$\nu\{(a, b]\} + |\alpha| + \lambda \neq 0,$$

$$\alpha_i = 0 \quad \text{se } d^i(\bar{x}(b)) < 0$$

e

$$\int_{[a, b]} \eta_u(t) \nu(dt) + \sum_{i=1}^k \alpha_i \gamma_i \cdot y_u(b) + \lambda \xi \cdot y_u(b) \geq 0 \quad (7.12)$$

para todo $u \in \mathcal{U}^*$.

onde γ_i e ξ são os vectores cuja existência é assegurada pela Prop. 7.2.1.

Proposição 7.2.3 *Seja $q : (a, b] \rightarrow \mathbb{R}^n$ a função mensurável definida por*

$$q(t) = \int_{[t,b]} h_x(s, \bar{x}(s)) \Phi(s, t) (s - a)^{-1} \nu(ds) + \left[\sum_{i=1}^k \alpha_i \gamma_i + \lambda \xi \right] \Phi(b, t).$$

Então,

$$q(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) = \min_{u \in \Omega} q(t) \cdot f(t, \bar{x}(t), u) \quad a.e. \ t \in (a, b]. \quad (7.13)$$

Proposição 7.2.4

$$\lim_{t \downarrow a} \int_{[t,b]} (s - a)^{-1} \nu(ds) < \infty$$

Com estes resultados e repetindo passo a passo as demonstrações efectuadas no Capítulo 4, obtém-se a demonstração dos Teoremas 4.1.1 e 4.1.2 com as novas hipóteses e a demonstração dos Teoremas 7.1.1 e 7.1.2.

Capítulo 8

Conclusões

Principais Contribuições

Esta tese teve como tema o desenvolvimento de condições necessárias não triviais, sob a forma do Princípio do Máximo, para certos problemas de Controlo Ótimo com restrições de estado.

O Princípio do Máximo estabelece um conjunto de condições necessárias que, quando o problema tem o estado inicial fixo e sobre a fronteira do conjunto de restrições, são trivialmente satisfeitas por um conjunto de multiplicadores não nulos, simultaneamente.

Neste caso, o Princípio do Máximo não fornece qualquer informação para a determinação da solução ótima do problema.

A contribuição essencial desta tese esteve no desenvolvimento de uma nova versão do Princípio do Máximo que garante a existência de outros multiplicadores, para além desses triviais, associados à solução ótima. Para isso foi necessário impôr uma hipótese adicional aos dados do problema, designada por qualificação das restrições.

Garantir que condições necessárias para problemas de optimização sejam úteis, no sentido de fornecer multiplicadores não triviais, tem sido objecto de estudo de vários autores. Este estudo está no entanto dirigido no sentido de estabelecer alguma qualificação das restrições, de forma a garantir que tais multiplicadores não possam simplesmente existir. *Todos* os multiplicadores são não triviais na presença dessas hipóteses adicionais sobre os dados do problema.

Os resultados que desenvolvemos, aceitando os triviais como um dado adquirido, garantem a existência de outros multiplicadores para além destes.

Foram considerados três tipos de problemas de controlo ótimo com restrições de estado.

O primeiro, servindo de base aos restantes, envolveu, além das restrições de estado, restrições sobre o estado final definidas através de uma desigualdade. Quando se substitui esta desigualdade por uma igualdade, a abordagem efectuada deixa de ser válida. Daqui resultou um segundo problema para o qual outras técnicas de abordagem foram utilizadas de forma a obter resultados análogos.

Para não tornar a notação muito pesada, o segundo problema considerado não foi tomado com restrições sobre o estado final dos dois tipos simultaneamente, igualdade e desigualdade. No entanto, uma observação cuidada dos passos efectuados e resultados intermédios, permite deduzir a possibilidade de generalizar os resultados

obtidos para este caso. Com essa generalização o segundo problema passará a conter o primeiro. É no entanto interessante, sob o ponto de vista técnico, descrever a abordagem que inicialmente foi tomada, quando apenas restrições de desigualdade sobre o estado final estavam presentes.

Em ambos os problemas foram assumidas hipóteses de diferenciabilidade sobre a função custo. O terceiro problema considerado, estruturalmente semelhante ao primeiro, dispunha de hipóteses mais fracas. A função custo e a função que definia restrições sobre o estado final deixam de ser continuamente diferenciáveis e passam a ser apenas Lipschitz contínuas. O tratamento deste problema passou a fazer-se num ambiente de cálculo não diferenciável, utilizando-se técnicas próprias deste contexto.

Pode parecer à partida que a qualificação das restrições, assumida no Princípio do Máximo não trivial, é de difícil verificação, uma vez que é definida à custa da solução óptima. Há no entanto problemas para os quais essa condição é facilmente verificável a partir dos dados do problema. Foi apresentado um problema tipo onde tal acontece, problema esse que se insere no clássico Cálculo das Variações.

Perspectivas de Desenvolvimento

Há toda uma perspectiva de desenvolvimento, no sentido de generalizar os resultados obtidos a problemas satisfazendo hipóteses mais fracas. Em particular, a hipótese considerada do conjunto de velocidades ser convexo, parece-nos ser importante enfraquecer, de forma a que os resultados obtidos sejam aplicáveis a um maior número de problemas. Problemas de controlo óptimo com funções custo envolvendo um termo integral podem ser facilmente reformulados como problemas com função custo dependendo apenas do estado final, através da introdução de uma variável de estado adicional. A dinâmica neste caso é também alterada e de uma maneira geral, mesmo que inicialmente o conjunto de velocidades seja convexo, o problema reformulado não mantém esta característica.

Restrições mais genéricas, por exemplo o estado inicial pertencer a um dado conjunto não singular, podem ser ainda objecto de investigação futura no que se refere ao tipo de trivialidade estudada neste trabalho.

Bibliografia

- [1] L. Ambrosio, O. Ascenzi, and G. Buttazzo. Lipschitz regularity for minimizers of integral functionals with highly discontinuous integrands. *J. Math. Anal. Appl.*, To appear.
- [2] A. V. Arutyunov and N. T. Tynyanskiy. The maximum principle in a problem with phase constraints. *Soviet J. Comput. Systems Sci.*, 23(1):28–35, 1985.
- [3] R. J. Aumann. Measurable utility and the measurable choice theorem. *La Decision, Actes Coll. Int. du CNRS, Aix-en-Provence*, 1967.
- [4] A. V. Balakrishnan. *Control Theory and the Calculus of Variations*. Academic Press, New York, 1969.
- [5] L. Cesari. *Optimization-Theory and Applications*. Springer-Verlag, New York, 1983.
- [6] F. H. Clarke. Generalized gradients and applications. *American Mathematical Society*, 205:1078–1091, 1975.
- [7] F. H. Clarke. The maximum principle under minimal hypotheses. *SIAM J. Control and Optim.*, 14:1078–1091, 1976.
- [8] F. H. Clarke. *Methods of Dynamic and Nonsmooth Optimization*. John Wiley, New York, 1983.
- [9] F. H. Clarke, P. D. Lowen, and R. B. Vinter. Differential inclusions with free time. *Ann. Inst. Henri Poincare*, 5:573–593, 1988.
- [10] F. H. Clarke and R. B. Vinter. Regularity properties of solutions to the basic problem in the calculus of variations. *Trans. Amer. Math. Soc.*, 289:73–98, 1985.

- [11] F. H. Clarke and R. B. Vinter. Applications of optimal multiprocesses. *SIAM J. Control and Opt.*, 27:1048–1071, 1989.
- [12] F. H. Clarke and R. B. Vinter. Optimal multiprocesses. *SIAM J. Control and Opt.*, 27:1072–1091, 1989.
- [13] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North-Holland, Amsterdam, 1976.
- [14] R. V. Gamkrelidze. On some extremal problems in the theory of differential equations with applications to the theory of optimal control. *SIAM J. Control and Opt.*, 3:106–128, 1965.
- [15] M. R. Hestenes. *Calculus of Variations and Optimal Control Theory*. John Wiley, New York, 1966.
- [16] A. D. Ioffe and V. M. Tihomirov. *Theory of Extremal Problems*. North-Holland, New York, 1974.
- [17] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley, New York, 1969.
- [18] D. Q. Mayne and H. Michalska. Receding horizon control of nonlinear systems. *IEEE Trans. A. C.*, 35:814–824, 1990.
- [19] L. W. Neustadt. A general theory of extremals. *J. of Computer and System Sci.*, 3:57–92, 1969.
- [20] L. S. Pontryagin and al. *The Mathematical Theory of Optimal Processes*. John Wiley, New York, 1962.
- [21] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. Vol1: Functional Analysis*. Academic Press, New York, 1980.
- [22] R. T. Rockafellar. State constraints in convex problems of bolza. *SIAM J. Control and Optim.*, 10:691–715, 1972.
- [23] R. T. Rockafellar. Integral functionals, normal integrands and measurable selections. *Nonlinear Operators and the Calculus of Variations*, L. Waelbroeck, Ed., *Lecture Notes in Mathematics*, 543:157–207, 1976.

- [24] J. D. L. Rowland and R. B. Vinter. Dynamic optimization problems with free time and active state constraints. *SIAM J. Control and Optim.*, To appear.
- [25] W. Rudin. *Real and Complex Analysis*. McGraw-Hill International Editions, 3rd edition, 1986.
- [26] R. B. Vinter and G. Pappas. A maximum principle for non-smooth optimal control problems with state constraints. *J. Math. An. and Applic.*, 89:212–232, 1982.
- [27] J. Warga. *Optimal Control of Differential and Functional Equations*. Academic Press, New York, 1972.