

# **Cálculo Variacional e Controle Ótimo**



Publicações Matemáticas

**Cálculo Variacional e  
Controle Ótimo**

Antonio Leitão  
UFSC

**impa**



23<sup>º</sup> Colóquio Brasileiro de Matemática

Copyright © 2001 by Antonio Leitão  
Direitos reservados, 2001 pela Associação Instituto  
Nacional de Matemática Pura e Aplicada - IMPA  
Estrada Dona Castorina, 110  
22460-320 Rio de Janeiro, RJ

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger

### **23<sup>o</sup> Colóquio Brasileiro de Matemática**

- Aspectos de Modelagem Matemática em Dinâmica dos Fluidos – André Nachbin
- **Cálculo Variacional e Controle Ótimo – Antonio Leitão**
- Computational Methods in the Local Theory of Curves – Abramo Hefez e Marcelo Escudeiro Hernandes
- Contornos, Conjuntos Convexos e Autômatos Celulares – André Toom
- Inteiros Quadráticos e o Grupo de Classes – Antonio J. Engler e Paulo Brumatti
- Introduction to Toric Varieties – Jean-Paul Brasselet
- Introdução aos Espaços de Escala (EDPs em Processamento de Imagens) – Ralph Teixeira
- Índices de Campos Holomorfos e Aplicações – Marcio G. Soares e Rogério S. Mol
- Notes on Morse Theory – Daniel V. Tausk, Francesco Mercuri e Paolo Piccione
- One Dimensional Dynamics: the Mathematical Tools – Edson de Faria e Welington de Melo
- Partial Regularity of Solutions of the 3-D Incompressible Navier-Stokes Equations – Hermano Frid e Mikhail Perepelitsa
- Perfect Simulation of Spatial Processes – Nancy L. Garcia
- Riemannian and Submanifold Geometry – Carlos Olmos
- Tópicos em Combinatória Contemporânea – Carlos Gustavo Moreira e Yoshiharu Kohayakawa
- Uma Introdução Sucinta a Algoritmos de Aproximação – Cristina G. Fernandes et al

#### **Distribuição:**

IMPA  
Estrada Dona Castorina, 110  
22460-320 Rio de Janeiro, RJ  
e-mail: [dic@impa.br](mailto:dic@impa.br)  
<http://www.impa.br>

ISBN: 85-244-0171-0

# Conteúdo

<b>1</b>	<b>Introdução ao Cálculo de Variações</b>	<b>3</b>
1.1	Problemas Variacionais e Convexidade . . . . .	3
1.2	Lemas de du Bois-Reymond e Lagrange . . . . .	9
1.3	Equação de Euler-Lagrange . . . . .	10
1.4	Três Problemas Variacionais Clássicos . . . . .	14
1.5	Extremais Diferenciáveis por Partes . . . . .	17
1.6	Problemas Vetoriais . . . . .	24
1.7	Exercícios . . . . .	25
<b>2</b>	<b>Problemas Variacionais e Controle Ótimo</b>	<b>27</b>
2.1	Problemas de Controle Ótimo: Apresentação . . . . .	27
2.2	Problemas Variacionais com Restrição . . . . .	28
2.3	Extremais Singulares e Trajetórias Ótimas . . . . .	38
2.4	Convexidade I: condições suficientes . . . . .	41
2.5	Convexidade II: condições necessárias . . . . .	45
2.6	Exercícios . . . . .	46
<b>3</b>	<b>Princípio do Máximo de Pontryagin</b>	<b>49</b>
3.1	Problemas com Horizonte Finito . . . . .	49
3.2	Problemas com Horizonte Infinito . . . . .	54
3.3	Problemas Impulsivos . . . . .	57
3.4	Aplicações do Princípio do Máximo . . . . .	59
3.5	Exercícios . . . . .	79
<b>4</b>	<b>Demonstração do Princípio do Máximo</b>	<b>81</b>
4.1	Otimização Infinita . . . . .	81
4.2	Um Problema Auxiliar . . . . .	83
4.3	Condições Necessárias de Otimalidade . . . . .	90
4.4	Exercícios . . . . .	93
	<b>Apêndice</b>	<b>95</b>

<b>A</b>	<b>Otimização Infinita</b>	<b>95</b>
A.1	Um Problema Abstrato de Otimização . . . . .	95
A.2	Linearização do Problema de Otimização . . . . .	96
A.3	Condições Necessárias para o Problema Abstrato . . . . .	102
	<b>Bibliografia</b>	<b>105</b>
	<b>Índice</b>	<b>110</b>

# Lista de Figuras

1.1	Candidato a geodésica na esfera . . . . .	15
1.2	Ciclóide: solução do problema da Braquistócrona . . . . .	17
1.3	Suavização de funções $\hat{C}^1[a, b]$ . . . . .	20
2.1	Condição de contorno transversal e funções admissíveis . . . . .	30
2.2	Extremal Singular: Caso I . . . . .	40
2.3	Extremal Singular: Caso II . . . . .	41
3.1	Algoritmo do método de shooting para resolver um sistema Hamiltoniano. . . . .	53
3.2	Trajatórias ótimas para os controles $u^* \equiv 1$ e $u^* \equiv -1$ . . . . .	61
3.3	Trajatórias correspondentes a controles constantes . . . . .	61
3.4	Trajatórias ótimas correspondentes aos controles $u^* \equiv 1$ e $u^* \equiv -1$ . . . . .	62
3.5	Condições iniciais $(h, v)$ que são levadas pelo controle $u^* \equiv 1$ à condição final $(0, 0, m(T))^*$ com $m(T) \geq M$ . . . . .	64
3.6	Trajatória ótima para o problema da alunissagem . . . . .	65
3.7	Curvas $\Gamma_1$ e $\Sigma^*$ . . . . .	74
3.8	Curva $\Sigma_0$ . . . . .	76
3.9	Cenário para o problema da pescaria ótima . . . . .	77
3.10	Regiões principais do plano de fase . . . . .	78
A.1	Cones tangenciais $C(x)$ e $T(C, x)$ . . . . .	97





# Capítulo 1

## Introdução ao Cálculo de Variações

Este capítulo é dedicado ao estudo de uma família particular de problemas de otimização, denominados na literatura moderna por *problemas variacionais*. A análise formal de tais problemas leva o nome de *cálculo variacional*. Tais problemas de otimização se caracterizam por estarem definidos em espaços (de funções) de dimensão infinita e pelo fato da função objetivo ser descrita por um operador integral.

De especial interesse é o estudo de condições necessárias e/ou suficientes para caracterização das soluções (pontos de máximo/mínimo) destes problemas de otimização. Simplificadamente, tais condições são obtidas a partir da extensão aos problemas variacionais de resultados da análise em várias variáveis, que permitem caracterizar os candidatos a solução dos problemas de otimização em espaços de dimensão finita.

Neste capítulo estudamos problemas variacionais sem restrições (somente com condições de contorno). Os problemas variacionais sujeitos a restrições são analisados no Capítulo 2, onde o teorema de multiplicadores de Lagrange desempenha papel fundamental.

No Parágrafo 1.1 são introduzidos os problemas variacionais e apresentados os conceitos de *extremos locais* e de *variações de Gâteaux*. Sob hipóteses adicionais de convexidade, é obtido um resultado preliminar sobre condições suficientes para otimalidade. No Parágrafo 1.2 é apresentado um resultado auxiliar (lema de du Bois-Raymond), essencial para obtenção das condições necessárias. No Parágrafo 1.3 analisamos condições necessárias para otimalidade, que são apresentadas na forma da *equação de Euler-Lagrange* e de sua *primeira integral*. No Parágrafo 1.5 consideramos problemas com trajetórias contínuas por partes, enquanto que no Parágrafo 1.6 são estudados problemas vetoriais.

### 1.1 Problemas Variacionais e Convexidade

Um problema clássico do cálculo de variações é a minimização de funcionais do tipo

$$I(y) := \int_a^b L(t, y(t), y'(t)) dt, \quad (1.1)$$

onde o intervalo  $[a, b]$  é fixo e a aplicação  $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  é conhecida. A minimização do funcional definido em (1.1) é realizada entre as funções  $y : [a, b] \rightarrow \mathbb{R}$  continuamente

diferenciáveis em  $[a, b]$ .<sup>1</sup>

O problema de minimizar o funcional em (1.1) é, via de regra, sujeito a diferentes tipos de restrição como:

- Impor condições a  $y$  nas extremidades do intervalo, i.e.  $y(a) = y_a$  e/ou  $y(b) = y_b$ ;
- Exigir que  $g(t, y(t), y'(t)) \equiv 0$  para  $t \in [a, b]$ , onde  $g : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ ;
- Exigir que  $\int_a^b g(t, y(t), y'(t)) dt = c$ , onde  $g : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  e  $c \in \mathbb{R}$ ;

O primeiro tipo de restrição é denominado **condição de contorno** e pode ser exigido em ambos ou em apenas um dos extremos do intervalo  $[a, b]$ . O segundo tipo de restrição é denominada **restrição Lagrangeana**, devido a sua semelhança com as restrições presentes nos problemas da mecânica Lagrangeana. O terceiro tipo é denominada **restrição isoperimétrica** (ou integral), uma vez que os primeiros problemas de interesse relacionados à esta restrição apresentavam a exigência dos candidatos  $y$  terem todos o mesmo *comprimento* (fato relacionado com a escolha  $g(t, y, y') = |y'|$ ).

O intervalo  $[a, b]$  não precisa ser necessariamente fixo. Este é o caso quando uma das condições de contorno é descrita pela curva de nível de uma função  $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Temos assim uma restrição do tipo:

- $\sigma(T, y(T)) = 0$ , com  $T > a$ ,

que é denominada **condição (de contorno) transversal**. Note que a função objetivo é da forma  $I(y, T) := \int_a^T L(t, y(t), y'(t)) dt$  e  $T > a$ , o extremo direito do intervalo, também precisa ser determinado.

As restrições acima podem aparecer de forma combinada, de modo que um mesmo problema pode estar sujeito a restrições de diferentes tipos, ou a várias restrições do mesmo tipo. Analisamos neste capítulo somente os problemas variacionais com condições de contorno. Problemas sujeitos a restrições lagrangeanas e isoperimétricas, assim como problemas com condições de contorno transversais são discutidos no Parágrafo 2.2.

Considere a seguinte família de problemas variacionais

$$\begin{cases} \text{Minimizar } I(y) := \int_a^b L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ y \in Y_{ad} := \{y \in C^1[a, b] \mid y(a) = y_a, y(b) = y_b\}. \end{cases} \quad (1.2)$$

O conjunto  $Y_{ad}$  é denominado conjunto das **funções admissíveis** (ou viáveis). Da análise real, sabemos que é preciso distinguir entre mínimos locais e globais do funcional  $I$ :

**Definição 1.1.1**  $\bar{y} \in Y_{ad}$  é denominado **mínimo global** de  $I : Y_{ad} \rightarrow \mathbb{R}$  quando a desigualdade

$$I(y) \geq I(\bar{y}),$$

<sup>1</sup>Adotamos no texto a seguinte notação:

$$\begin{aligned} C[a, b] &:= \{y : [a, b] \rightarrow \mathbb{R} \mid y \text{ contínua}\}, \\ C^1[a, b] &:= \{y : [a, b] \rightarrow \mathbb{R} \mid y \text{ diferenciável em } [a, b], y' \in C[a, b]\}. \end{aligned}$$

é satisfeita para todo  $y \in Y_{ad}$ . □ □ □

Definimos na seqüência os mínimos locais de  $I$ . A importância destes mínimos vem do fato de que condições apenas necessárias, em geral, os englobam juntamente com os mínimos globais. Antes de apresentar a definição, relembremos alguns fatos importantes sobre a topologia dos espaços  $C[a, b]$  e  $C^1[a, b]$ .

**Observação 1.1.2** O espaço vetorial  $C[a, b]$  é completo (Banach) com a norma do supremo (ou de Tschebyscheff)

$$\|f\|_\infty := \sup_{t \in [a, b]} |f(t)|, \quad f \in C[a, b].$$

Logo,  $d(f, g) := \|f - g\|_\infty$  define uma distância (métrica) em  $C[a, b]$ . De forma análoga é possível construir uma métrica para  $C^1[a, b]$  com a distância  $\|f - g\|_{1, \infty} := \|f - g\|_\infty + \|f' - g'\|_\infty, f, g \in C^1[a, b]$ . □ □ □

**Definição 1.1.3**  $\bar{y} \in Y_{ad}$  é denominado **mínimo local fraco** de  $I$  quando

$$\exists \delta > 0 \text{ tq } \forall y \in Y_{ad} \text{ com } \|y - \bar{y}\|_{1, \infty} < \delta \text{ temos } I(y) \geq I(\bar{y}).$$

$\bar{y} \in Y_{ad}$  é denominado **mínimo local forte** de  $I$  quando

$$\exists \delta > 0 \text{ tq } \forall y \in Y_{ad} \text{ com } \|y - \bar{y}\|_\infty < \delta \text{ temos } I(y) \geq I(\bar{y}).$$

□ □ □

De especial interesse para os problemas variacionais são os mínimos locais fracos. Note que uma vizinhança forte de  $y \in Y_{ad}$  possui mais elementos do que a vizinhança fraca correspondente. Portanto, a exigência de  $y$  ser mínimo local forte é mais restritiva.

A seguir, é apresentada uma definição que estende o conceito de *derivada direcional* de aplicações  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e é de importância fundamental para a caracterização de soluções do problema variacional (1.2).

**Definição 1.1.4** Seja  $Y$  um espaço vetorial e  $I : Y \rightarrow \mathbb{R}$ . Dados  $y, v \in Y$  definimos a **variação de Gâteaux** de  $I$  em  $y$  na direção  $v$  através de

$$\delta I(y; v) := \lim_{\varepsilon \rightarrow 0} \frac{I(y + \varepsilon v) - I(y)}{\varepsilon},$$

quando o limite existe. □ □ □

Note que, caso a derivada total em relação à variável  $\varepsilon$  da função real  $\mathbb{R} \ni \varepsilon \mapsto I(y + \varepsilon v) \in \mathbb{R}$  esteja bem definida em  $\varepsilon = 0$ , podemos escrever

$$\delta I(y; v) = \left. \frac{d}{d\varepsilon} I(y + \varepsilon v) \right|_{\varepsilon=0}.$$

Da Definição 1.1.4 segue ainda a linearidade da variação de Gâteaux em relação ao funcional  $I$ , i.e., fixados  $y, v \in Y$  temos

$$\delta(\alpha I_1 + \beta I_2)(y; v) = \alpha \delta I_1(y; v) + \beta \delta I_2(y; v), \quad \alpha, \beta \in \mathbb{R}. \tag{1.3}$$

(Pressuposto, obviamente, que as variações envolvidas existam.) Outra propriedade interessante se verifica em relação à direção de derivação:

$$\delta I(y; \alpha v) = \alpha \delta I(y; v), \quad \alpha \in \mathbb{R}. \quad (1.4)$$

(Novamente supondo a existência da variação.)

**Exemplo 1.1.5** Seja  $Y = \mathbb{R}^n$  e  $f \in C^1(Y; \mathbb{R})$ . Dados  $y, v \in Y$  temos

$$\delta f(y; v) = \lim_{\varepsilon \rightarrow 0} \frac{f(y + \varepsilon v) - f(y)}{\varepsilon} = \langle \nabla f(y), v \rangle,$$

que é a derivada parcial de  $f$  em  $y$  na direção do vetor  $v$ . □ □ □

**Exemplo 1.1.6** Seja  $Y = C^1[a, b]$  e  $I : Y \ni y \mapsto \int_a^b L(t, y(t), y'(t)) dt \in \mathbb{R}$ , onde  $L$  é uma aplicação  $C^1([a, b] \times \mathbb{R}^2; \mathbb{R})$ . Dados  $y, v \in Y$  temos

$$\begin{aligned} \delta I(y; v) &= \left. \frac{d}{d\varepsilon} I(y + \varepsilon v) \right|_{\varepsilon=0} \\ &= \int_a^b \left[ \left. \frac{d}{d\varepsilon} L(t, y + \varepsilon v, y' + \varepsilon v') \right]_{\varepsilon=0} dt \\ &= \int_a^b [L_y(t, y(t), y'(t))v(t) + L_{y'}(t, y(t), y'(t))v'(t)] dt. \end{aligned}$$

□ □ □

Os primeiros resultados sobre condições suficientes de otimalidade para o problema (1.2) surgem quando fazemos hipóteses de convexidade sobre a função objetivo. Isto justifica a introdução do seguinte conceito.

**Definição 1.1.7** Dado um espaço vetorial  $Y$  e um funcional  $I : D \subset Y \rightarrow \mathbb{R}$ , dizemos que  $I$  é **convexo em  $D$**  quando para todo par de elementos  $y, v \in Y$  temos

$$I(\alpha y + (1 - \alpha)v) \leq \alpha I(y) + (1 - \alpha)I(v), \quad \alpha \in [0, 1].$$

$I$  é denominado **estritamente convexo em  $D$** , quando a desigualdade na expressão acima for estrita. □ □ □

Caso a variação de Gâteaux  $\delta I(y; v)$  do funcional  $I$  esteja definida para todos elementos  $y, y + v \in D$ , é possível fornecer uma definição equivalente de convexidade:

**Lema 1.1.8** *Seja  $Y$  um espaço vetorial e  $I : D \subset Y \rightarrow \mathbb{R}$  tal que para todo par de elementos  $y, v$  de  $Y$  satisfazendo  $y, y + v \in D$ , a variação de Gâteaux  $\delta I(y; v)$  existe. São equivalentes as afirmações:*

- a)  $I$  é convexo em  $D$ ;
- b) Para todo par de elementos  $y, v \in Y$  com  $y, y + v \in D$ , temos

$$I(y + v) - I(y) \geq \delta I(y; v).$$

Temos ainda que  $I$  é estritamente convexo em  $D$ , quando a igualdade na desigualdade do item  $b$ ) ocorre se e somente se  $v = 0$ .

**Demonstração:** É bastante demonstrarmos a equivalência entre  $a$ ) e  $b$ ).

$(a) \implies (b)$  Note que para  $0 < \varepsilon < 1$  podemos escrever  $y + \varepsilon v$  na forma da combinação linear convexa

$$y + \varepsilon v = (1 - \alpha)y + \alpha(y + v),$$

com  $\alpha = \varepsilon \in [0, 1]$ . Da definição de convexidade segue que

$$I(y + \varepsilon v) \leq (1 - \varepsilon)I(y) + \varepsilon I(y + v),$$

de onde obtemos

$$\frac{1}{\varepsilon} [I(y + \varepsilon v) - I(y)] \leq I(y + v) - I(y).$$

Tomando agora o limite quando  $\varepsilon \rightarrow 0$ , obtemos a desigualdade em  $b$ ).

$(b) \implies (a)$  Dados  $y, v \in Y$  e  $\alpha \in [0, 1]$ , defina  $w := \alpha y + (1 - \alpha)v$ . Da hipótese  $b$ ) segue

$$\delta I(w; h_1) \leq I(v) - I(w), \quad \delta I(w; h_2) \leq I(y) - I(w),$$

para  $h_1 = \alpha(v - y)$ ,  $h_2 = (1 - \alpha)(y - v)$ . De (1.4) segue agora que

$$\frac{1}{\alpha} [I(v) - I(w)] \geq \delta I(w; (v - y)) \geq \frac{1}{(1 - \alpha)} [I(w) - I(y)],$$

de onde obtemos  $\alpha I(y) + (1 - \alpha)I(v) \geq I(w)$  e o teorema fica provado.  $\blacksquare$

Apresentamos a seguir uma condição suficiente de otimalidade para um problema abstrato em que a função objetivo é convexa.

**Lema 1.1.9** Dado um espaço vetorial  $Y$  e um funcional convexo  $I : D \subset Y \rightarrow \mathbb{R}$ , então cada  $\bar{y}$  que satisfaz

$$\delta I(\bar{y}; v) = 0, \quad \forall \bar{y} + v \in D,$$

minimiza  $I$  em  $D$ . Se  $I$  é estritamente convexo, então  $\bar{y}$  é único.

**Demonstração:** Dado  $y \in D$  defina  $v := y - \bar{y} \in Y$ . Logo

$$I(y) - I(\bar{y}) = I(\bar{y} + v) - I(\bar{y}) \geq \delta I(\bar{y}; v) = 0.$$

A unicidade de  $\bar{y}$  é consequência imediata da definição de convexidade estrita.  $\blacksquare$

Estamos agora em condições de formular um resultado que fornece condições suficientes de otimalidade para o problema (1.2).

**Teorema 1.1.10** Seja  $Y = C^1[a, b]$ ,  $Y_{ad} = \{y \in Y; y(a) = y_a, y(b) = y_b\}$ ,  $I : Y_{ad} \ni y \mapsto \int_a^b L(t, y, y') dt \in \mathbb{R}$ , onde  $L \in C^1([a, b] \times \mathbb{R}^2; \mathbb{R})$  satisfaz

$$L(t, y + v, z + w) - L(t, y, z) \geq L_y(t, y, z)v + L_z(t, y, z)w, \quad (1.5)$$

para todo  $(t, y, z)$ ,  $(t, y + z, z + w) \in [a, b] \times \mathbb{R}^2$ . São verdadeiras as afirmações:

- a)  $I$  é convexo em  $Y_{ad}$ ;  
 b) Se  $L$  é tal que a igualdade em (1.5) ocorre sse  $vw = 0$ , então  $I$  é estritamente convexo em  $Y_{ad}$ ;  
 c) Cada  $\bar{y} \in Y_{ad}$  que satisfaz a equação diferencial

$$\frac{d}{dt}L_{y'}(t, \bar{y}, \bar{y}') = L_y(t, \bar{y}, \bar{y}'), \quad t \in [a, b] \quad (1.6)$$

é um mínimo global de  $I$  em  $Y_{ad}$ .

- d) Se a hipótese do item (b) é verificada, o elemento  $\bar{y} \in Y_{ad}$  que satisfaz a equação diferencial do item c) é o único mínimo global de  $I$  em  $Y_{ad}$ .

**Demonstração:** Dados  $y, y + v \in Y_{ad}$ , a desigualdade (1.5) implica em

$$L(t, y + v, y' + v') - L(t, y, y') \geq L_y(t, y, y')v + L_{y'}(t, y, y')v'. \quad (1.7)$$

Integrando obtemos

$$\int_a^b [L(t, y + v, y' + v') - L(t, y, y')] dt \geq \int_a^b [L_y(t, y, y')v + L_{y'}(t, y, y')v'] dt,$$

i.e.  $I(y + v) - I(y) \geq \delta I(y; v)$ , provando o item (a). Se a hipótese em (b) é verdadeira, então dados  $y, y + v \in Y_{ad}$  a igualdade em (1.7) ocorre sse  $v(t)v'(t) \equiv 0$  em  $[a, b]$ . Note que  $v'(t)v'(t) = 1/2(v^2(t))'$ , logo  $v(t)v'(t) \equiv 0$  sse  $v^2(t)$  é constante. Note ainda que  $v^2(a) = 0$ .<sup>2</sup> Portanto  $I(y + v) - I(y) = \delta L(y; v)$  sse  $v \equiv 0$  e o item (b) fica provado.

Os itens (c) e (d) são os mais importantes do teorema. Suponha que  $\bar{y} \in Y_{ad}$  é solução da equação diferencial (1.6). Logo, para  $v \in Y$  com  $\bar{y} + v \in Y_{ad}$  temos

$$\begin{aligned} \delta I(\bar{y}; v) &= \int_a^b [L_y(t, \bar{y}, \bar{y}')v + L_{y'}(t, \bar{y}, \bar{y}')v'] dt \\ &= \int_a^b \left[ \frac{d}{dt}L_{y'}(t, \bar{y}, \bar{y}')v \right] dt \\ &= [L_{y'}(t, \bar{y}(t), \bar{y}'(t))v(t)]_a^b = 0. \end{aligned}$$

Como  $I$  é convexo (veja o item (a)), o item (c) segue do Lema 1.1.9. Se a hipótese do item (b) é verificada, então  $I$  é estritamente convexo e a unicidade de  $\bar{y}$  segue da segunda parte do Lema 1.1.9. ■

A equação diferencial (1.6) é denominada **equação de Euler-Lagrange** e desempenha um papel ímpar no cálculo variacional. O Teorema 1.1.10, além de fornecer condições suficientes para garantir a convexidade do funcional  $I$ , caracteriza (no caso  $I$  convexo) a equação de Euler-Lagrange como condição suficiente de otimalidade para o problema 1.2. Voltaremos a analisar esta equação diferencial de segunda ordem no Parágrafo 1.3, onde são investigadas condições necessárias para otimalidade.

<sup>2</sup>De fato, como  $y$  e  $y + v$  pertencem a  $Y_{ad}$ , então  $v(a) = v(b) = 0$ .

## 1.2 Lemas de du Bois-Reymond e Lagrange

Neste parágrafo estudamos um lema, cuja formalização impulsionou fortemente o desenvolvimento do cálculo variacional. Trata-se do **lema de du Bois-Reymond**, que foi objeto de investigação por matemáticos como Euler e Lagrange, tendo sido apresentado por P. du Bois-Reymond no ano de 1879. Foi entretanto K. Weierstraß, que em seus seminários (1875–1882) apresentou pela primeira vez uma demonstração clara e completa. A demonstração deste lema utiliza um resultado auxiliar que enunciamos a seguir.

**Lema 1.2.1** *Seja  $h \in C[a, b]$  tal que*

$$\int_a^b h(t) v'(t) dt = 0,$$

*para todo  $v \in C_0^1[a, b] := \{w \in C^1[a, b] \mid w(a) = w(b) = 0\}$ . Então a função  $h$  é constante no intervalo  $[a, b]$ .*

**Demonstração:** Dado  $c \in \mathbb{R}$ , defina a função  $v(t) := \int_a^t (h(s) - c) ds$ ,  $t \in [a, b]$ . Por construção  $v \in C^1[a, b]$  e  $v(a) = 0$ . A escolha particular  $\bar{c} := 1/(b-a) \int_a^b h(s) ds$  gera uma função  $\bar{v}$  que satisfaz

$$0 \leq \int_a^b (h(t) - \bar{c})^2 dt = \int_a^b (h(t) - \bar{c}) \bar{v}'(t) dt = \int_a^b h(t) \bar{v}'(t) dt - [\bar{c}\bar{v}]_a^b = 0,$$

provando assim que  $h(t) = \bar{c}$ , para  $t \in [a, b]$ . ■

**Lema 1.2.2 (du Bois-Reymond)** *Seja  $f \in C[a, b]$  tal que para todo  $v \in C^1[a, b]$  com  $v(a) = v(b) = 0$  tenhamos*

$$\int_a^b f(t) v(t) dt = 0.$$

*Então  $f \equiv 0$  em  $[a, b]$ .*

**Demonstração:** Definindo  $g(t) := \int_a^t f(s) ds$ , segue da hipótese (integrando por partes) que

$$-\int_a^b g(s) v'(t) dt = 0, \quad \forall v \in C_0^1[a, b].$$

O Lema 1.2.1 implica que  $g(t) = c$ ,  $t \in [a, b]$ . O lema de du Bois-Reymond segue agora da identidade  $g' = f$ . ■

As funções  $v$  utilizadas nos Lemas 1.2.1 e 1.2.2 são denominadas **funções teste**, devido ao papel por elas desempenhado. O espaço  $C_0^1[a, b]$  é também denominado **espaço de funções teste**. Discutimos a seguir uma generalização do lema de du Bois-Reymond, formulada por Lagrange.

**Lema 1.2.3 (Lagrange)** *Seja  $f \in C[a, b]$  tal que para todo  $v \in C_0^k[a, b] := \{v \in C^k[a, b] \mid v^{(j)}(a) = v^{(j)}(b) = 0, j = 0, 1, \dots, k\}$  tenhamos*

$$\int_a^b f(t) v(t) dt = 0.$$

Então  $f \equiv 0$  em  $[a, b]$ .

**Demonstração:** Suponha que  $f(t_0) > 0$  para algum  $t_0 \in (a, b)$ . Como  $f$  é contínua, existe  $[c, d] \subset (a, b)$  tal que

$$t_0 \in (c, d) \quad \text{e} \quad f(t) \geq \frac{1}{2} f(t_0) > 0, \quad t \in (c, d).$$

Definindo a função

$$v(t) := \begin{cases} [(t-c)(d-t)]^{k+1}, & t \in [c, d] \\ 0, & t \in [a, b] \setminus [c, d] \end{cases}$$

observamos que  $v$  é não negativa e  $v \in C_0^k[a, b]$ . Por construção temos que  $\int_a^b f v dt = \int_c^d f v dt > 0$ , o que contradiz a hipótese. Logo  $f(t) \leq 0, t \in (a, b)$ . A continuidade de  $f$  implica que  $f(t) \leq 0$  em  $[a, b]$ . Analogamente, provamos que  $f(t) \geq 0$  em  $[a, b]$ . ■

### 1.3 Equação de Euler–Lagrange

Suponha que  $L : [a, b] \times \mathbb{R}^2 \rightarrow \mathbb{R}$  é duas vezes continuamente diferenciável e que o funcional  $I$  possui um mínimo local fraco  $\bar{y} \in Y_{ad} := \{y \in C^1[a, b] \mid y(a) = y_a, y(b) = y_b\}$  satisfazendo  $\bar{y} \in C^2[a, b]$ . Dados  $\eta \in C_0^1[a, b]$  e  $\varepsilon_0 > 0$  suficientemente pequeno, definimos a família de funções admissíveis

$$y(\cdot; \varepsilon) := \bar{y}(\cdot) + \varepsilon \eta(\cdot), \quad \varepsilon \in (-\varepsilon_0, \varepsilon_0).$$

Por construção temos  $\|y(\cdot; \varepsilon) - \bar{y}\|_{1, \infty} = |\varepsilon| \|\eta\|_{1, \infty}$ . Note ainda que, pelo fato de  $\bar{y}$  ser mínimo local fraco de  $I$  em  $Y_{ad}$ , temos

$$J(\varepsilon) := I(y(\cdot; \varepsilon)) \geq I(\bar{y}) = J(0), \quad \forall \varepsilon \in (-\varepsilon_0, \varepsilon_0).$$

Portanto,  $\bar{\varepsilon} = 0$  é mínimo local da função real  $J$ , definida pela composição de  $I$  com a família  $y(\cdot; \varepsilon)$ . Da análise real, sabemos que uma condição necessária para que isto aconteça (supondo  $J$  diferenciável) é que

$$\left. \frac{d}{d\varepsilon} J(\varepsilon) \right|_{\varepsilon=0} = 0.$$

Da Definição 1.1.4, segue que

$$\delta I(\bar{y}; \eta) = \left. \frac{d}{d\varepsilon} I(\bar{y} + \varepsilon \eta) \right|_{\varepsilon=0} = \left. \frac{d}{d\varepsilon} J(\varepsilon) \right|_{\varepsilon=0} = 0 \quad (1.8)$$



é uma condição necessária para que  $\bar{y}$  seja mínimo local de  $I$  em  $Y_{ad}$ . Uma vez que  $L$  é, em particular, continuamente diferenciável, temos (veja Exemplo 1.1.6)

$$\delta I(\bar{y}; \eta) = \int_a^b [L_y(t, \bar{y}(t), \bar{y}'(t))\eta(t) + L_{y'}(t, \bar{y}(t), \bar{y}'(t))\eta'(t)] dt. \quad (1.9)$$

Note que a função

$$\lambda : [a, b] \ni t \mapsto L_{y'}(t, \bar{y}(t), \bar{y}'(t)) \in \mathbb{R}$$

é continuamente diferenciável em  $[a, b]$ , devido as hipóteses feitas em  $L$  e  $\bar{y}$ . Como  $\eta$  satisfaz as condições de contorno  $\eta(a) = \eta(b) = 0$ , obtemos integrando (1.9) por partes

$$\delta I(\bar{y}; \eta) = \int_a^b \left[ L_y(t, \bar{y}(t), \bar{y}'(t)) - \frac{d}{dt} \lambda(t) \right] \eta(t) dt. \quad (1.10)$$

Note que (1.10) é válida para todo  $\eta \in C_0^1[a, b]$ . Logo, segue do lema de du Bois-Reymond (Lema 1.2.2) que

$$L_y(t, \bar{y}(t), \bar{y}'(t)) - \frac{d}{dt} L_{y'}(t, \bar{y}(t), \bar{y}'(t)) = 0, \quad \forall t \in [a, b]. \quad (1.11)$$

Em (1.11) podemos reconhecer a equação de Euler-Lagrange, obtida no Parágrafo 1.1 (veja equação (1.6)). Esta equação diferencial de segunda ordem nos fornece uma condição necessária para determinação de mínimos locais de um funcional. Como foi visto no Teorema 1.1.10, esta condição é também suficiente para otimalidade, caso o funcional  $I$  seja convexo. Podemos resumir a discussão acima no seguinte teorema:

**Teorema 1.3.1** *Seja  $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  duas vezes continuamente diferenciável e  $\bar{y} \in C^2[a, b]$  um mínimo local fraco do problema (1.2). Então  $\bar{y}$  é solução do problema de valor de contorno*

$$\begin{cases} L_y(t, y, y') - \frac{d}{dt} L_{y'}(t, y, y') = 0, & t \in (a, b) \\ y(a) = y_a, & y(b) = y_b. \end{cases} \quad (1.12)$$

**Demonstração:** Veja acima. ■

**Observação 1.3.2 (problemas com fronteira livre)** Suponha que apenas uma condição de contorno seja fornecida no problema (1.2), por exemplo  $y(b) = y_b$ . Construindo a família de funções admissíveis

$$y(\cdot; \varepsilon) := \bar{y}(\cdot) + \varepsilon \eta(\cdot), \quad \eta \in C^1[a, b], \quad \eta(b) = 0,$$

obtemos através de um desenvolvimento análogo ao anterior, a condição necessária

$$\begin{aligned} 0 &= \delta I(\bar{y}; \eta) \\ &= L_{y'}(a, \bar{y}(a), \bar{y}'(a))\eta(a) - \int_a^b \left[ L_y(t, \bar{y}, \bar{y}') - \frac{d}{dt} L_{y'}(t, \bar{y}, \bar{y}') \right] \eta(t) dt, \end{aligned} \quad (1.13)$$

para todo  $\eta \in C^1[a, b]$  com  $\eta(b) = 0$ . Note que (1.13) vale em particular para  $\eta \in C_0^1[a, b]$ , de onde segue (novamente argumentando com o Lema 1.2.2) a equação de Euler–Lagrange. Tomando agora em (1.13)  $\eta \in C^1[a, b]$  com  $\eta(b) = 0$  e  $\eta(a) \neq 0$ , obtemos a condição extra

$$L_{y'}(a, \bar{y}(a), \bar{y}'(a)) = 0. \quad (1.14)$$

Sendo assim, para problemas com fronteira livre obtemos, além da equação de Euler–Lagrange, uma condição de contorno para  $\bar{y}$  no extremo do intervalo  $[a, b]$  onde o problema não possui nenhuma restrição. Tal condição surge *naturalmente* da formulação variacional do problema de otimização, sendo por isso denominada **condição de contorno natural**. Resumindo, para que  $\bar{y}$  seja solução do problema

$$\begin{cases} \text{Minimizar } I(y) := \int_a^b L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ y \in \{y \in C^1[a, b] \mid y(b) = y_b\} \end{cases}$$

é necessário que  $\bar{y}$  seja solução do problema de valor de contorno

$$\begin{cases} L_y(t, y, y') - \frac{d}{dt} L_{y'}(t, y, y') = 0 \\ y(b) = y_b, L_{y'}(a, y(a), y'(a)) = 0. \end{cases}$$

Observe que, como no Teorema 1.3.1, a condição necessária é expressa na forma (implícita) de uma equação diferencial de segunda ordem (equação de Euler–Lagrange) com duas condições de contorno.  $\square \square \square$

As hipóteses  $\bar{y} \in C^2[a, b]$  e  $L \in C^2([a, b] \times \mathbb{R}^2; \mathbb{R})$  são por demasiado restritivas e devem ser enfraquecidas. Com isso em mente, analisamos uma aplicação do lema de du Bois–Reymond semelhante a utilizada na obtenção da equação (1.10) (veja Exercício 1.2).

**Teorema 1.3.3** *Seja  $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  continuamente diferenciável e  $\bar{y} \in C^1[a, b]$  um mínimo local fraco do o problema (1.2). Então a aplicação*

$$[a, b] \ni t \mapsto L_{y'}(t, \bar{y}(t), \bar{y}'(t)) \in \mathbb{R}$$

*é continuamente diferenciável e  $\bar{\eta}$  é solução do problema de valor de contorno (1.12).*

**Demonstração:** Repetindo a argumentação feita na demonstração do Teorema 1.3.1, obtemos para  $\eta \in C_0^1[a, b]$

$$\int_b^a [L_y(t, \bar{y}(t), \bar{y}'(t)) \eta(t) + L_{y'}(t, \bar{y}(t), \bar{y}'(t)) \eta'(t)] dt = 0. \quad (1.15)$$

Integrando por partes e observando que  $\eta(a) = \eta(b) = 0$ , segue que

$$\int_a^b \left[ - \int_a^t L_y(s, \bar{y}(s), \bar{y}'(s)) ds + L_{y'}(t, \bar{y}(t), \bar{y}'(t)) \right] \eta'(t) dt = 0.$$

O Lema 1.2.1 implica na existência de uma constante  $c$ , que satisfaz

$$-\int_a^t L_y(s, \bar{y}(s), \bar{y}'(s)) ds + L_{y'}(t, \bar{y}(t), \bar{y}'(t)) = c, \quad t \in [a, b]. \quad (1.16)$$

Como a aplicação

$$[a, b] \ni t \mapsto -\int_a^t L_y(s, \bar{y}(s), \bar{y}'(s)) ds \in \mathbb{R}$$

é continuamente diferenciável, segue de (1.16) que a aplicação

$$[a, b] \ni t \mapsto L_{y'}(t, \bar{y}(t), \bar{y}'(t)) \in \mathbb{R}$$

também é continuamente diferenciável (apesar de exigirmos nas hipóteses do teorema apenas  $\bar{y} \in C^1[a, b]$ ). Podemos então derivar a expressão (1.16) em relação a  $t$ , obtendo assim a equação de Euler-Lagrange. ■

Não é possível deduzir a equação de Euler-Lagrange integrando (1.15) por partes, como na demonstração do Teorema 1.3.1, pois para diferenciar a aplicação  $t \mapsto L_{y'}(t, \bar{y}(t), \bar{y}'(t))$  é preciso usar a regra de cadeia e a função  $\bar{y}$ , envolvida na composição, não é suficientemente regular. Entretanto é possível provar a diferenciabilidade de  $\bar{y}'$  em todo  $t_0 \in [a, b]$  com  $L_{y'y'}(t_0, \bar{y}(t_0), \bar{y}'(t_0)) \neq 0$ , caso  $L$  seja uma função  $C^3$ .

As considerações feitas na Observação 1.3.2, sobre as condições de contorno naturais, continuam válidas no caso  $L \in C^1$ ,  $\bar{y} \in C^1$ , como o leitor pode facilmente verificar.

**Definição 1.3.4** As soluções da equação diferencial de Euler-Lagrange (desconsiderando as condições de contorno) são denominadas **funções estacionárias** ou **extremais**, independente do fato de serem ou não soluções do problema variacional. □ □ □

**Observação 1.3.5** Suponha que  $L$  é uma aplicação continuamente diferenciável e que  $y \in Y_{ad}$  é uma função estacionária. Integrando a equação de Euler-Lagrange obtemos

$$L_{y'}(t, y, y') = \int_a^t L_y(s, y(s), y'(s)) ds + const. \quad (1.17)$$

Fazendo a hipótese extra  $y \in C^2[a, b]$ , segue da equação de Euler-Lagrange

$$\frac{d}{dt} L(t, y, y') = L_t(t, y, y') + L_y(t, y, y')y' + L_{y'}(t, y, y')y'' = L_t(t, y, y') + \frac{d}{dt} [L_{y'}(t, y, y')y'],$$

que pode ser reescrito como

$$L_t(t, y, y') = \frac{d}{dt} [L(t, y, y') - y' L_{y'}(t, y, y')].$$

Integrando esta expressão, obtemos a equação

$$L(t, y, y') - y' L_{y'}(t, y, y') = \int_a^t L_t(s, y(s), y'(s)) ds + const. \quad (1.18)$$

que é conhecida como **primeira integral** da equação de Euler–Lagrange. Note a semelhança com a equação (1.17), assim como o fato da exigência  $\bar{y} \in C^2[a, b]$  não aparecer explicitamente na equação.

A equação (1.18) pode ainda ser deduzida no caso geral em que o extremal  $y$  é uma função apenas  $C^1[a, b]$ . Uma demonstração simples, porém trabalhosa, pode ser encontrada em [Tr, Capítulo 6]. A demonstração é levada a cabo acoplando-se a abordagem variacional com mudanças de coordenadas convenientes.  $\square \square \square$

## 1.4 Três Problemas Variacionais Clássicos

Os exemplos a seguir ilustram como a equação de Euler–Lagrange é utilizada para determinar a solução de alguns problemas clássicos do cálculo variacional. São esses: determinação de geodésicas no plano; determinação de geodésicas na esfera; a Braquistócrona.

**Exemplo 1.4.1** Analisamos o problema de encontrar, dados dois pontos  $(t_0, y_0)$  e  $(t_1, y_1)$  no plano, a curva de menor comprimento que os une. Representamos as curvas admissíveis com parametrizações do tipo

$$y : [t_0, t_1] \ni t \mapsto y(t) \in \mathbb{R}; \quad y(t_0) = y_0, \quad y(t_1) = y_1.$$

Supondo as curvas admissíveis continuamente diferenciáveis, obtemos o seu comprimento pela expressão

$$J(y) := \int_{t_0}^{t_1} \sqrt{1 + y'(t)} dt.$$

Logo, o problema variacional pode ser formulado como:

$$\begin{cases} \text{Minimizar } J(y) = \int_{t_0}^{t_1} L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ y \in \{C^1[t_0, t_1] \mid y(t_0) = y_0, y(t_1) = y_1\} \end{cases}$$

com  $L(t, y, y') = (1 + (y')^2)^{1/2}$ . Da equação de Euler–Lagrange concluímos que um extremal satisfaz

$$0 = L_y - \frac{d}{dt} L_{y'} = 0 - \frac{d}{dt} \left[ \frac{y'}{(1 + (y')^2)^{1/2}} \right].$$

Portanto,

$$\frac{y'}{(1 + (y')^2)^{1/2}} = \text{const.}$$

O que implica em  $y'$  ser constante, ou seja, em  $y$  ser linear.  $\square \square \square$

**Exemplo 1.4.2** Consideramos agora o problema de, fixados dois pontos  $P$  e  $Q$  na superfície da esfera de raio  $r$ , encontrar a curva de menor comprimento que os une. Estando a esfera centrada na origem, parametrizamos-a localmente pelas coordenadas:

$$r(\cos t \cos y, \sin t \cos y, \sin y) \quad \text{com } t \in (0, 2\pi), y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

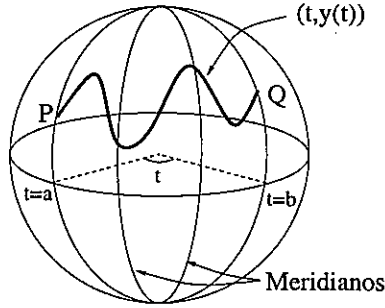


Figura 1.1: Curva admissível unindo os pontos P e Q.

por simplicidade são consideradas apenas curvas que *cortam* cada meridiano em apenas um ponto. Tais curvas podem ser parametrizadas usando a *latitude*  $t$  como parâmetro. A *longitude*  $y$  de um ponto da curva é dada por uma função  $y(t)$  da variável independente  $t$  (veja Figura 1.1). Uma curva é então representada por

$$[a, b] \ni t \mapsto r(\cos t \cos y(t), \sin t \cos y(t), \sin y(t)) \in \mathbb{R}^3.$$

Suponha que os pontos a serem unidos sejam parametrizados por:  $(a, y_a), (b, y_b)$ . Quando  $y$  é continuamente diferenciável, o comprimento da curva correspondente é dado por

$$I(y) = \int_a^b r \sqrt{\cos^2 y(t) + y'(t)^2} dt.$$

Logo,  $L(t, y, y') = r(\cos^2 y(t) + y'(t)^2)^{1/2}$ . Da primeira integral da equação de Euler (1.18) obtemos a identidade

$$-\cos^2 y = A \sqrt{\cos^2 y + (y')^2},$$

com  $A$  constante. Obviamente  $A \in (-1, 0)$  e ainda

$$A^2 (y')^2 = \cos^4 y - A^2 \cos^2 y.$$

Usando separação de variáveis, concluímos que

$$\frac{A dy}{(\cos^4 y - A^2 \cos^2 y)^{1/2}} = dt.$$

Logo

$$\int \frac{A dy}{\cos y (\cos^2 y - A^2)^{1/2}} = \pm(t - a) + B,$$

com  $B$  constante. Fazendo a mudança de variáveis  $u = \tan y$ , obtemos finalmente

$$y(t) = \arctan\left(\frac{\sqrt{1 - A^2}}{A} \sin(\pm(t - a) + B)\right),$$

onde as constantes  $A$  e  $B$  são determinadas a partir dos dados  $a, b, y_a, y_b$ .

Note que o equador  $\bar{y} = 0$  ( $A = -1$ ) é um extremal. Contudo, se  $b - a > \pi$ , uma das partes do equador não é o caminho mais curto. Tal fato nos leva a concluir que *pequenas* partes desta curva determinam o caminho mais curto, enquanto que *longas* partes não o fazem.<sup>3</sup>  $\square \square \square$

**Exemplo 1.4.3 (Braquistócrona)** Discutimos neste exemplo um problema que influenciou fortemente o desenvolvimento inicial do cálculo de variações. Em 1696 Johann Bernoulli propôs o seguinte problema:

Sejam  $P_0$  e  $P_1$  dois pontos dados sobre um plano vertical. Deve ser encontrada uma curva unindo esses dois pontos do sorte que um ponto de massam partindo de  $P_0$  que a percorra sob a influência somente de seu próprio peso, alcance  $P_1$  no menor tempo possível. Considere ainda a velocidade inicial  $v_0$  dada.

Tal problema foi formulado originalmente por Galileo em 1638, que acreditava ser a solução do problema um arco de circunferência. Outros matemáticos da época como Leibniz, Jakob Bernoulli e Newton (que resolveu o problema anonimamente, usando um pseudônimo) se interessaram pelo problema, que possui a peculiaridade de ter sido proposto por Johann Bernoulli na forma de desafio.<sup>4</sup>

Influenciado por Leibniz, J. Bernoulli deu ao problema o nome de **Braquistócrona** (do grego *brachystos* – mínimo, *chronos* – tempo) e usando o princípio de refração de Fermat (veja, e.g, [Wei, Capítulo 5]) obteve a seguinte formulação variacional para o problema:<sup>5</sup>

$$\left\{ \begin{array}{l} \text{Minimizar } J(y) := \int_{x_0}^{x_1} \left( \frac{1 + (y')^2}{2gy + c} \right)^{1/2} dx \\ \text{sujeito a} \\ y \in \{C^1[x_0, x_1] \mid y(x_0) = y_0, y(x_1) = y_1\} \end{array} \right.$$

onde  $g$  é a constante gravitacional e  $c = c_{g, v_0, y_0}$  é a energia total do corpo (cinética + potencial) no instante inicial. Note que o integrando é autônomo, i.e.  $L = L(y, y')$ . Logo, segue da primeira integral da equação de Euler (1.18) que

$$\text{const.} = L(y, y') - y' L_{y'}(y, y') = \frac{(1 + (y')^2)^{1/2}}{(2gy + c)^{1/2}} - y' \frac{(1 + (y')^2)^{-1/2}}{(2gy + c)^{1/2}} y'.$$

Após alguma manipulação algébrica, obtemos a equação diferencial

$$(2gy + c)(1 + (y')^2) = \text{const.} \quad (1.19)$$

<sup>3</sup>O tratamento de problemas como o surgido neste exemplo foi motivo de investigação por Karl Jakobi, que através da teoria de campos de extremais conseguiu esclarecer quando uma função estacionária deixa de ser solução do problema variacional. Maiores detalhes podem ser encontrados, e.g. em [Tr, Capítulo 9].

<sup>4</sup>Para mais detalhes históricos consulte [Gol].

<sup>5</sup>Para detalhes veja, e.g., [Gol], [Tr].

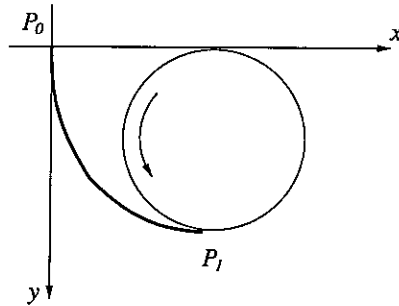


Figura 1.2: Ciclóide: solução do problema da Braquistócrona

Fazendo a hipótese simplificadora  $c = 0$  (que corresponde ao caso particular  $v_0 = y_0 = 0$ ), obtemos no lugar de (1.19) a equação diferencial

$$y(1 + (y')^2) = A. \quad (1.20)$$

Para resolver esta equação diferencial, supomos o extremal  $y$  da forma  $y(x) = A \sin^2(\frac{1}{2}\theta(x))$ . Substituindo esta expressão em (1.20), obtemos  $y' = \sqrt{(A - y)/y}$ ; de onde segue

$$A \sin(\frac{\theta}{2}) \cos(\frac{\theta}{2}) \frac{d\theta}{dx} = \frac{\cos(\frac{\theta}{2})}{\sin(\frac{\theta}{2})}.$$

Portanto,  $dx = A \sin^2(\frac{\theta}{2}) d\theta$ , ou ainda

$$x = B + \frac{1}{2}A(\theta - \sin \theta).$$

Sendo assim, obtemos para o extremal  $y$  a seguinte parametrização:

$$\gamma : [\theta_0, \theta_1] \ni \theta \mapsto (b + a(\theta + \sin \theta), a(1 + \cos \theta)) \in \mathbb{R}^2,$$

onde os parâmetros  $a, b$  são determinados pelas condições de contorno  $\gamma(\theta_0) = (x_0, y_0)$ ,  $\gamma(\theta_1) = (x_1, y_1)$ . Tal curva é denominada cicloide (veja Figura 1.2).  $\square \square \square$

## 1.5 Extremais Diferenciáveis por Partes

No Parágrafo 1.3, consideramos como admissíveis apenas as funções continuamente diferenciáveis. Essa hipótese é bastante restritiva e faz com que determinados problemas variacionais não possuam solução (veja, e.g., Exemplo 1.5.1). Neste parágrafo exigimos menos regularidade dos candidatos à solução do problema variacional e novamente buscamos condições necessárias para otimalidade.

**Exemplo 1.5.1** Considere um problema variacional escalar com integrando

$$L(t, y(t), y'(t)) := (1 - (y')^2)^2.$$

Obviamente temos  $L(t, y, y') = 0$  quando  $|y'| = 1$  e  $L(t, y, y') > 0$  caso contrário. Portanto, as funções com derivada igual a  $\pm 1$  por partes são as candidatas naturais à solução do problema variacional

$$\begin{cases} \text{Minimizar } I(y) := \int_a^b (1 - [y'(t)]^2)^2 dt \\ \text{sujeito a} \\ y \in Y_{ad} := \{y \in C^1[0, 1] \mid y(0) = y_0, y(1) = y_1\}. \end{cases}$$

Entretanto, apenas para poucos pares de condições de contorno  $\{y_0, y_1\}$ , existe uma solução  $C^1[0, 1]$  para o problema variacional. Em muito maior número são as condições de contorno para as quais os extremais são funções apenas contínuas, do tipo zig-zag.  $\square \square \square$

O Exemplo 1.5.1 ilustra a necessidade de definirmos classes de funções, que são contínuas a menos de um número finito de pontos.

**Definição 1.5.2** Uma função  $y : [a, b] \rightarrow \mathbb{R}$  é denominada **contínua por partes**, nota-se  $y \in \hat{C}[a, b]$ , quando existe uma partição  $a = t_0 < \dots < t_{n+1} = b$  tal que para todo  $i = 0, \dots, n$  a restrição  $y|_{(t_i, t_{i+1})}$  possui uma extensão contínua ao intervalo  $[a, b]$ .

Uma função contínua  $y : [a, b] \rightarrow \mathbb{R}$  é denominada **continuamente diferenciável por partes**, nota-se  $y \in \hat{C}^1[a, b]$ , quando existe uma partição  $a = t_0 < \dots < t_{n+1} = b$  tal que para todo  $i = 0, \dots, n$  a restrição  $y|_{(t_i, t_{i+1})}$  possui uma extensão continuamente diferenciável ao intervalo  $[a, b]$ .  $\square \square \square$

**Observação 1.5.3** Os conjuntos  $\hat{C}[a, b]$  e  $\hat{C}^1[a, b]$  são espaços vetoriais (normalisáveis) sobre  $\mathbb{R}$ . Como  $\hat{C}^1[a, b] \subset C[a, b]$ , então  $\|\cdot\|_\infty$  define uma norma em  $\hat{C}^1[a, b]$ . Entretanto, de especial interesse para este espaço é a norma definida por

$$\|f\|_{1,\infty} := \max_{t \in [a,b]} |f(t)| + \max_{t \in [a,b]} |f'(t)|.$$

É importante ressaltar que o espaço vetorial normado assim obtido não é completo (Banach), pois o número de pontos de discontinuidade em uma sequência de funções pode se tornar ilimitado.  $\square \square \square$

No lema a seguir é apresentada uma versão do teorema fundamental do cálculo para funções em  $\hat{C}^1[a, b]$ . Este resultado será utilizado no final deste parágrafo, para obter as desejadas condições necessárias de otimalidade.

**Lema 1.5.4** *Seja  $y \in \hat{C}^1[a, b]$ . Então vale a identidade*

$$y(t) = y(a) + \int_a^t y'(s) ds, \quad t \in [a, b].$$



**Demonstração:** Note que  $y'$  é contínua a menos de um número finito de pontos, portanto existe a integral (de Riemann)

$$\int_a^t y'(s) ds, \quad t \in [a, b].$$

Sejam  $t_1, \dots, t_n$  os pontos de discontinuidade de  $y'$ . Para  $t \in [t_k, t_{k+1}]$  temos

$$\begin{aligned} y(t) - y(a) &= y(t) - y(t_k) + \sum_{l=0}^{k-1} [y(t_{l+1}) - y(t_l)] \\ &= \int_{t_k}^t y'(s) ds + \sum_{l=0}^{k-1} \int_{t_l}^{t_{l+1}} y'(s) ds = \int_a^t y'(s) ds \end{aligned}$$

e o lema fica provado. ■

**Corolário 1.5.5** *Seja  $y \in \hat{C}^1[a, b]$ . São verdadeiras as afirmações:*

- a) Se  $\int_a^b y'(s)^2 ds = 0$  então  $y' \equiv 0$ ;  
 b) Se  $y'(t) = 0$  em todos os pontos  $t$  onde  $y$  é diferenciável, então  $y$  é constante.

**Demonstração:** Sejam  $t_1, \dots, t_n$  os pontos de discontinuidade de  $y'$ ,  $t_0 = a$  e  $t_{n+1} = b$ . Da identidade

$$0 = \int_a^b y'(s)^2 ds = \sum_{k=0}^n \int_{t_k}^{t_{k+1}} y'(s)^2 ds,$$

temos que

$$y'|_{[t_k, t_{k+1}]} = 0, \quad 0 \leq k \leq n.$$

Provando assim o item (a). O item (b) segue imediatamente do Lema 1.5.4. ■

O lema a seguir discute um modo de encontrar uma aproximação *suave* para uma função diferenciável por partes, isto é, como aproximar uma função apenas  $\hat{C}^1$  por outra  $C^1$ . É importante observar que, no processo descrito, a derivada da aproximação é mantida limitada pela derivada da função original.

**Lema 1.5.6** *Seja  $\hat{y} \in \hat{C}^1[a, b]$  e  $t_1, \dots, t_n$  os pontos de discontinuidade de  $\hat{y}'$ . Então é possível encontrar constantes  $A \in \mathbb{R}$  e  $\delta_0 > 0$ , tais que para todo  $\delta \in (0, \delta_0)$  existe  $y \in C^1[a, b]$  satisfazendo*

- a)  $y|_R = \hat{y}$ , onde  $R = [a, b] \setminus \bigcup_{i=0}^n [t_i - \delta, t_i + \delta]$ ;  
 b)  $\max_{t \in [a, b]} |y'(t)| \leq 4 \max_{t \in [a, b]} |\hat{y}'(t)|$ ;  
 c)  $\max_{t \in [a, b]} |y(t) - \hat{y}(t)| \leq A\delta$ .

**Demonstração:** Basta provar o resultado para partições com apenas um ponto  $\hat{t}$ . Escolha  $\delta_0 > 0$  tal que  $[\hat{t} - \delta_0, \hat{t} + \delta_0] \subset (a, b)$ . Seja  $\delta \in (0, \delta_0)$ .

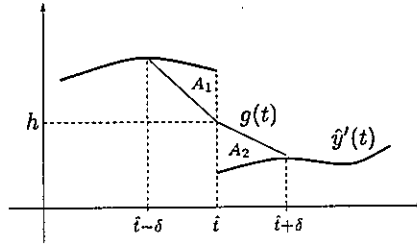


Figura 1.3: Construção da função  $g$  no Teorema 1.5.6

A idéia por trás da *suavização* consiste em substituir a derivada  $\hat{y}'$  por uma poligonal contínua em  $[\hat{t} - \delta, \hat{t} + \delta]$  e não alterar  $\hat{y}$  fora deste intervalo. Dada uma constante  $h \in \mathbb{R}$  define a função

$$g(t) := \begin{cases} \hat{y}'(t) & , |t - \hat{t}| > \delta \\ -(t - \hat{t})\delta^{-1}\hat{y}'(\hat{t} - \delta) + (t - \hat{t} + \delta)\delta^{-1}h & , t \in [\hat{t} - \delta, \hat{t}] \\ -(t - \hat{t} - \delta)\delta^{-1}h + (t - \hat{t})\delta^{-1}\hat{y}'(\hat{t} + \delta) & , t \in [\hat{t}, \hat{t} + \delta] \end{cases}$$

Apartir daí, defina a função  $y \in C^1[a, b]$

$$y(t) := \hat{y}(a) + \int_a^t g(s) ds, \quad t \in [a, b].$$

Escolhemos agora  $h$  de modo que  $y = \hat{y}$  em  $[a, b] \setminus [\hat{t} - \delta, \hat{t} + \delta]$ , satisfazendo assim (a). Este é o caso quando a condição

$$\int_{\hat{t}-\delta}^{\hat{t}+\delta} g(s) ds = \int_{\hat{t}-\delta}^{\hat{t}+\delta} \hat{y}'(s) ds =: A_\delta$$

é satisfeita. Na Figura 1.3, observamos que isto equivale a exigir que as áreas  $A_1$  e  $A_2$  sejam iguais. Note que  $A_\delta = h\delta + \frac{\delta}{2}[\hat{y}'(\hat{t} - \delta) + \hat{y}'(\hat{t} + \delta)]$ . Escolhemos portanto

$$h = \frac{A_\delta}{\delta} - \frac{1}{2}[\hat{y}'(\hat{t} - \delta) + \hat{y}'(\hat{t} + \delta)].$$

Definindo agora  $M := \max_{t \in [a, b]} |\hat{y}'(t)|$  obtemos

$$|A_\delta| \leq 2\delta M \quad \text{e} \quad |h| \leq 3M.$$

Da definição de  $g$ , segue que para todo  $t \in [a, b]$

$$\begin{aligned} |y'(t)| &= |g(t)| \\ &\leq \max\{|\hat{y}'(t)|, |\hat{y}'(\hat{t} - \delta)| + |h|, |h| + |\hat{y}'(\hat{t} + \delta)|\} \\ &\leq M + |h| \\ &\leq 4M, \end{aligned}$$

provando assim (b). Para determinar  $A$  em (c) basta observar que

$$|y(t) - \hat{y}(t)| \leq \int_{\hat{t}-\delta}^t |y'(s) - \hat{y}'(s)| ds \leq \int_{\hat{t}-\delta}^{\hat{t}+\delta} (|y'(s)| + |\hat{y}'(s)|) ds \leq 10M \delta.$$

Fica assim provado o teorema. ■

Estamos agora em condições de obter condições necessárias de otimalidade para problemas variacionais formulados em  $\hat{Y}_{ad}$ . Antes porém, analisamos uma importante correspondência existente entre as soluções do problema (1.2) e do problema

$$\begin{cases} \text{Minimizar } I(y) := \int_a^b L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ y \in \hat{Y}_{ad} := \{y \in \hat{C}^1[a, b] \mid y(a) = y_a, y(b) = y_b\}. \end{cases} \quad (1.21)$$

A fim de simplificar a notação, denominamos os problemas (1.2) e (1.21) por problemas  $(P)$  e  $(\hat{P})$  respectivamente. Observe que o conceito de mínimo local fraco na Definição 1.1.3 é naturalmente estendido ao problema  $(\hat{P})$ , bastando para isso substituir o espaço  $Y_{ad}$  por  $\hat{Y}_{ad}$  naquela definição.

**Teorema 1.5.7** *Seja  $L : [a, b] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  uma aplicação contínua. Se  $\bar{y} \in Y_{ad}$  é mínimo local fraco do problema  $(P)$ , então  $\bar{y}$  também é mínimo local fraco do problema  $(\hat{P})$ .*

**Demonstração:** Se  $\bar{y}$  é mínimo local fraco de  $(P)$  então existe  $\varepsilon > 0$  tal que

$$I(\bar{y}) \leq I(v), \quad \forall v \in S_\varepsilon := \{v \in Y_{ad} \mid \|v - \bar{y}\|_{1,\infty} < \varepsilon\}.$$

Tome  $\hat{y} \in \hat{Y}_{ad}$  com  $\|\hat{y} - \bar{y}\|_{1,\infty} < \varepsilon/5$ . Dado  $\delta > 0$  suficientemente pequeno, o Lema 1.5.6 garante a existência de  $v_{\varepsilon,\delta} \in C^1[a, b]$  tal que

$$\|v_{\varepsilon,\delta} - v\|_\infty \leq A\delta \quad \text{e} \quad \|v'_{\varepsilon,\delta}\|_\infty \leq 4\|v'\|_\infty,$$

onde  $v := \hat{y} - \bar{y} \in \hat{C}^1[a, b]$  e a constante  $A$  depende apenas de  $\hat{y}$  e  $\bar{y}$ . Definindo a função  $y_\varepsilon := \bar{y} + v_{\varepsilon,\delta} \in Y_{ad}$ , temos

$$\begin{aligned} \|y_\varepsilon - \bar{y}\|_{1,\infty} &= \|v_{\varepsilon,\delta}\|_{1,\infty} = \|v_{\varepsilon,\delta}\|_\infty + \|v'_{\varepsilon,\delta}\|_\infty \\ &\leq \|v_{\varepsilon,\delta} - v\|_\infty + \|v\|_\infty + 4\|v'\|_\infty \\ &\leq A\delta + 4\|\hat{y} - \bar{y}\|_{1,\infty} \\ &\leq A\delta + \frac{4\varepsilon}{5}. \end{aligned}$$

Logo,  $y_\varepsilon \in S_\varepsilon$  se escolhermos  $\delta < \varepsilon/5A$ . Temos assim

$$\begin{aligned} I(\hat{y}) &\geq I(y_\varepsilon) - |I(y_\varepsilon) - I(\hat{y})| \\ &\geq I(\bar{y}) - |I(\bar{y} + v_{\varepsilon,\delta}) - I(\bar{y} + v)| \\ &\geq I(\bar{y}) - \int_{\hat{t}-\delta}^{\hat{t}+\delta} |L(t, \bar{y} + v_{\varepsilon,\delta}, (\bar{y} + v_{\varepsilon,\delta})') - L(t, \bar{y} + v, (\bar{y} + v)')| dt. \end{aligned}$$

Tomando o limite  $\delta \rightarrow 0$ , a integral do lado direito converge a zero e obtemos a desigualdade  $I(\hat{y}) \geq I(\bar{y})$ . ■

Resumindo, o Teorema 1.5.7 garante que toda solução  $C^1$  é também uma solução  $\hat{C}^1$  de um problema variacional. Este resultado permanece válido para problemas com condições de contorno naturais (veja Observação 1.3.2). É ainda possível provar um resultado análogo para mínimos globais.

A recíproca do Teorema 1.5.7 não é verdadeira, conforme ilustra o exemplo a seguir.

**Exemplo 1.5.8** Considere a função  $L(t, y, y') := y^2(1 - y')^2$  definida em  $[a, b] = [-1, 1]$ , e as condições de contorno  $y_a = 0$ ,  $y_b = 1$ . Observe que a função

$$\bar{y}(t) := \begin{cases} 0 & , t \leq 0 \\ t & , t \geq 0 \end{cases}$$

é um mínimo global de  $I$  em  $\hat{Y}_{ad}$ , pois  $I(\bar{y}) = 0$ . Verificamos a seguir que não existe outro mínimo global. Suponha que  $y \in \hat{Y}_{ad}$  é um outro mínimo. Como  $y(1) = 1$ , o escalar  $\alpha$  definido por

$$\alpha := \inf\{\beta \in [-1, 1] \mid y(t) > 0, \beta < t \leq 1\}$$

necessariamente satisfaz  $\alpha \in [-1, 1]$ . Da identidade  $I(y) = I(\bar{y}) = 0$ , obtemos  $y'(t) = 1$  para  $t \in (\alpha, 1]$ , i.e.

$$y(t) = t, \quad t \in [\alpha, 1].$$

Como  $y$  é contínua, segue da definição de  $\alpha$  que  $\alpha = 0$ . Logo

$$y(t) = t, \quad t \in [0, 1].$$

Note agora que  $y(-1) = y(0) = 0$ . Note ainda que para todo  $t \in (-1, 0)$  com  $y(t) \neq 0$ , temos necessariamente  $y'(t) = 1$ . Logo, podemos concluir que  $y(t) = 0$ ,  $t \in [-1, 0]$ . Conseqüentemente, o problema variacional não possui solução em  $Y_{ad}$  e a única solução em  $\hat{Y}_{ad}$  é a função  $\bar{y}$ . □ □ □

O próximo teorema fornece as procuradas condições necessárias de otimalidade para o problema variacional  $(\hat{P})$ .

**Teorema 1.5.9** *Seja  $L : [a, b] \times \mathbb{R} \times \mathbb{R}$  continuamente diferenciável. Se  $\hat{y} \in \hat{Y}_{ad}$  é um mínimo local fraco do problema  $(\hat{P})$ , então*

a) *existe uma constante  $c \in \mathbb{R}$  tal que*

$$L_{y'}(t, \hat{y}(t), \hat{y}'(t)) = \int_a^t L_y(s, \hat{y}(s), \hat{y}'(s)) ds + c, \quad t \in [a, b]; \quad (1.22)$$

b)  *$\hat{y}$  satisfaz a condição de Weierstrass-Erdmann*

$$L_{y'}(t, \hat{y}(t-), \hat{y}'(t-)) = L_{y'}(t, \hat{y}(t+), \hat{y}'(t+)), \quad t \in (a, b); \quad (1.23)$$

c)  *$\hat{y}$  satisfaz a equação de Euler-Lagrange nos pontos  $t \in (a, b)$  de continuidade de  $\hat{y}'$ .*

**Demonstração:** Dado  $\varepsilon_0 > 0$  escolha  $\eta \in \hat{C}^1[a, b]$  com  $\eta(a) = \eta(b) = 0$ . Como  $\hat{y}$  é mínimo local temos  $\delta I(\hat{y}; \eta) = 0$ . Da hipótese de diferenciabilidade de  $L$ , segue que  $0 = \delta I(\hat{y}; \eta) = \frac{d}{d\varepsilon} I(\hat{y} + \varepsilon\eta)|_{\varepsilon=0}$ , i.e.

$$\int_a^b [L_y(s, \hat{y}(s), \hat{y}'(s)) \eta(s) + L_{y'}(s, \hat{y}(s), \hat{y}'(s)) \eta'(s)] ds = 0.$$

Integrando por partes obtemos

$$\int_a^b \left[ - \int_a^s L_y(r, \hat{y}(r), \hat{y}'(r)) dr + L_{y'}(s, \hat{y}(s), \hat{y}'(s)) \right] \eta'(s) ds = 0. \quad (1.24)$$

Escolhamos agora a constante  $c$  (usando o teorema do valor médio) de forma que a função  $v \in \hat{C}^1[a, b]$  definida por

$$v(t) := \int_a^t \left[ - \int_a^s L_y(r, \hat{y}(r), \hat{y}'(r)) dr + L_{y'}(s, \hat{y}(s), \hat{y}'(s)) - c \right] ds$$

satisfaça a condição

$$v(a) = v(b) = 0.$$

Tomando agora  $f(s) := - \int_a^s L_y(r, \hat{y}(r), \hat{y}'(r)) dr + L_{y'}(s, \hat{y}(s), \hat{y}'(s))$ ,  $s \in [a, b]$ , segue da equação (1.24)

$$\begin{aligned} \int_a^b v'(s)^2 ds &= \int_a^b [f(s) - c]^2 ds = \int_a^b f(s) [f(s) - c] ds - c \int_a^b [f(s) - c] ds \\ &= 0 - c \int_a^b f(s) ds + c^2(b-a) = 0. \end{aligned}$$

O Corolário 1.5.5 nos permite concluir que  $v' \equiv 0$  em  $[a, b]$ , o que implica em

$$- \int_a^t L_y(s, \hat{y}(s), \hat{y}'(s)) ds + L_{y'}(t, \hat{y}(t), \hat{y}'(t)) - c = 0, \quad t \in [a, b].$$

Fica assim provado o item *a*). Para provar *b*), basta observar que (1.22) implica na continuidade da aplicação  $t \mapsto L_{y'}(t, \hat{y}(t), \hat{y}'(t))$  em  $[a, b]$ .

Provamos agora *c*). Se  $\hat{y}'$  é contínua no ponto  $t \in (a, b)$ , segue de (1.22) que a aplicação

$$s \mapsto L_{y'}(s, \hat{y}(s), \hat{y}'(s))$$

é continuamente diferenciável em uma vizinhança de  $t$ . Logo, a Equação de Euler-Lagrange é obtida simplesmente derivando-se (1.22), como na demonstração do Teorema 1.3.3. ■

Resumindo, o Teorema 1.5.9 garante que um mínimo local fraco do problema  $(\hat{P})$  satisfaz a equação de Euler-Lagrange nos pontos de continuidade de sua derivada. Mais ainda, as únicas discontinuidades possíveis em um mínimo local fraco, são aquelas que preservam a continuidade da aplicação

$$t \mapsto L_{y'}(t, \hat{y}(t), \hat{y}'(t)).$$

Esta condição (veja também (1.23)) é denominada **condição de Weierstraß–Erdmann**. É possível ainda provar que um mínimo local fraco do problema  $(\tilde{P})$  satisfaz a primeira integral de equação de Euler–Lagrange e que a aplicação

$$t \longmapsto L(t, \hat{y}(t), \hat{y}'(t)) - \hat{y}'(t) L_{y'}(t, \hat{y}(t), \hat{y}'(t))$$

é contínua (veja Exercícios 1.3 e 1.4). Esta última condição é conhecida como **segunda condição de Weierstraß–Erdmann**.

## 1.6 Problemas Vetoriais

Extendemos neste parágrafo, para problemas variacionais vetoriais, as condições necessárias de otimalidade obtidas para os problemas escalares. Considere para tanto o problema variacional

$$\begin{cases} \text{Minimizar } I(y) := \int_a^b L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ y \in \hat{Y}_{ad} := \{y \in \hat{C}^1([a, b]; \mathbb{R}^n) \mid y(a) = y_a, y(b) = y_b\}, \end{cases} \quad (1.25)$$

onde  $L \in C^1([a, b] \times \mathbb{R}^{2n}; \mathbb{R})$  e  $y_a, y_b \in \mathbb{R}^n$  são dados. Nos problemas vetoriais as funções admissíveis são aplicações vetoriais:  $y(t) = (y_1(t), \dots, y_n(t))$ , com  $y_j \in \hat{C}^1[a, b]$ .

O resultado a seguir generaliza para o problema (1.25) as condições necessárias obtidas para o problema variacional escalar (1.21).

**Teorema 1.6.1** *Seja  $L : [a, b] \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$  continuamente diferenciável. Se  $\hat{y} \in \hat{Y}_{ad}$  é mínimo local fraco do problema (1.25), então  $\hat{y}$  satisfaz a equação vetorial de Euler–Lagrange*

$$\frac{d}{dt} L_{y'}(t, \hat{y}, \hat{y}') = L_y(t, \hat{y}, \hat{y}')$$

nos pontos  $t \in [a, b]$  de continuidade de  $\hat{y}'$ . Também é satisfeita a primeira integral da equação de Euler<sup>6</sup>

$$L(t, \hat{y}, \hat{y}') - \langle \hat{y}'(t), L_{y'}(t, \hat{y}, \hat{y}') \rangle = \int_a^t L_t(s, \hat{y}(s), \hat{y}'(s)) ds + \text{const.}, \quad t \in [a, b].$$

Além disso, nos pontos  $t \in [a, b]$  de discontinuidade de  $\hat{y}'$  são satisfeitas as condições de Weierstraß–Erdmann

$$L_{y'}(t-, \hat{y}(t-), \hat{y}'(t-)) = L_{y'}(t+, \hat{y}(t+), \hat{y}'(t+)) \quad \text{e} \quad [L - \langle \hat{y}', L_{y'} \rangle](t-) = [L - \langle \hat{y}', L_{y'} \rangle](t+).$$

(Note que a primeira condição de Weierstraß–Erdmann é vetorial, enquanto que a segunda é dada por uma equação escalar.)

**Demonstração:** É deixada como exercício por ser análoga a demonstração do Teorema 1.5.9 (basta deduzir as equações vetoriais componente a componente). ■

<sup>6</sup>Note que a primeira integral é uma equação escalar, enquanto que a equação de Euler–Lagrange é vetorial.

Nos problemas vetoriais considerados neste parágrafo a função objetivo é escalar. Em muitos problemas de controle ótimo (e em seus correspondentes variacionais) é conveniente considerar funções objetivo vetoriais, i.e.  $I : \hat{Y}_{ad} \rightarrow \mathbb{R}^k$ . Uma alternativa neste caso é utilizar o conceito de Pareto otimalidade para eleger uma estratégia ótima (veja, e.g., [Lei, Capítulo 17]).

## 1.7 Exercícios

1.1 Calcule a primeira integral da Equação de Euler–Lagrange quando  $L = L(t)$ . Mostre que as funções lineares são estacionárias.

1.2 Prove que se  $g, h \in C[a, b]$  são tais que

$$\int_a^b [g(t)v(t) + h(t)v'(t)] dt = 0,$$

para todo  $v \in C_0^1[a, b]$ , então  $h \in C^1[a, b]$  e  $h' = g$ .  
(Sugestão: Veja demonstração do Teorema 1.3.3.)

1.3 Seja  $L$  uma aplicação  $C^1$ . Mostre que todo mínimo local fraco  $\hat{y} \in \hat{Y}_{ad} \cap C^2[a, b]$  do problema  $(\hat{P})$  satisfaz a primeira integral da Equação de Euler–Lagrange

$$L(\hat{y}, \hat{y}') - \hat{y}' L_{y'}(\hat{y}, \hat{y}') = \text{const.}$$

1.4 Seja  $L$  uma aplicação  $C^1$  e  $\hat{y} \in \hat{Y}_{ad}$  um mínimo local fraco do problema  $(\hat{P})$ . Mostre que a segunda condição de Weierstrass–Erdmann

$$(L - \hat{y}' L_{y'})(t^-) = (L - \hat{y}' L_{y'})(t^+), \quad t \in (a, b)$$

é satisfeita.

(Sugestão: Use o fato de  $\hat{y}$  satisfazer a primeira integral da equação de Euler–Lagrange.)





## Capítulo 2

# Problemas Variacionais e Controle Ótimo

Neste capítulo apresentamos um paralelo entre problemas do cálculo de variações e problemas de controle ótimo. O objetivo principal é comparar as condições de otimalidade para ambas as famílias de problemas.

Iniciamos o capítulo apresentando os problemas de controle ótimo. A seguir, estudamos condições necessárias de otimalidade para problemas variacionais restritos e aplicamo-as a problemas de controle ótimo correspondentes. Os dois últimos parágrafos são dedicados a análise de condições de otimalidade para problemas de controle.

O Parágrafo 2.1 é dedicado a formulação matemática dos problemas de controle ótimo. Na seqüência, são estudados no Parágrafo 2.2 problemas variacionais com condições de contorno transversais, restrições isoperimétricas e restrições lagrangeanas (veja Parágrafo 1.1). As condições de otimalidade obtidas para estes problemas são aplicadas na análise de problemas de controle ótimo correspondentes.

No Parágrafo 2.3 analisamos problemas de controle ótimo sujeitos a um tipo de restrição muito utilizada em aplicações modeladas por equações escalares. O desenvolvimento é baseado na análise dos problemas variacionais correspondentes. O interesse em tais problemas de controle está no fato das soluções serem do tipo bang-bang.

Nos dois últimos Parágrafos, 2.4 e 2.5, estudamos, sob hipóteses adicionais de convexidade, a suficiência e necessidade para otimalidade de um conjunto especial de condições (princípio do máximo).

### 2.1 Problemas de Controle Ótimo: Apresentação

Em um problema de controle, uma variável de estado  $z = z(t) \in \mathbb{R}^n$  dependente do tempo evolui de acordo com uma dada dinâmica

$$z'(t) = f(t, z(t), u(t)), \quad t > t_0 \tag{2.1}$$

a partir de um estado (condição) inicial  $z(t_0) = z_0$ . Aqui  $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  corresponde ao modelo estudado,  $z_0 \in \mathbb{R}^n$  é o estado inicial do sistema e  $u : \mathbb{R} \rightarrow \mathbb{R}^m$  é um parâmetro

livre influenciando a dinâmica, denominado **controle** do sistema. Em muitos problemas é também fornecida uma condição de contorno final  $z(t_1) = z_1$ , ou ainda uma condição de contorno transversal. A equação (2.1) é denominada **equação de estado**.

Nos problemas de controle ótimo considerados neste capítulo, a tarefa que se impõe é a de minimizar funcionais do tipo

$$J(u, z) := \int_{t_0}^{t_1} L(t, z(t), u(t)) dt,$$

com  $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ , onde  $z$  e  $u$  estão relacionados pela dinâmica  $z' = f(t, z, u)$ ,  $t \in (t_0, t_1)$  e ainda  $z(0) = z_0$ ,  $z(t_1) = z_1$ ,  $u \in \mathcal{U}_{ad}$ .

O conjunto  $\mathcal{U}_{ad}$  é denominado **conjunto de controles admissíveis**, sendo usualmente escolhido como subconjunto de  $C([t_0, t_1]; \mathbb{R}^m)$ ,  $\tilde{C}([t_0, t_1]; \mathbb{R}^m)$  ou  $L^1([t_0, t_1]; \mathbb{R}^m)$ . Note que diferentes escolhas de  $\mathcal{U}_{ad}$  implicam em diferentes graus de regularidade de  $z$  e também determinam o conceito que deve ser utilizado para definir a solução do problema de valor inicial em (2.1).

Podemos formular o problema de controle ótimo descrito acima na forma resumida

$$\begin{cases} \text{Minimizar } J(u, z) := \int_{t_0}^{t_1} L(t, z(t), u(t)) dt \\ \text{sujeito a } u \in \mathcal{U}_{ad}, \\ z' = f(t, z, u), \quad t \in (t_0, t_1), \quad z(0) = z_0, \quad z(t_1) = z_1. \end{cases}$$

A analogia entre os problemas de controle ótimo e o problema do cálculo de variações se torna evidente quando observamos que, no caso particular  $f(t, z, u) = u$ , o problema acima toma a forma do problema variacional (1.2). É exatamente esta semelhança que nos permite comparar as condições de otimalidade para ambas as classes de problemas.

Assim como os problemas variacionais, os problemas de controle ótimo podem ser formulados com condições de contorno transversais, restrições lagrangeanas ou restrições isoperimétricas. No próximos dois parágrafos tratamos problemas variacionais restritos e problemas de controle que podem ser analisados usando as condições obtidas para tais problemas variacionais.

No problema acima os tempos inicial e final são dados (**tempo fixo**), porém o tempo final pode ser uma das incógnitas no problema de controle ótimo (**tempo livre**). De interesse são ainda os problemas formulados no intervalo  $[t_0, \infty)$  (**horizonte infinito**), assim como problemas em que são admitidas discontinuidades na variável de estado  $z$  (**problemas impulsivos**). Tais problemas de controle são abordados no capítulo seguinte, juntamente com uma formulação bastante geral do problema acima.

## 2.2 Problemas Variacionais com Restrição

Com a intenção de contruir um paralelo entre as condições de otimalidade do cálculo variacional e do controle ótimo, analisamos neste parágrafo problemas variacionais em que são impostas restrições as trajetórias admissíveis.

O interesse no estudo dos problemas (2.2), (2.6) e (2.10) para a teoria de controle ótimo se deve ao fato de que diversas famílias de problemas possuem equação de dinâmica da forma:  $z'(t) = u(t)$ ; ou ainda  $z'(t) = f_1(t, z) + f_2(t, z)u(t)$ , com  $f_2(t, z) \neq 0$ . Neste caso é possível reescrever os problemas de controle ótimo como problemas variacionais e assim utilizar as condições necessárias apresentadas neste parágrafo.

Antes de iniciarmos a análise dos problemas variacionais restritos, apresentamos uma versão do teorema de multiplicadores de Lagrange em espaços vetoriais genéricos. Para tanto, consideramos que tanto a função objetivo quanto a restrição são descritas por aplicações Gâteaux diferenciáveis.

**Teorema 2.2.1 (Lagrange)** *Seja  $Y$  um espaço vetorial normado e  $I, G$  aplicações de  $Y$  em  $\mathbb{R}$ . Suponha que as variações de Gâteaux  $\delta I(y, v)$ ,  $\delta G(y, v)$  estão bem definidas para  $y, v \in Y$  e ainda que para cada  $y, v \in Y$  tenhamos  $\delta I(y_n, v) \rightarrow \delta I(y, v)$ ,  $\delta G(y_n, v) \rightarrow \delta G(y, v)$  sempre que  $y_n \rightarrow y$  em  $Y$ .<sup>1</sup> Se  $\bar{y}$  é um mínimo local de  $I$  sujeito a restrição  $G(y) = 0$ , então ou  $\delta G(\bar{y}, \cdot) \equiv \theta$ , em  $Y$  ou existe um multiplicador  $\lambda \in \mathbb{R}$ , tal que*

$$\delta I(\bar{y}, v) = \lambda \delta G(\bar{y}, v), \quad \forall v \in Y.$$

**Demonstração:** Veja, e.g., [Tr, Capítulo 5]. ■

Na realidade o teorema acima continua válido se exigimos continuidade fraca de  $\delta I$  e  $\delta G$  apenas em  $\bar{y}$ . Uma consequência imediata do teorema de Lagrange é que os conjuntos de nível  $\{y \in Y; I(y) = I(\bar{y})\}$  e  $\{y \in Y; G(y) = 0\}$  possuem o mesmo hiperplano tangente em  $\bar{y}$ .

### Condições de Contorno Transversais

Começamos por estudar os problemas variacionais com condições de contorno transversais. Considere o problema

$$\begin{cases} \text{Minimizar } I(y, T) := \int_a^T L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ T > a; \sigma(T, y(T)) = 0; y \in \{y \in C^1[a, T] \mid y(a) = y_a\}; \end{cases} \quad (2.2)$$

onde  $L \in C^1([a, b] \times \mathbb{R}^2; \mathbb{R})$ ,  $y_a \in \mathbb{R}$  e  $\sigma \in C^1([a, b] \times \mathbb{R}; \mathbb{R})$  com  $\nabla \sigma \neq \theta$  (a constante  $b \in \mathbb{R}$  é tomada grande o bastante, de forma a não interferir na formulação do problema).

Note que nem o tempo final nem a condição de contorno final são especificados. A condição final para as funções admissíveis é fornecida pela curva de nível zero da função  $\sigma$  (veja Figura 2.1).

Suponha que  $(\bar{y}, \bar{T})$  com  $\bar{y} \in C^1[a, \bar{T}]$  é um mínimo local fraco para o problema (2.2). Logo, para funções teste  $\eta \in C_0^1[a, \bar{T}]$ , temos

$$\left. \frac{d}{d\varepsilon} I(\bar{y} + \varepsilon\eta, \bar{T}) \right|_{\varepsilon=0} = 0.$$

<sup>1</sup>Tal propriedade é denominada continuidade fraca da variação de Gâteaux.

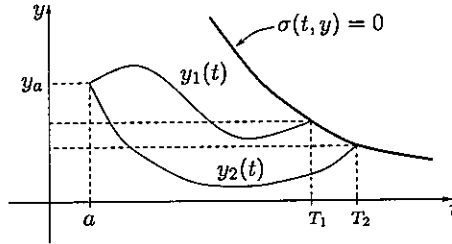


Figura 2.1: Condição de contorno transversal e funções admissíveis

Argumentando agora como no Parágrafo 1.3, concluímos que  $\bar{y}$  satisfaz a equação de Euler-Lagrange em  $(a, \bar{T})$ . Temos assim as condições necessárias

$$y(a) = y_a, \quad \frac{d}{dt} L_{y'}(t, \bar{y}, \bar{y}') = L_y(t, \bar{y}, \bar{y}'), \quad t \in (a, \bar{T}).$$

Falta ainda obter, como na Observação 1.3.2, uma condição de contorno natural que relacione  $\bar{T}$ , o tempo final, com a função  $\sigma$ , que descreve a condição final. Essa condição natural é obtida como consequência do teorema de multiplicadores de Lagrange (veja Teorema 2.2.1).

Defina  $Y_{ad} := \{y \in C^1[a, b] \mid y(a) = y_a\}$ . Note que os candidatos a solução do problema variacional (2.2) são pares da forma  $(y, T) \in Y_{ad} \times \mathbb{R} \subset C^1[a, b] \times \mathbb{R} =: Y$ . Note ainda que  $Y$  é um espaço vetorial normado com a norma do produto cartesiano  $\|(y, T)\| := \|y\|_{1, \infty} + |T|$ . A função objetivo  $I$  está bem definida como uma aplicação de  $Y$  em  $\mathbb{R}$  e a variação de  $I$  em  $(\bar{y}, \bar{T})$  na direção  $(\eta, \gamma) \in Y$  é dada por<sup>2</sup>

$$\delta I(\bar{y}, \bar{T}; \eta, \gamma) = L(\bar{T}, \bar{y}, \bar{y}') \gamma + \int_a^{\bar{T}} [L_y(t, \bar{y}, \bar{y}') \eta(t) + L_{y'}(t, \bar{y}, \bar{y}') \eta'(t)] dt.$$

Como  $\bar{y}$  é estacionária em  $(a, \bar{T})$ , podemos usar a equação de Euler-Lagrange para reescrever o integrando da expressão acima. Obtemos assim

$$\begin{aligned} \delta I(\bar{y}, \bar{T}; \eta, \gamma) &= L(\bar{T}, \bar{y}, \bar{y}') \gamma + \int_a^{\bar{T}} \frac{d}{dt} [L_{y'}(t, \bar{y}, \bar{y}') \eta(t)] dt \\ &= L(\bar{T}, \bar{y}, \bar{y}') \gamma + [L_{y'}(t, \bar{y}, \bar{y}') \eta(t)]_a^{\bar{T}}. \end{aligned}$$

Observe que a condição final transversal pode ser reescrita como  $G(y, T) = 0$ , onde  $G : Y \ni (y, T) \mapsto \sigma(T, y(T)) \in \mathbb{R}$ . Calculando a variação de Gâteaux de  $G$  em  $(\bar{y}, \bar{T})$  na direção  $(\eta, \gamma) \in Y$  obtemos<sup>3</sup>

$$\delta G(\bar{y}, \bar{T}; \eta, \gamma) = \sigma_t(\bar{T}, \bar{y}) \gamma + \sigma_y(\bar{T}, \bar{y}) [\bar{y}'(\bar{T}) \gamma + \eta(\bar{T})].$$

<sup>2</sup>Lembre que  $\delta I(\bar{y}, \bar{T}; \eta, \gamma) = (d/d\varepsilon) I(\bar{y} + \varepsilon \eta, \bar{T} + \varepsilon \gamma)|_{\varepsilon=0}$ .

<sup>3</sup>Note que  $\delta G(\bar{y}, \bar{T}; \eta, \gamma) = (d/d\varepsilon) \sigma(\bar{T} + \varepsilon \gamma, (\bar{y} + \varepsilon \eta)(\bar{T} + \varepsilon \gamma))|_{\varepsilon=0}$ . Para simplificar a notação escrevemos  $\sigma(\bar{T}, \bar{y})$  ao invés de  $\sigma(\bar{T}, \bar{y}(\bar{T}))$ .

O teorema de multiplicadores de Lagrange garante que se  $(\bar{y}, \bar{T})$  é um mínimo local de  $I$  sujeito a restrição  $G(y, T) = 0$ , então ou  $\delta G(\bar{y}, \bar{T}; \cdot, \cdot) \equiv 0$ , ou existe um multiplicador  $\lambda \in \mathbb{R}$  tal que  $\delta(I + \lambda G)(\bar{y}, \bar{T}; \cdot, \cdot) \equiv 0$ . Como o primeiro caso implica em  $\nabla \sigma(\bar{T}, \bar{y}(\bar{T})) = \theta$ , podemos descartá-lo devido as hipóteses do problema (2.2).

Tomando funções teste  $\eta \in C_0^1[a, \bar{T}]$  e  $\gamma \neq 0$ , temos

$$0 = \delta(I + \lambda G)(\bar{y}, \bar{T}; \eta, \gamma) = L(\bar{T}, \bar{y}, \bar{y}')\gamma + \lambda [\sigma_t(\bar{T}, \bar{y}) + \sigma_y(\bar{T}, \bar{y}) \bar{y}'(\bar{T})]\gamma. \quad (2.3)$$

Enquanto que escolhendo  $(\eta, \gamma)$  com  $\gamma = \eta(a) = 0$  e  $\eta(\bar{T}) \neq 0$ , obtemos

$$L_{y'}(\bar{T}, \bar{y}, \bar{y}') \eta(\bar{T}) + \lambda \sigma_y(\bar{T}, \bar{y}) \eta(\bar{T}) = 0,$$

i.e.  $\lambda = -L_{y'}(\bar{T}, \bar{y}, \bar{y}') / \sigma_y(\bar{T}, \bar{y})$ . Substituindo em (2.3) temos

$$L(\bar{T}, \bar{y}, \bar{y}') \sigma_y(\bar{T}, \bar{y}) = L_{y'}(\bar{T}, \bar{y}, \bar{y}') [\sigma_t(\bar{T}, \bar{y}) + \sigma_y(\bar{T}, \bar{y}) \bar{y}'(\bar{T})], \quad (2.4)$$

que é a condição natural procurada. A condição (2.4) é denominada **condição de transversalidade**. Podemos resumir a discussão acima na forma do seguinte teorema:

**Teorema 2.2.2** *Seja  $L \in C^1([a, b] \times \mathbb{R}^2; \mathbb{R})$ ,  $\sigma \in C^1([a, b] \times \mathbb{R}; \mathbb{R})$  com  $\nabla \sigma \neq \theta$ . Se  $(\bar{y}, \bar{T})$  é um mínimo local fraco de  $I$  em  $Y_{ad} \times \mathbb{R}$  restrito a  $\sigma(T, y(T)) = 0$ , então  $(\bar{y}, \bar{T})$  é solução do problema de valor de contorno*

$$\begin{cases} L_y(t, y, y') - \frac{d}{dt} L_{y'}(t, y, y') = 0, & t \in (a, T) \\ y(a) = y_a, & L(T, y, y') \sigma_y(T, y) = L_{y'}(T, y, y') [\sigma_t(T, y) + \sigma_y(T, y) y'(T)]. \end{cases}$$

**Demonstração:** Veja acima. ■

**Observação 2.2.3** Tomando  $\sigma(T, y) = t - b$  no Teorema 2.2.2, fixamos o tempo final  $T = b$  e deixamos livre a condição de contorno final. Podemos assim representar os *problemas com fronteira livre* vistos no Parágrafo 1.3. Neste caso a condição de transversalidade se escreve como

$$L_{y'}(b, y(b), y'(b)) = 0,$$

que é exatamente a condição de contorno natural obtida na Observação 1.3.2.

Tomando  $\sigma(T, y) = y - y_b$  no Teorema 2.2.2, deixamos o tempo final livre e fixamos a condição de contorno final  $y(T) = y_b$ . Obtemos assim os denominados de **problemas com horizonte livre**. Neste caso a condição de transversalidade toma a forma

$$L(T, y(T), y'(T)) = y'(T) L_{y'}(T, y(T), y'(T)).$$

□ □ □

**Exemplo 2.2.4 (Controle Ótimo a uma Curva Alvo)** Consideramos um problema de controle ótimo sujeito a uma condição de contorno transversal. Suponha que a condição inicial  $z(t_0) = z_0 \in \mathbb{R}^n$  é fornecida e que, no tempo final  $t = t_1$ , o estado final tenha que

pertencer a uma dada curva  $\mathcal{C}$ , a qual é descrita por  $\mathcal{C} := \{z \in \mathbb{R}^n \mid \sigma(z) = 0\}$ . Temos assim o seguinte problema de controle:

$$\begin{cases} \text{Minimizar } J(z, u) := \int_{t_0}^{t_1} L(t, z(t), u(t)) dt \\ \text{sujeito a } u \in \mathcal{U}_{ad} := \hat{C}([t_0, t_1]; \mathbb{R}^m), \\ z' = f(t, z, u), \quad t \in (t_0, t_1), \quad z(0) = z_0, \quad \sigma(z(1)) = 0, \end{cases}$$

com  $f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}$ . No caso particular  $m = n$ ,  $f(t, z, u) = u$ , temos o seguinte problema variacional correspondente:

$$\begin{cases} \text{Minimizar } I(y) := \int_{t_0}^{t_1} L(t, z(t), z'(t)) dt \\ \text{sujeito a} \\ y \in Y_{ad} := \{y \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n) \mid y(t_0) = z_0\}, \quad \sigma(y(t_1)) = 0. \end{cases}$$

Condições necessárias de otimalidade para o problema acima são dadas pela versão do Teorema 2.2.2 para problemas vetoriais (veja Parágrafo 1.6), na forma da equação vetorial de Euler-Lagrange e da condição de transversalidade

$$L(t_1, y, y') \nabla_y \sigma(t_1, y) = L_{y'}(t_1, y, y') [\sigma_t(t_1, y) + \langle \nabla_y \sigma(t_1, y), y' \rangle]$$

(Verifique!). Como  $\sigma = \sigma(y)$ , esta condição se reduz a

$$L(t_1, y, y') \nabla \sigma(y) = \langle \nabla \sigma(y), y' \rangle L_{y'}(t_1, y, y') = c L_{y'}(t_1, y, y').$$

Usando a equação de Euler-Lagrange obtemos

$$L(t_1, y, y') \nabla \sigma(y) = -c L_{y'}(t_1, y, y').$$

Definindo o multiplicador de Lagrange (Princípio do Máximo)  $\lambda(t) := -L_{y'}(t, y, y')$  temos

$$\lambda(t_1) = c \nabla \sigma(y(t_1)). \quad (2.5)$$

Portanto, a condição de transversalidade se traduz no fato do multiplicador de Lagrange ser perpendicular a curva  $\mathcal{C}$  no tempo final  $t_1$ . O problema de controle ótimo acima se inclui na classe de problemas abordada no Parágrafo 3.1. Compare a equação (2.5) com a condição de contorno no item *ii*) do Teorema 3.1.2 (tome  $p = 1$ ,  $L_1 \equiv 0$ ).

No caso escalar  $m = n = 1$ , o problema variacional se reduz a um *problema com horizonte livre* e a condição de transversalidade é dada na Observação 2.2.3.  $\square \square \square$

## Restrições Lagrangeanas

Obtemos agora condições necessárias para problemas variacionais com restrições lagrangeanas. Considere o problema

$$\left\{ \begin{array}{l} \text{Minimizar } I(y) := \int_a^b L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ y \in Y_{ad} := \{y \in C^1([a, b]; \mathbb{R}^n) \mid y(a) = y_a, y(b) = y_b\}; \\ G(t, y(t)) = 0, t \in [a, b]; \end{array} \right. \quad (2.6)$$

onde  $L : [a, b] \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$ ,  $G : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}$  e  $y_a, y_b \in \mathbb{R}^n$  são dados. Note que no caso escalar  $n = 1$ , o problema de minimização perde o sentido, pois na maioria dos casos a restrição  $G(t, y(t)) = 0$  determina  $y$  completamente.

**Teorema 2.2.5** *Sejam  $L \in C^2([a, b] \times \mathbb{R}^{2n}; \mathbb{R})$ ,  $G \in C^2([a, b] \times \mathbb{R}^n; \mathbb{R})$  com  $G_y(t, y) \neq \theta$ . Se  $\bar{y} \in Y_{ad}$  é um mínimo local fraco de  $I$  em  $Y_{ad}$  sujeito a restrição lagrangeana  $G(t, y(t)) = 0$ ,  $t \in [a, b]$ , então existe uma função (multiplicador)  $\lambda \in C[a, b]$  tal que*

$$\frac{d}{dt}[L + \lambda G]_{y'} = [L + \lambda G]_y, \quad t \in (a, b).$$

**Demonstração:** Seja  $\tau \in (a, b)$ . Como  $G_y(\tau, \bar{y}(\tau)) \neq \theta$ , então para pelo menos um índice  $1 \leq j \leq n$  temos  $G_{y_j}(\tau, \bar{y}(\tau)) \neq 0$ . Por conveniência representamos  $y \in \mathbb{R}^n$  por  $(y_j, Y)$ , onde  $y_j$  é a  $j$ -ésima componente de  $y$  e  $Y \in \mathbb{R}^{n-1}$  é o vetor contendo as demais componentes.<sup>4</sup> O teorema da função implícita (ver [Ru2] ou [Wa1]) garante a existência de uma função  $g \in C^2(\mathbb{R} \times \mathbb{R}^{n-1}, \mathbb{R})$  definida em uma vizinhança de  $(\tau, \bar{Y}(\tau))$  tal que  $G(t, y) = 0$  sse  $y_j = g(t, Y)$ . Sendo assim, podemos reescrever (localmente) a restrição do problema (2.6) como:  $\exists [c, d] \subset (a, b)$  com  $\tau \in (c, d)$  e  $y_j(t) = g(t, Y(t))$ ,  $t \in [c, d]$ .

Definindo o conjunto  $V := \{Y \in C^1([a, b]; \mathbb{R}^{n-1}) \mid Y(t) = \bar{Y}(t), t \notin (c, d)\}$ , podemos construir para cada  $Y \in V$  a função

$$y_j(t) := \begin{cases} g(t, Y(t)), & t \in (c, d) \\ \bar{y}_j(t), & t \notin (c, d) \end{cases}$$

obtendo assim  $y(t) := (y_j(t), Y(t)) \in Y_{ad}$  que obviamente satisfaz  $G(t, y(t)) = 0$ ,  $t \in [a, b]$ . Definimos agora o funcional

$$\begin{aligned} \tilde{I} : C^1([a, b]; \mathbb{R}^{n-1}) &\longrightarrow \mathbb{R} \\ Y(t) &\longmapsto I(g(\cdot, Y), Y) \end{aligned}$$

Por construção temos

$$\tilde{I}(Y) = \int_a^b L(t, y_j, Y, y'_j, Y') dt,$$

<sup>4</sup>Em particular temos  $\bar{y}(t) = (\bar{y}_j(t), \bar{Y}(t))$ .

onde  $y_j(t) = g(t, Y(t))$  e  $y'_j(t) = (d/dt)g(t, Y(t)) = g_t(t, Y) + g_Y(t, Y)Y'$ . Definindo  $\tilde{L}(t, Y, Y') := L(t, y_j, Y, y'_j, Y')$  podemos escrever

$$\tilde{I}(Y) = \int_a^b \tilde{L}(t, Y, Y') dt = \int_c^d \tilde{L}(t, Y, Y') dt + \text{const.}$$

para todo  $Y \in V$ . Como  $\bar{y}$  é mínimo local do problema (2.6), então  $\bar{Y}$  é mínimo local de  $\tilde{I}$  em  $V$  (por que?). Portanto  $\bar{Y}$  satisfaz

$$\frac{d}{dt} \tilde{L}_{Y'} = \tilde{L}_Y, \quad t \in (c, d). \quad (2.7)$$

Por construção temos  $\tilde{L}_{Y'} = L_{y'_j} g_Y + L_{Y'}$  e  $\tilde{L}_Y = L_{y_j} g_Y + L_Y + L_{y'_j} [g_t + g_Y Y']_Y$ . Substituindo em (2.7) temos

$$\frac{d}{dt} [L_{y'_j} g_Y + L_{Y'}] = L_{y_j} g_Y + L_Y + L_{y'_j} [g_t + g_Y Y']_Y, \quad t \in (c, d). \quad (2.8)$$

Usando agora a identidade  $[g_t + g_Y Y']_Y = (d/dt)g_Y$  obtemos de (2.8)

$$\frac{d}{dt} L_{Y'} - L_Y = \left[ -\frac{d}{dt} L_{y'_j} + L_{y_j} \right] g_Y, \quad t \in (c, d).$$

Como  $G(t, g(t, Y), Y) \equiv 0$  em uma vizinhança de  $(\tau, \bar{Y}(\tau))$ , temos que  $G_{y_j}(t, g(t, Y), Y)g_Y + G_Y(t, g(t, Y), Y) = 0$ . Logo, para qualquer função  $\lambda \in C[c, d]$  temos

$$\begin{aligned} \frac{d}{dt} [L + \lambda G]_{Y'} - [L + \lambda G]'_Y &= \frac{d}{dt} L_{Y'} - L_Y - \lambda G_Y \\ &= \frac{d}{dt} L_{Y'} - L_Y + \lambda G_{y_j} g_Y \\ &= \left[ -\frac{d}{dt} L_{y'_j} + L_{y_j} + \lambda G_{y_j} \right] g_Y. \end{aligned}$$

A escolha  $\lambda(t) = [(d/dt) L_{y'_j} - L_{y_j}] / G_{y_j} \in C[c, d]$  (lembre que  $G_{y_j} \neq 0$  em  $[c, d]$ ) nos fornece a equação

$$\frac{d}{dt} [L + \lambda G]_{Y'} - [L + \lambda G]'_Y = 0, \quad t \in (c, d). \quad (2.9)$$

Provamos assim que para essa escolha de multiplicador, a equação de Euler-Lagrange é satisfeita em uma vizinhança de  $t = \tau$ . Como  $\tau \in (a, b)$  é arbitrário, concluímos que o teorema é válido localmente no intervalo  $(a, b)$ .

Note que é possível *cobrir* o intervalo compacto  $[a, b]$  com uma família finita de sub-intervalos, tais que em cada um deles  $G_{y_j} \neq 0$  para algum  $j$ . A escolha do multiplicador  $\lambda$  em cada sub-intervalo depende apenas deste fato, o que justifica a conclusão de que é possível escolher  $\lambda \in C[a, b]$  de modo que (2.9) seja satisfeita em todo intervalo  $(a, b)$ . ■

**Observação 2.2.6** É possível formular o problema (2.6) com  $G = G(t, y, y')$  e, ainda assim, obter um resultado análogo ao apresentado no Teorema 2.2.5. A verificação deste resultado mais geral pode ser encontrada por exemplo em [Tr, Capítulo 11]. □ □ □

No Exemplo 2.2.9 é estudado um problema de controle ótimo com restrições lagrangeana e isoperimétrica.



**Restrições Isoperimétricas**

Tratamos a seguir dos problemas variacionais com restrições isoperimétricas (ou integrais). Considere o problema

$$\left\{ \begin{array}{l} \text{Minimizar } I(y) := \int_a^b L(t, y(t), y'(t)) dt \\ \text{sujeito a} \\ y \in Y_{ad} := \{y \in C^1[a, b] \mid y(a) = y_a, y(b) = y_b\}; \\ F(y) := \int_a^b G(t, y(t), y'(t)) dt = c; \end{array} \right. \quad (2.10)$$

onde  $L, G \in C^1([a, b] \times \mathbb{R}^2; \mathbb{R})$  e  $y_a, y_b, c \in \mathbb{R}$  são dados. Como consequência imediata do Teorema 2.2.1 (Lagrange), obtemos o seguinte resultado:

**Teorema 2.2.7** *Sejam  $L, G \in C^1([a, b] \times \mathbb{R}^2; \mathbb{R})$ ,  $F$  definida como em (2.10) e  $c \in \mathbb{R}$  tais que  $\delta F(y; \cdot) \neq 0$  no conjunto de nível  $c$  do funcional  $F$ , definido por*

$$F_c := \{y \in C^1[a, b] \mid F(y) = c\}.$$

*Se  $\bar{y}$  é um mínimo local fraco do problema (2.10), então existe um multiplicador  $\lambda \in \mathbb{R}$  tal que  $\bar{y}$  é solução do problema de valor de contorno*

$$\left\{ \begin{array}{l} (L + \lambda G)_y(t, y, y') - \frac{d}{dt}(L + \lambda G)_{y'}(t, y, y') = 0, \quad t \in (a, b) \\ y(a) = y_a, \quad y(b) = y_b. \end{array} \right.$$

**Demonstração:** É deixada como exercício para o leitor. ■

**Exemplo 2.2.8** Considere o problema de controle ótimo

$$\left\{ \begin{array}{l} \text{Minimizar } J(u, z) = \int_0^1 u(t)^2 dt \\ \text{sujeito a } u \in \hat{C}[0, 1], \\ z' = u, \quad z(0) = z_0, \quad z(1) = z_1, \quad F(u, z) = \int_0^1 tu(t) dt = c. \end{array} \right.$$

O problema variacional correspondente é

$$\left\{ \begin{array}{l} \text{Minimizar } I(y) = \int_0^1 y'(t)^2 dt \\ \text{sujeito a} \\ y \in Y_{ad} = \{y \in \hat{C}^1[0, 1] \mid y(0) = z_0, y(1) = z_1\} \int_0^1 ty'(t) dt = c. \end{array} \right.$$

Do Teorema 2.2.7, temos que se  $y$  é solução, então existe  $\lambda \in \mathbb{R}$  com

$$\frac{d}{dt}(2y'(t) + \lambda t) = 0.$$

Logo,  $2y'(t) + \lambda t = \text{const.}$  Temos então

$$y(t) = -\frac{1}{4}\lambda t^2 + at + b.$$

Usando as condições de contorno obtemos  $a$  e  $b$  em função de  $\lambda$  (além é claro de  $z_0$  e  $z_1$ , que são dados). Para determinar o valor de  $\lambda$ , basta substituir  $y' = a(\lambda) - \frac{1}{2}\lambda t$  na restrição isoperimétrica.

Em muitas aplicações as restrições isoperimétrica e/ou lagrangeana são dadas na forma de uma desigualdade. No exemplo a seguir estudamos tal situação.

**Exemplo 2.2.9 (Lançamento de um Foguete)** Um foguete deve ser lançado a partir do solo na vertical e deve, a partir da escolha de uma estratégia de consumo de combustível, alcançar a maior altitude possível. O modelo para o problema é representado por

- $t$  : tempo;
- $m(t)$  : massa do foguete;
- $v(t)$  : velocidade do foguete;
- $h(t)$  : altura do foguete.

Uma equação de balanço de forças (Taxa de variação do impulso = Soma das forças agindo sobre o corpo) nos fornece a dinâmica

$$m v' + \rho m' = -m G(h) - D(v, h),$$

onde  $\rho > 0$  é a velocidade com que o combustível, depois de queimado, é expelido,  $G$  é a força gravitacional (depende da altura),  $D$  é a resistência do ar (depende da velocidade e da altura).

A fim de simplificar o modelo supomos  $G(h) = g$  (constante) e desprezamos a resistência do ar ( $D(v, h) \equiv 0$ ). Logo,

$$m v' = -g m - \rho m'. \quad (2.11)$$

Suponha que podemos controlar o empuxo do foguete (dado por  $-\rho \dot{m}$ ). Note que, se denotamos por  $M_f > 0$  a massa do foguete sem combustível, temos que  $m(t) \geq M_f > 0$ . Portanto, escolhemos como variável de controle  $u := \rho m' / m$ , obtendo o sistema

$$\begin{cases} h' = v(t) & , h(0) = 0 \\ v' = u(t) - g, & v(0) = 0 \end{cases}$$

Supomos ainda o controle limitado, i.e.,  $0 \leq u(t) \leq u_{\max}$ . O problema exige que, dada uma quantidade fixa de combustível  $M_c$ , a estratégia de controle deve ser tal que o foguete alcance a maior altitude possível. Logo, temos as condições de contorno  $m(0) = M_f + M_c$  e  $m(T) = M_f$ . Note que a última condição de contorno pode ser escrita na forma da restrição isoperimétrica

$$\int_0^T u(t) dt = \rho \log[(M_f + M_c)/M_f] =: K.$$

A função objetivo é dada por

$$J(u, z) = h(T) = \int_0^T v(t) dt = \int_0^T (T - t)v'(t) dt = \int_0^T (T - t)u(t) dt - gT^2/2.$$

Definindo as variáveis de estado  $z_1 = h$ ,  $z_2 = v$ , podemos escrever o problema de controle ótimo na forma:

$$\left\{ \begin{array}{l} \text{Minimizar } J(u, z) = z_1(T) = \int_0^T (t - T)u(t) dt \\ \text{sujeito a } u \in \hat{C}[0, T], \\ z'_1 = z_2, z'_2 = u - g, z_1(0) = z_2(0) = 0, \\ F(u, z) = \int_0^T u(t) dt = K, 0 \leq u(t) \leq u_{max} \end{array} \right.$$

A restrição isoperimétrica juntamente com as condições de contorno  $z_2(0) = z_2(T) = 0$  nos permitem determinar o tempo final:  $T = K/g$ .

Observe que neste problema de controle temos, além da restrição isoperimétrica, uma restrição lagrangeana (na forma de desigualdade) para a variável de controle. Note ainda que tanto a função objetivo quanto as restrições dependem somente da variável de controle. Podemos escrever a restrição  $0 \leq u \leq u_{max}$  na forma de uma única restrição quadrática:  $u(u - u_{max}) \leq 0$ . A Função hamiltoniana se escreve então

$$H(t, u) = (t - T)u + \lambda u + \Lambda(t)(u^2 - u_{max}u),$$

com  $\lambda \in \mathbb{R}$  e  $\Lambda \in C[0, T]$ .

Se conseguirmos encontrar multiplicadores  $\lambda$ ,  $\Lambda$  e controle  $\bar{u}$  tais que

$$0 = \frac{d}{du}H(t, \bar{u}) = (t - T + \lambda) + \Lambda(t)(2\bar{u} - u_{max}),$$

$$\Lambda(t)\bar{u}(t)(\bar{u} - u_{max}) = 0, \text{ com } \Lambda(t) > 0.$$

e ainda  $\int_0^T \bar{u}(t)dt = K$ , então  $\bar{u}$  é solução (e é única, pois a hamiltoniana é estritamente convexa para  $\Lambda > 0$ ). Tal fato se justifica por estarmos trabalhando com uma restrição lagrangeana na forma de desigualdade (a segunda condição acima significa que o multiplicador  $\Lambda$  é positivo quando a restrição lagrangeana se torna ativa).

Uma análise de sinais nos permite concluir que  $\Lambda(t) = 0$  somente para  $t = \tau := T - \lambda$ . Para  $t \neq \tau$  temos  $\Lambda(t) > 0$ , o que implica em  $\bar{u}(t) = 0$  ou  $u_{max}$ . A fim de satisfazer a restrição isoperimétrica, escolhemos  $\lambda = T - K/u_{max} < T$ . Temos assim o controle ótimo

$$\bar{u}(t) := \begin{cases} u_{max} & , t \in [0, \tau) \\ 0 & , t \in (\tau, T] \end{cases}$$

do tipo **bang-bang** (i.e., assume apenas valores extremos).

□ □ □

### 2.3 Extremais Singulares e Trajetórias Ótimas

Neste parágrafo utilizamos o formalismo do cálculo variacional para obter, através da análise de condições necessárias, a solução para uma importante família de problemas de controle ótimo, a saber: problemas em que as variáveis de controle são limitadas por funções que dependem do tempo e da variável de estado.

Sob hipóteses adequadas é possível encontrar uma relação entre as soluções da equação de Euler–Lagrange (extremais) e as trajetórias ótimas do problema de controle correspondente. Começamos por analisar o seguinte problema variacional

$$\left\{ \begin{array}{l} \text{Minimizar } I(z) := \int_{t_0}^{t_1} [G(t, z(t)) + H(t, z(t))z'(t)] dt \\ \text{sujeito a} \\ A(t, z(t)) \leq z'(t) \leq B(t, z(t)), \quad t \in [t_0, t_1]; \\ z(t_0) = z_0, \quad z(t_1) = z_1. \end{array} \right. \quad (2.12)$$

Note que no problema (2.12) são feitas restrições sobre a derivada das funções admissíveis. Importante na análise a seguir é o fato de  $L(t, z, z')$  ser uma função afim na variável  $z'$ . Observe ainda que o problema (2.12) pode ser interpretado como decorrente do seguinte problema de controle ótimo

$$\left\{ \begin{array}{l} \text{Minimizar } I(z) := \int_{t_0}^{t_1} [L_1(t, z(t)) + L_2(t, z(t))u(t)] dt \\ \text{sujeito a} \\ z' = g(t, z) + f(t, z)u(t), \quad t \in [t_0, t_1]; \\ z(t_0) = z_0, \quad z(t_1) = z_1, \quad u(t) \in \Omega := [u_m, u_M]. \end{array} \right.$$

De fato, explicitando  $u$  em função de  $z$  e  $z'$  na equação diferencial (supondo  $f(t, z) \neq 0$ ) e substituindo no funcional  $I$  obtemos um problema como (2.12) onde

$$G(t, z) = L_1(t, z) - L_2(t, z) \frac{g(t, z)}{f(t, z)}; \quad H(t, z) = L_2(t, z) \frac{1}{f(t, z)};$$

$$A(t, z) = \min_{u(t) \in \Omega} \{g(t, z) + f(t, z)u\}; \quad B(t, z) = \max_{u(t) \in \Omega} \{g(t, z) + f(t, z)u\}.$$

Suponha agora as aplicações  $G, H, A, B : [t_0, t_1] \times \mathbb{R} \rightarrow \mathbb{R}$  continuamente diferenciáveis e defina o conjunto das trajetórias admissíveis

$$Z_{ad} := \{z \in \hat{C}^1[t_0, t_1] \mid A(t, z(t)) \leq z'(t) \leq B(t, z(t)), \quad t \in [t_0, t_1]; \quad z(t_0) = z_0, \quad z(t_1) = z_1\}.$$

Esquecendo por um momento as restrições sobre a derivada das funções admissíveis, o problema (2.12) pode ser escrito como um problema usual do cálculo das variações:

$$\left\{ \begin{array}{l} \text{Minimizar } I(z) := \int_{t_0}^{t_1} [G(t, z(t)) + H(t, z(t))z'(t)] dt \\ \text{sujeito a } z \in \hat{C}^1[t_0, t_1], \quad z(t_0) = z_0, \quad z(t_1) = z_1. \end{array} \right.$$

A equação de Euler–Lagrange associada a esse problema é

$$\frac{\partial G}{\partial z}(t, z(t)) - \frac{\partial H}{\partial t}(t, z(t)) = 0, \quad t \in (t_0, t_1). \tag{2.13}$$

Fazemos agora as seguintes hipóteses:

h1) Existe um **extremal singular**  $z^*$  para o problema variacional, isto é uma função  $z^* \in \hat{C}^1[t_0, t_1]$  que é solução da equação diferencial (2.13) e satisfaz simultaneamente a restrição  $A(t, z^*(t)) \leq z^{*\prime}(t) \leq B(t, z^*(t))$ ,  $t \in (t_0, t_1)$ .<sup>5</sup>

h2) O extremal singular  $z^*$  divide a faixa  $S := [t_0, t_1] \times \mathbb{R}$  em duas semi-faixas:

$$E_+ := \{(t, z) \in [t_0, t_1] \times \mathbb{R} \mid z > z^*(t)\}, \quad E_- := \{(t, z) \in [t_0, t_1] \times \mathbb{R} \mid z < z^*(t)\},$$

nas quais se verifica:

$$\frac{\partial G}{\partial z}(t, z) - \frac{\partial H}{\partial t}(t, z) > 0, \quad (t, z) \in E_+, \quad \frac{\partial G}{\partial z}(t, z) - \frac{\partial H}{\partial t}(t, z) < 0, \quad (t, z) \in E_-.$$

Como o funcional  $I$  é linear no argumento  $z'$ , é de se esperar que, nos pontos em que uma solução não coincide com o extremal singular, ela está na fronteira da admissibilidade, i.e.  $z'(t) = A(t, z(t))$  ou  $z'(t) = B(t, z(t))$ . A fim de formular mais precisamente esta idéia definimos a **função de identificação**  $\Sigma : S \rightarrow \mathbb{R}$

$$\Sigma(t, z) := \begin{cases} A(t, z), & (t, z) \in E_+ \\ z^*(t), & z = z^*(t) \\ B(t, z), & (t, z) \in E_- \end{cases} \tag{2.14}$$

Estamos agora em condições de formular um resultado que nos permite identificar as soluções do problema variacional (2.12) e do problema de controle ótimo correspondente.

**Teorema 2.3.1** *Suponha que  $G, H, A$  e  $B$  são funções continuamente diferenciáveis e que as condições (h1) e (h2) são satisfeitas. Seja  $\Sigma$  a função de identificação definida em (2.14). Se existir uma trajetória admissível  $z \in Z_{ad}$  que intercepta o extremal singular  $z^*$  em um ponto de  $(t_0, t_1)$ , então a função  $\bar{z}$  definida por*

$$\begin{cases} \bar{z}' = \Sigma(t, \bar{z}(t)), & t \in [t_0, t_1] \\ \bar{z}(t_0) = z_0, \quad \bar{z}(t_1) = z_1 \end{cases}$$

é uma trajetória ótima de (2.12).

**Demonstração:** Suponha  $z^*(0) < z_0$  (o caso  $z_0 < z^*(0)$  é análogo). Seja  $z$  uma trajetória admissível qualquer que intercepta  $z^*$  no ponto  $\tau \in (t_0, t_1)$ . Considere agora o problema de valor inicial

$$z' = A(t, z(t)), \quad z(t_0) = z_0.$$

<sup>5</sup>Note que  $z^*$  não pertence necessariamente a  $Z_{ad}$ .

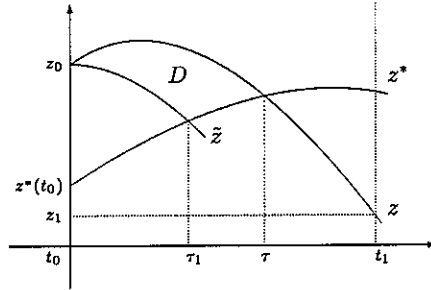


Figura 2.2: Trajetória admissível interceptando o extremal singular

Tal problema possui uma solução local, que denominamos  $\tilde{z}$ . Como

$$z'(t_0) \geq A(t_0, z(t_0)) = \tilde{z}'(t_0),$$

então  $z(t) \geq \tilde{z}(t)$  para  $t \in [t_0, t_0 + \varepsilon]$ . Por hipótese  $z(\tau) = z^*(\tau)$  então existe  $\tau_1 \in (t_0, \tau]$  tal que  $\tilde{z}(\tau_1) = z^*(\tau_1)$  (veja Figura 2.2). Note ainda que da desigualdade  $z^*(t_0) < z_0 = \tilde{z}(t_0)$  temos

$$\tilde{z}(t) > z^*(t), \quad t \in (t_0, \tau_1).$$

Defina agora a função  $\hat{z}$  por

$$\hat{z}(t) := \begin{cases} \tilde{z}(t) & , t \in [t_0, \tau_1] \\ z^*(t) & , t \in [\tau_1, \tau] \\ z(t) & , t \in (\tau, t_1] \end{cases}$$

que é admissível por construção. Seja  $D \subset E_+$  a região cuja fronteira  $\partial D$  é a curva de Jordan formada pelos segmentos

$$(t, \tilde{z}(t)), \quad t \in [t_0, \tau_1]; \quad (t, z^*(t)), \quad t \in [\tau_1, \tau]; \quad (t, z(t)), \quad t \in [t_0, \tau_1];$$

orientada no sentido anti-horário. Como  $\hat{z}(t) = z(t)$ ,  $t \geq \tau$ , temos

$$\begin{aligned} I(z) - I(\hat{z}) &= \int_{t_0}^{\tau} [G(t, z(t)) + H(t, z(t))z'(t)] dt - \int_{t_0}^{\tau} [G(t, \hat{z}(t)) + H(t, \hat{z}(t))\hat{z}'(t)] dt \\ &= \oint_{\partial D} [G(t, z) dt + H(t, z) dz] \\ &= \iint_D \left[ \frac{\partial G}{\partial z} - \frac{\partial H}{\partial t} \right] (t, z) dz dt \end{aligned}$$

(para obter a última igualdade utilizamos a fórmula de Green). Como a hipótese (h2) implica que  $[\frac{\partial G}{\partial z} - \frac{\partial H}{\partial t}](t, z) > 0$ , para  $(t, z) \in E_+$ , temos então

$$I(z) - I(\hat{z}) > 0,$$

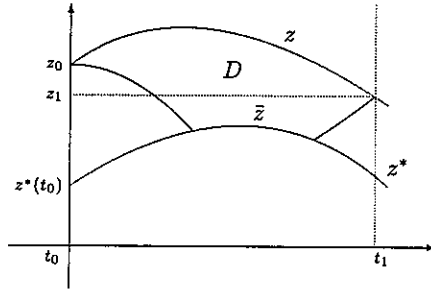


Figura 2.3: Trajetória admissível que não intercepta o extremal singular

quando  $\tau_1 < \tau$ . (Caso contrário obtemos apenas  $I(z) - I(\hat{z}) \geq 0$ .) Repetindo a argumentação na extremidade  $t_1$  do intervalo obtemos  $I(z) \geq I(\hat{z})$ .

Para completar a demonstração basta provar para toda trajetória admissível  $z$  que não intercepta  $z^*$  temos  $I(z) \geq I(\hat{z})$ . Como supomos  $z^*(0) < z_0$ , uma condição necessária para que exista uma trajetória admissível nessas condições é  $z^*(t_1) < z_1$ . Temos então

$$\begin{aligned} I(z) - I(\hat{z}) &= \int_{t_0}^{t_1} [G(t, z) + H(t, z)z'] dt - \int_{t_0}^{t_1} [G(t, z^*) + H(t, z^*)(z^*)'] dt \\ &= \iint_D \left[ \frac{\partial G}{\partial z} - \frac{\partial H}{\partial t} \right] (t, z) dz dt > 0, \end{aligned}$$

onde  $D$  é a região situada entre as curvas  $z$  e  $\hat{z}$ , conforme a Figura 2.3 ■

**Observação 2.3.2** Caso uma das condições de contorno não seja fornecida, por exemplo  $z(t_1) = z_1$ , obtemos em seu lugar a condição de contorno natural (veja Observação 1.3.2)

$$L_{y'}(t_1, z(t_1), z'(t_1)) = H(t_1, z(t_1)) = 0.$$

Neste caso, mesmo que a condição  $H(t_1, z(t_1)) = 0$  defina  $z(t_1)$  de forma única, é ainda possível aplicar o Teorema 2.3.1. □ □ □

**Observação 2.3.3** O Teorema 2.3.1 garante que a trajetória ótima para o problema é da forma **bang-singular-bang**.

No caso particular em que o extremal satisfaça  $z^*(t_0) = z_0$  ou  $z^*(t_1) = z_1$ , as trajetórias ótimas são do tipo **singular-bang** e **bang-singular** respectivamente. □ □ □

## 2.4 Convexidade I: condições suficientes

Continuamos neste parágrafo a analogia iniciada no Parágrafo 2.2 entre os problemas do cálculo variacional e os problemas de controle ótimo. Nosso objetivo é mostrar que, sob certas hipóteses de convexidade, a equação de Euler-Lagrange é uma condição suficiente de otimalidade para problemas que não possuem restrições nas variáveis de controle.

O leitor atento nota a semelhança entre o resultado principal deste parágrafo (Teorema 2.4.1) e o Teorema de condições suficientes 1.1.10. Começamos por considerar o problema de controle ótimo

$$\begin{cases} \text{Minimizar } I(z, u) := \int_{t_0}^{t_1} L(t, z(t), u(t)) dt \\ \text{sujeito a } z \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n); u \in \hat{C}([t_0, t_1]; \mathbb{R}^m); \\ z'(t) = f(t, z, u), t \in (t_0, t_1); \\ z(t_0) = z_0, z(t_1) = z_1; \end{cases} \quad (2.15)$$

onde os tempos inicial e final  $t_0, t_1$  são fixos,  $L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $z_0, z_1 \in \mathbb{R}^n$  são dados. O conjunto das trajetórias admissíveis é dado por

$$Z_{ad} := \{z \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n) \mid z(t_0) = z_0, z(t_1) = z_1\}.$$

Como não são feitas restrições as variáveis de controle em (2.15), o conjunto dos controles admissíveis é simplesmente  $U_{ad} := \hat{C}([t_0, t_1]; \mathbb{R}^m)$ .

Os candidatos a solução do problema de otimização restrita (2.15) são os processos admissíveis  $(z, u) \in Z_{ad} \times U_{ad}$ . Podemos então considerar o problema (2.15) como um problema variacional com restrições lagrangeanas

$$\begin{cases} \text{Minimizar } I(z, u) \\ \text{sujeito a } (z, u) \in Z_{ad} \times U_{ad} \\ G(t, z, u) = f(t, z, u) - z' = \theta, t \in (t_0, t_1); \end{cases} \quad (2.16)$$

Nossa experiência com problemas variacionais sujeitos a restrições lagrangeanas (veja Parágrafo 2.2) sugere que a solução de (2.16) está relacionada com a minimização do funcional estendido

$$\tilde{I}(z, u, \lambda) := \int_{t_0}^{t_1} \tilde{L}(t, z(t), u(t)) dt,$$

onde  $\tilde{L}(t, z, u) = L(t, z, u) + \lambda(t)G(t, z, u)$  e  $\lambda \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n)$  é um multiplicador. Definimos agora a função auxiliar

$$\begin{aligned} H : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (t, z, u, \lambda) &\longmapsto \langle \lambda, f(t, z, u) \rangle + L(t, z, u), \end{aligned}$$

denominada **função de Hamilton**. Por construção temos  $\tilde{L}(t, z, u) = H(t, z, u, \lambda) - \lambda(t)z'$ . A equação de Euler-Lagrange para o problema variacional irrestrito em  $\tilde{I}$  possui uma componente relativa a  $z$  e outra a  $u$ .<sup>6</sup> Temos assim

$$\frac{d}{dt}[L + \lambda G]_{z'} = [L + \lambda G]_z \quad \text{e} \quad \frac{d}{dt}[L + \lambda G]_{u'} = [L + \lambda G]_u.$$

<sup>6</sup>A componente da equação de Euler-Lagrange relativa ao multiplicador  $\lambda$  é simplesmente a restrição  $G(t, z, u) \equiv \theta$  do problema (2.16), como o leitor pode facilmente verificar.



Utilizando a definição de  $H$ , podemos reescrever as equações acima na forma

$$\lambda'(t) = -H_z(t, z, u, \lambda) \quad \text{e} \quad \theta = H_u(t, z, u, \lambda).$$

Note ainda que a equação de evolução do problema (2.15),  $z' = f(t, z, u)$ , (que corresponde a restrição lagrangeana do problema (2.16)) pode ser reescrita como  $z' = H_\lambda(t, z, u, \lambda)$ .

Observe que as condições de Weierstraß–Erdmann, que dizem respeito a continuidade das aplicações

$$\tilde{L}_{z'}(t, z, u) = -\lambda(t) \quad \text{e} \quad \tilde{L}_u(t, z, u) = \theta,$$

são trivialmente satisfeitas devido as hipóteses de regularidade do multiplicador  $\lambda$ .

A pergunta que investigamos neste parágrafo é a seguinte: As condições

$$z'(t) = H_\lambda(t, z, u, \lambda), \quad \lambda'(t) = -H_z(t, z, u, \lambda) \quad \text{e} \quad H_u(t, z, u, \lambda) = \theta$$

são suficientes para caracterizar um mínimo local (ou global) do problema (2.15)? A resposta é sim, desde que façamos hipóteses adequadas sobre a convexidade da função de Hamilton. No Parágrafo 2.5 investigamos a necessidade destas mesmas condições.

O desenvolvimento conduzido neste parágrafo utiliza apenas argumentos elementares de análise real, devido ao fato de supormos convexidade parcial da função de Hamilton, em relação as variáveis  $z$  e  $u$ . No capítulo seguinte analisamos o papel das condições acima em problemas de controle genéricos.

**Teorema 2.4.1** *Seja  $L \in C^1([t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m; \mathbb{R})$ ,  $f \in C^1([t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^n)$ . Suponha que a função de Hamilton  $H$  satisfaz*

$$H(t, z + v, u + w, \lambda) - H(t, z, u, \lambda) \geq H_z(t, z, u, \lambda) v + H_u(t, z, u, \lambda) w, \quad v \in \mathbb{R}^n, \quad w \in \mathbb{R}^m \quad (2.17)$$

em  $[t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ . São verdadeiras as afirmações:

a) Se  $(\bar{z}, \bar{u}, \bar{\lambda}) \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n) \times \hat{C}^1([t_0, t_1]; \mathbb{R}^m) \times \hat{C}^1([t_0, t_1]; \mathbb{R}^n)$  é solução de

$$\begin{cases} z'(t) = H_\lambda(t, z, u, \lambda), & \lambda'(t) = -H_z(t, z, u, \lambda), & H_u(t, z, u, \lambda) = \theta, & t \in (t_0, t_1), \\ z(t_0) = z_0, & z(t_1) = z_1, & & \end{cases} \quad (2.18)$$

então  $(\bar{z}, \bar{u})$  é mínimo global de  $I$  em  $Z_{ad} \times U_{ad}$  sujeito a restrição lagrangeana  $G(t, z, u) = f(t, z, u) - z' = \theta$ ,  $t \in [t_0, t_1]$ .

b) Caso a igualdade em (2.17) ocorra sse  $|v||w| = 0$ , então o processo  $(\bar{z}, \bar{u})$  no item (b) é o único mínimo global de  $I$  em  $Z_{ad} \times U_{ad}$ .

**Demonstração:** Da hipótese (2.18) segue imediatamente que  $(\bar{z}, \bar{u}) \in Z_{ad} \times U_{ad}$  e ainda que  $G(t, \bar{z}, \bar{u}) \equiv \theta$  em  $[t_0, t_1]$ . Seja  $(z, u) \in Z_{ad} \times U_{ad}$  um processo admissível satisfazendo a

restrição  $G(t, z, u) \equiv \theta$ . Temos então

$$\begin{aligned}
 I(z, u) - I(\bar{z}, \bar{u}) &= \tilde{I}(z, u, \bar{\lambda}) - \tilde{I}(\bar{z}, \bar{u}, \bar{\lambda}) \\
 &= \int_{t_0}^{t_1} [H(t, z, u, \bar{\lambda}) - H(t, \bar{z}, \bar{u}, \bar{\lambda}) - \bar{\lambda}(t)(z'(t) - \bar{z}'(t))] dt \\
 &\geq \int_{t_0}^{t_1} [H_z(t, \bar{z}, \bar{u}, \bar{\lambda})(z - \bar{z}) + H_u(t, \bar{z}, \bar{u}, \bar{\lambda})(u - \bar{u}) - \bar{\lambda}(z - \bar{z})'] dt \quad (2.19) \\
 &= \int_{t_0}^{t_1} [-\bar{\lambda}'(z - \bar{z})' - \bar{\lambda}(z - \bar{z})'] dt \\
 &= [-\bar{\lambda}(z - \bar{z})]_{t_0}^{t_1} = 0,
 \end{aligned}$$

provando o item (a). Para provar o item (b), basta observar que se  $(z, u) \neq (\bar{z}, \bar{u})$  é outro processo admissível satisfazendo a restrição  $G(t, z, u) \equiv \theta$ , então (2.19) é obtida com desigualdade estrita, provando assim que  $I(z, u) - I(\bar{z}, \bar{u}) > 0$ . ■

**Observação 2.4.2** Caso a condição de contorno  $z(t_1) = z_1$  não seja fornecida, o conjunto das trajetórias admissíveis correspondente é  $Z_{ad} := \{z \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n) \mid z(t_0) = z_0\}$  e o Teorema 2.4.1 continua válido se substituímos a condição de contorno  $z(t_1) = z_1$  em (2.18) pela condição *natural*  $\lambda(t_1) = \theta$ .

Analogamente, se a condição de contorno final para o problema (2.15) for da forma  $\sigma(z(t_1)) = \theta$ , onde a aplicação  $\sigma: \mathbb{R}^n \rightarrow \mathbb{R}^p$  satisfaz  $\nabla \sigma \neq \theta$ , o conjunto das trajetórias admissíveis correspondente é  $Z_{ad} := \{z \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n) \mid z(t_0) = z_0, \sigma(z(t_1)) = \theta\}$ .

Neste caso o Teorema 2.4.1 continua válido se substituímos a condição  $z(t_1) = z_1$  em (2.18) pela condição *natural* de transversalidade  $\lambda(t_1) = \Lambda \sigma_z(z(t_1))$  para  $\Lambda \in \mathbb{R}^p$  e pela exigência que  $\Lambda \sigma(z): \mathbb{R}^n \rightarrow \mathbb{R}$  seja uma função convexa. De fato, Se  $(z, u) \in Z_{ad} \times U_{ad}$  satisfaz  $G(t, z, u) \equiv \theta$ , definimos  $\tilde{I}(z, u, \lambda, \Lambda) := \tilde{I}(z, u, \lambda) + \Lambda \sigma(z(t_1))$  e obtemos

$$\begin{aligned}
 I(z, u) - I(\bar{z}, \bar{u}) &= \tilde{I}(z, u, \bar{\lambda}, \Lambda) - \tilde{I}(\bar{z}, \bar{u}, \bar{\lambda}, \Lambda) \\
 &\geq [-\bar{\lambda}(z - \bar{z})]_{t_0}^{t_1} + \Lambda \sigma(z(t_1)) - \Lambda \sigma(\bar{z}(t_1)) \\
 &\geq [-\bar{\lambda}(z - \bar{z})]_{t_0}^{t_1} + \Lambda \sigma_z(\bar{z}(t_1)) [z(t_1) - \bar{z}(t_1)] \\
 &= [-\bar{\lambda}(t_1) + \Lambda \sigma_z(\bar{z}(t_1))] [z(t_1) - \bar{z}(t_1)] = 0.
 \end{aligned}$$

Para obter a última das desigualdade acima usamos a convexidade de  $\Lambda \sigma(z)$ . Note que  $\Lambda \in \mathbb{R}^p$  é o multiplicador associado com a restrição  $\sigma(z(t_1)) = \theta$ . □ □ □

Das hipóteses do Teorema 2.4.1 concluímos que, fixados os valores de  $t, z, \lambda$ , a função  $H(t, z, \cdot, \lambda): \mathbb{R}^m \rightarrow \mathbb{R}$  é convexa, no sentido da Definição 1.1.8. De fato, tomando  $v = \theta$  em (2.17) temos

$$H(t, z, u + w, \lambda) - H(t, z, u, \lambda) \geq H_u(t, z, u, \lambda) w, \quad u, w \in \mathbb{R}^m.$$

Logo, uma consequência da condição  $H_u(t, \bar{z}(t), \bar{u}(t), \bar{\lambda}(t)) = \theta, t \in (t_0, t_1)$  em (2.18) é que  $\bar{u}(t)$  é um mínimo local de  $H(t, \bar{z}(t), \cdot, \bar{\lambda}(t))$  para cada  $t \in (t_0, t_1)$ , fato que pode ser escrito

na forma

$$H(t, \bar{z}(t), \bar{u}(t), \bar{\lambda}(t)) = \min_{u \in \mathbb{R}^m} \{H(t, \bar{z}(t), u, \bar{\lambda}(t))\}, \quad t \in (t_0, t_1). \quad (2.20)$$

A condição expressa em (2.20), que é conseqüência do pacote de condições suficientes (2.18), é extremamente importante na teoria de controle ótimo. No Parágrafo 2.5, ainda utilizando a hipótese de convexidade parcial da função de Hamilton, provamos que esta condição é também necessária para otimalidade de soluções do problema 2.15. A verificação da necessidade dessa condição no caso geral é tecnicamente mais complicada e é discutida no Capítulo 3.

A condição (2.20) é denominada **condição de máximo** ou de otimalidade, apesar de ser aqui apresentada na forma de mínimo. Isto, entretanto, se deve tão somente a natureza do problema de controle ótimo a ser estudado, cujo objetivo pode ser tanto maximizar quanto minimizar o funcional  $I$ .

## 2.5 Convexidade II: condições necessárias

Verificamos neste parágrafo que as condições propostas em (2.18) e em especial a condição de máximo (2.20) são, sob hipóteses adequadas, necessárias para otimalidade de soluções do problema de controle ótimo (2.15). Como nos problemas variacionais restritos, as condições necessárias são obtidas utilizando um teorema de multiplicadores.

Suponha que, fixados  $(t, z, \lambda) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ , a aplicação  $H(t, z, \cdot, \lambda) : \mathbb{R}^m \rightarrow \mathbb{R}$  é convexa (utilizamos a notação do Parágrafo 2.4). Seja  $(\bar{z}, \bar{u}) \in Y_{ad} \times U_{ad}$  um processo ótimo para o problema (2.15). O teorema de multiplicadores de Lagrange garante a existência de  $\bar{\lambda} \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n)$  tal que  $(\bar{z}, \bar{u}, \bar{\lambda})$  é mínimo local do funcional  $I$ .

Dada uma tripla qualquer de funções teste  $(\eta_1, \eta_2, \eta_3) \in \hat{C}_0^1 \times \hat{C}_0 \times \hat{C}_0^1$ , sabemos que a variação

$$\begin{aligned} \delta I(\bar{z}, \bar{u}, \bar{\lambda}; \eta_1, \eta_2, \eta_3) &= \int_{t_0}^{t_1} [\eta_1(L_z + \bar{\lambda}f_z) - \eta_1' \bar{\lambda} + \eta_2(L_u + \bar{\lambda}f_u) + \eta_3(f - \bar{z}')] dt \\ &= \int_{t_0}^{t_1} [\eta_1(L_z + \bar{\lambda}f_z + \bar{\lambda}') + \eta_2(L_u + \bar{\lambda}f_u) + \eta_3(f - \bar{z}')] dt \end{aligned}$$

sempre se anula. Sendo assim, escolhendo  $\eta_1 = \eta_2 \equiv \theta$  e  $\eta_3 \in \hat{C}_0^1[t_0, t_1]$  qualquer, obtemos a equação de estado

$$\bar{z}'(t) = f(t, \bar{z}, \bar{u}), \quad t \in (t_0, t_1).$$

Tomando  $\eta_1 \in \hat{C}_0[t_0, t_1]$  qualquer e  $\eta_2 \equiv \theta$ , obtemos a equação adjunta

$$\bar{\lambda}' = -L_z(t, \bar{z}, \bar{u}) - \bar{\lambda}f_z(t, \bar{z}, \bar{u}), \quad t \in (t_0, t_1).$$

Tomando agora  $\eta_2 \in \hat{C}_0[t_0, t_1]$  qualquer, obtemos a equação de otimalidade

$$\theta = L_u(t, \bar{z}, \bar{u}) + \bar{\lambda}f_u(t, \bar{z}, \bar{u}) = H_u(t, \bar{z}, \bar{u}, \bar{\lambda}), \quad t \in (t_0, t_1).$$

Como  $\bar{z}$  satisfaz por construção as condições de contorno do problema (2.15), concluímos que  $(\bar{z}, \bar{u}, \bar{\lambda})$  satisfazem (2.18). A condição de máximo (2.20) é portanto satisfeita para o controle ótimo  $\bar{u}$ .

Podemos resumir a argumentação acima na forma do seguinte teorema:

**Teorema 2.5.1 (Princípio do Máximo)** *Sejam  $L$  e  $f$  aplicações continuamente diferenciáveis,  $t_0 < t_1$  fixos,  $z_0, z_1 \in \mathbb{R}^n$  dados. Suponha ainda que a função de Hamilton  $H$  é convexa na variável  $u$ . Se  $(\bar{z}, \bar{u})$  é um mínimo local fraco de  $I$  em  $Y_{ad} \times U_{ad}$ , sujeito a restrição  $G(t, z, u) := f(t, z, u) - z' \equiv \theta$ , então existe um multiplicador  $\bar{\lambda} \in \hat{C}^1([t_0, t_1]; \mathbb{R}^n)$  tal que*

$$\begin{cases} z'(t) = H_\lambda(t, z, u, \lambda), \\ \lambda'(t) = -H_z(t, z, u, \lambda), \quad t \in (t_0, t_1), \\ z(t_0) = z_0, \quad z(t_1) = z_1, \end{cases}$$

e ainda, o controle ótimo  $\bar{u}$  satisfaz a condição de máximo

$$H(t, \bar{z}(t), \bar{u}(t), \bar{\lambda}(t)) = \min_{u \in \mathbb{R}^m} \{H(t, \bar{z}(t), u, \bar{\lambda}(t))\}, \quad t \in (t_0, t_1).$$

**Demonstração:** Veja acima. ■

## 2.6 Exercícios

2.1 Verifique os detalhes da demonstração do Teorema 2.2.7.

2.2 Considere o problema

$$\begin{cases} \text{Minimizar } I(z, u) := \int_0^1 (u(t) + 1) dt \\ \text{s.a.} \\ z' = -u, \quad t \in (0, 1); \\ z(0) = z(1) = 0. \end{cases}$$

Mostre que cada controle  $\hat{u} \in \hat{C}[0, 1]$  que satisfaz  $\int_0^1 u(t) dt = 0$ , é ótimo para o problema.

2.3 Encontre a solução do problema

$$\begin{cases} \text{Minimizar } I(z, u) := \int_0^1 [z'(t) - 2t + 1]^2 dt \\ \text{s.a.} \\ z' = u, \quad t \in (0, 1); \\ z(0) = z(1) = 0; \quad z(t) \geq 0. \end{cases}$$

2.4 Considere o problema escalar

$$\left\{ \begin{array}{l} \text{Minimizar } \int_0^1 [3z(t)^2 + u(t)^2] dt \\ \text{s.a.} \\ z' = -z + u, \quad t \in (0, 1), \quad z(0) = 1. \end{array} \right.$$

- a) Obtenha o problema variacional correspondente.
- b) Encontre a equação de Euler-Lagrange e as condições de contorno correspondentes.
- c) Determine o controle ótimo.



## Capítulo 3

# Princípio do Máximo de Pontryagin

Neste capítulo é analisado um conjunto de condições necessárias de otimalidade para problemas de controle ótimo. Tal resultado é conhecido na literatura como **princípio do máximo**<sup>1</sup> e pode ser interpretado como um teorema de multiplicadores de Lagrange em espaços de dimensão infinita.

No Parágrafo 2.5, sob hipóteses adicionais de convexidade, obtivemos uma versão simplificada do princípio do máximo utilizando argumentos elementares de análise. Como foi visto, a condição de máximo (2.20) pode ser obtida diretamente a partir da equação de Euler-Lagrange. No caso geral (sem a hipótese de convexidade da Hamiltoniana) a demonstração da necessidade da condição de máximo é mais complexa e necessita de resultados provenientes da teoria de otimização em espaços de dimensão infinita. consulte o Apêndice A). A autoria do princípio do máximo é controversa. A maioria dos autores credita o resultado ao grupo liderado pelo matemático russo L.S. Pontryagin (1956). Entretanto, tal condição pode ser encontrada em um texto anterior, porém pouco divulgado, de M.R. Hestenes (1950). Para maiores detalhes consulte [He1], [PBG], assim como as referências em [He2].

No Parágrafo 3.1 apresentamos o princípio do máximo para problemas de controle ótimo com horizonte finito. Algumas variantes do teorema, correspondentes a diferentes tipos de condições de contorno, são analisadas. No Parágrafo 3.2 consideramos o princípio do máximo para problemas com horizonte infinito. No Parágrafo 3.3 são discutidas condições necessárias para problemas impulsivos (que admitem trajetórias descontínuas). No Parágrafo 3.4 são discutidas diversas aplicações, onde o princípio do máximo é utilizado na identificação de processos ótimos.

### 3.1 Problemas com Horizonte Finito

Começamos por apresentar uma formulação bastante geral para os problemas de controle ótimo com horizonte finito. Considere o problema de controle

---

<sup>1</sup>Também conhecido como princípio do mínimo, ou princípio de Pontryagin.

$$P(t_0, z_0) \quad \left\{ \begin{array}{l} \text{Minimizar } J(z, u) := L_1(t_1, z(t_1)) + \int_{t_0}^{t_1} L(t, z(t), u(t)) dt \\ \text{sujeito a} \\ t_1 \geq t_0; u \in L^1([t_0, t_1]; \mathbb{R}^m), u(t) \in \Omega \text{ q.s. em } [t_0, t_1]; \\ z(t) = z_0 + \int_{t_0}^t f(s, z(s), u(s)) ds, t \in [t_0, t_1]; \psi(t_1, z(t_1)) = \theta \end{array} \right.$$

onde

$$\begin{aligned} L &: [t_0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}, \quad L_1 : [t_0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}, \\ f &: [t_0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad \psi : [t_0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^p \end{aligned}$$

e  $\Omega \subset \mathbb{R}^m$ . O tempo inicial  $t_0$  e a condição inicial  $z_0$  são fornecidos, enquanto que o tempo final  $t_1$  e/ou a condição final podem ser ou não conhecidos. Por analogia ao cálculo variacional, tal problema é denominado **problema de Bolza**. Caso a função objetivo seja da forma

$$J(z, u) := \int_{t_0}^{t_1} L(t, z(t), u(t)) dt,$$

o problema é denominado **problema de Lagrange**. Uma terceira variante são os **problemas de Mayer**, em que a função objetivo é da forma  $J(z, u) := L_1(t_1, z(t_1))$ . A denominação aqui empregada corresponde a utilizada no clássico cálculo de variações. As três formulações acima são equivalentes, no sentido de que um mesmo problema de controle ótimo pode sempre ser colocado em qualquer uma das formas acima.

Note que o fato da dinâmica do sistema ser descrita por uma equação integral, ao invés de diferencial, nos permite considerar trajetórias admissíveis menos regulares. O conjunto das estratégias de controle admissíveis é

$$\mathcal{U}_{ad} := L^1_{loc}([0, \infty); \mathbb{R}^m),$$

de forma que as trajetórias correspondentes são funções absolutamente contínuas em  $[t_0, t_1]$ .

**Definição 3.1.1** Sejam  $f, L$  as funções definidas acima. A aplicação

$$\begin{aligned} H &: [t_0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}, \\ (t, z, \lambda, u) &\mapsto \langle \lambda, f(t, z, u) \rangle + \eta L(t, z, u). \end{aligned}$$

é denominada **função de Hamilton** (note que a origem da constante  $\eta$  precisa ainda ser esclarecida). □ □ □

Os problemas  $P(t_0, z_0)$  podem ser divididos em dois tipos: problemas com tempo final fixo ( $t_1 = T$  conhecido) ou com tempo final desconhecido. A demonstração do princípio do máximo para os problemas do primeiro tipo é ligeiramente mais simples e pode ser encontrada e.g. em [Tr, Teorema 11.8] (somente caso autônomo). Uma demonstração bastante compreensível para problemas de tempo ótimo (i.e., do tipo  $J(z, u) = \int_{t_0}^{t_1} dt$ ) autônomos com dinâmica linear pode ainda ser encontrada em [Ho, Capítulo 9].



No teorema a seguir apresentamos o **princípio do máximo**. Por ser longa e técnica, a demonstração é deixada para o Capítulo 4. A argumentação utilizada na demonstração segue a linha das notas de aula de M. Brokate (veja [Br]). Uma comparação com a demonstração proposta por W. Fleming e R. Rishel (veja [FIRi, Capítulo 2]) é inevitável: os autores também utilizam um problema (auxiliar) abstrato de otimização para obter as condições necessárias para otimalidade. Como referências auxiliares o leitor pode consultar as referências clássicas [PBG], [He1], [Ber], [Bo] ou ainda os textos modernos [FIRi], [Know], [LiYo], [MaSt], [Tr], [Za].

**Teorema 3.1.2** *Suponha que  $L, f$  são aplicações  $C^2$  e que  $L_1, \psi$  são  $C^1$ . Se  $(z^*, u^*, t_1^*)$  é uma solução do problema  $P(t_0, z_0)$  com  $\psi_z(t_1^*, z^*(t_1^*)) \neq \theta$  e  $(L_1)_z(t_1^*, z^*(t_1^*)) \neq \theta$ , então existe uma função  $\lambda : [t_0, t_1^*] \rightarrow \mathbb{R}^n$  e constantes  $\eta \geq 0, \mu \in \mathbb{R}^p$  que satisfazem:*

i) *Equação de estado*

$$z^*(t) = z_0 + \int_{t_0}^t f(s, z^*(s), u^*(s)) ds, \quad t \in [t_0, t_1^*];$$

ii) *Equação adjunta*

$$\begin{cases} \lambda(t) = \lambda_1 + \int_t^{t_1^*} \frac{\partial H}{\partial z}(s, z^*(s), \lambda(s), u^*(s)) ds, \quad t \in [t_0, t_1^*], \\ \lambda_1 := \eta \frac{\partial L_1}{\partial z}(t_1^*, z^*(t_1^*)) - \frac{\partial \psi}{\partial z}(t_1^*, z^*(t_1^*)) \mu; \end{cases}$$

iii) *Equação de evolução da função de Hamilton*

$$\begin{cases} H(t, z^*(t), \lambda(t), u^*(t)) = H_1 - \int_t^{t_1^*} \frac{\partial H}{\partial t}(s, z^*(s), \lambda(s), u^*(s)) ds, \quad t \in [t_0, t_1^*], \\ H_1 := -\eta \frac{\partial L_1}{\partial t}(t_1^*, z^*(t_1^*)) + \left\langle \frac{\partial \psi}{\partial t}(t_1^*, z^*(t_1^*)), \mu \right\rangle; \end{cases}$$

iv) *Condição de otimalidade*

$$H(t, z^*(t), \lambda(t), u^*(t)) = \min_{u \in \Omega} H(t, z^*(t), \lambda(t), u), \quad q.s. \text{ em } [t_0, t_1^*];$$

v) *Condição de não acoplamento  $\eta + |\mu| \neq 0$ .*

**Demonstração:** Veja Capítulo 4. ■

A demonstração do Teorema 3.1.2 se constitui na aplicação de um teorema de multiplicadores a um problema auxiliar, obtido de  $P(t_0, z_0)$  por uma mudança de variáveis denominada transformação no tempo e cuja solubilidade está relacionada à de  $P(t_0, z_0)$ . As constantes  $\eta$  e  $\mu$  surgem na demonstração como componentes de um vetor normal a um hiperplano, que separa conjuntos de nível associados a função objetivo  $J$  e a condição final  $\psi(t, z(t)) = \theta$ .

**Observação 3.1.3** A denominação princípio do máximo é motivada pelo item iv) do Teorema 3.1.2, que, entretanto, se refere à determinação de um mínimo. Este fato se deve à escolha do sinal do multiplicador de Lagrange  $\eta \geq 0$ . Tal escolha é arbitrária e a formulação do teorema com  $\eta \leq 0$  altera o min da condição iv) para um max. Sem dúvida, as condições mais interessantes do Teorema 3.1.2 são

$$\bullet \begin{cases} z' = H_\lambda(t, z, \lambda, u), & z(t_0) = z_0, \psi(t_1, z(t_1)) = 0, \\ \lambda' = -H_z(t, z, \lambda, u), & \lambda(t_1) = \eta \frac{\partial L_1}{\partial z}(t_1, z(t_1)) - \frac{\partial \psi}{\partial z}(t_1, z(t_1))^* \mu; \end{cases}$$

$$\bullet H(t, z, \lambda, u) = \min_{w \in \Omega} H(t, z(t), \lambda(t), w), \text{ q.s. em } [t_0, t_1].$$

Note que o par  $(z, \lambda)$  é solução de um sistema Hamiltoniano para a função  $H$ .

Os problemas em que o multiplicador escalar  $\eta$  assume o valor zero são denominados anormais. Neste caso o controle ótimo fornecido pela condição de máximo não depende da função objetivo, mas somente das restrições do problema de otimização.  $\square \square \square$

**Observação 3.1.4** Analogamente aos problemas variacionais, os problemas de controle ótimo também podem ser formulados com diferentes tipos de condições de contorno. A cada um destes tipos corresponde uma variante do Teorema 3.1.2, que se diferencia deste apenas pelas condições de contorno para a variável adjunta e para a função de Hamilton. Enunciamos a seguir algumas variantes do problema  $P(t_0, z_0)$  que surgem com maior frequência em aplicações. Apresentamos ainda as condições de contorno correspondentes (condições naturais de contorno).

Considere o problema  $P(t_0, z_0)$  com  $t_1$  fixo ( $t_1 > t_0$ ) e  $L_1 \equiv \theta$  (problema de Lagrange).

- Se a condição final é fixada ( $z(t_1) = z_1$ ) não há nenhuma condição para  $\lambda(t_1)$  (corresponde a escolha  $\psi(t, z) := (t - t_1, z - z_1)^* \in \mathbb{R}^2$ ).
- Se a condição final é livre ( $z(t_1)$  qualquer) a variável adjunta satisfaz  $\lambda(t_1) = \theta$  (corresponde a escolha  $\psi(t, z) := t - t_1 \in \mathbb{R}$ ).
- Se a condição final é da forma:  $z(t_1) \geq z_1$  (no caso escalar) a variável adjunta satisfaz:  $\lambda(t_1) \geq 0$  e a igualdade ocorre quando  $z(t_1) > z_1$ .

Considere agora o problema  $P(t_0, z_0)$  com  $t_1$  livre e  $L_1 \equiv \theta$ . Neste caso as condições para  $\lambda(t_1)$  discutidas acima não se alteram e, além disso,

$$H(t_1, z(t_1), \lambda(t_1), u(t_1)) = 0.$$

Esta equação extra corresponde a variável adicional do problema: o tempo final desconhecido  $t_1 \in (t_0, \infty)$ .  $\square \square \square$

O princípio do máximo pode, em alguns casos, ser utilizado para efetivamente determinar uma solução do problema  $P(t_0, z_0)$ . Para tanto, aplica-se a seguinte estratégia: Inicialmente a condição de máximo é usada para explicitar o controle ótimo  $u$  em função das variáveis  $z$  e  $\lambda$ . Obtemos assim

$$u(\cdot) = U_*(\cdot, z(\cdot), \lambda(\cdot)).$$

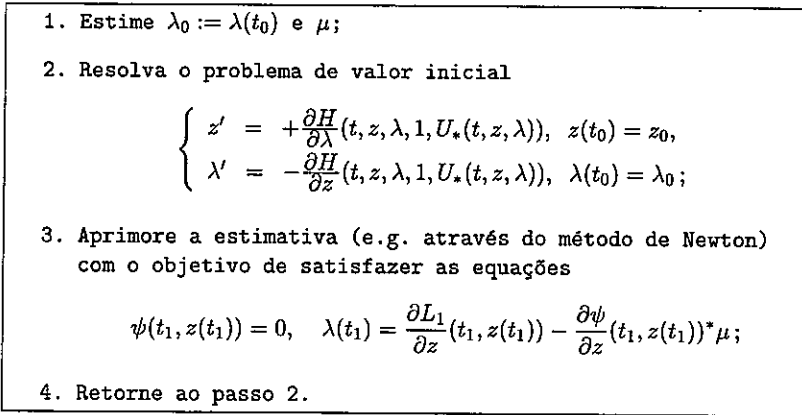


Figura 3.1: Algoritmo do método de shooting para resolver um sistema Hamiltoniano.

O passo seguinte consiste em substituir esta expressão no sistema Hamiltoniano (veja Observação 3.1.3), eliminando deste a variável  $u$ . Obtemos assim

$$z' = H_\lambda(t, z, \lambda, U_*), \quad z(t_0) = z_0, \quad \psi(t_1, z(t_1)) = 0; \quad (3.1)$$

$$\lambda' = -H_z(t, z, \lambda, U_*), \quad \lambda(t_1) = \eta \frac{\partial L_1}{\partial z}(t_1, z(t_1)) - \frac{\partial \psi}{\partial z}(t_1, z(t_1))^* \mu; \quad (3.2)$$

$$H' = \frac{\partial H}{\partial t}(t, z, \lambda, U_*), \quad (3.3)$$

$$H(t_1, z(t_1), \lambda(t_1), U_*(t_1)) = -\eta \frac{\partial L_1}{\partial t}(t_1, z(t_1)) + \left\langle \frac{\partial \psi}{\partial t}(t_1, z(t_1)), \mu \right\rangle$$

Resolvendo este sistema, obtemos um candidato a solução do problema  $P(t_0, z_0)$ .

Nem sempre é viável aplicar tal abordagem, pois em certos problemas não é possível obter a representação  $U_*(\cdot, z, \lambda)$  para o controle ótimo (e.g., o extremal singular no Parágrafo 2.3).

O sistema resultante de (3.1), (3.2), (3.3) pela substituição  $u(\cdot) = U_*(\cdot, z, \lambda)$  pode ser resolvido por um método do tipo shooting, conforme o esquema na Figura 3.1 (neste esquema aplicamos o Teorema 3.1.2 com  $\eta = 1$ ).

**Observação 3.1.5** O problema de valor de contorno (3.1), (3.2), (3.3) possui, tomando o controle  $u$  fixo,  $2n + 1$  variáveis  $(z, \lambda, H)$  e  $p + 1$  parâmetros  $(\eta, \mu)$ . Temos assim  $2n + p + 2$  graus de liberdade, os quais estão sujeitos a  $2n + p + 1$  equações. Aparentemente temos um grau de liberdade a mais. Note porém que a condição de não acoplamento  $v$  garante que  $\eta$  e  $\mu$  não são ambos nulos. Logo, é sempre possível simplificar o sistema (3.1), (3.2), (3.3) em relação a  $\eta$  ou a uma das componentes de  $\mu$ . Sendo assim, o Teorema 3.1.2 pode ser formulado alternativamente como:

... existem  $\lambda : [t_0, t_1^*] \rightarrow \mathbb{R}^n$ ,  $\eta = 0$  ou  $\eta = 1$ ,  $\mu \in \mathbb{R}^p$  que satisfazem  $i)$ , ... ,  $v)$ .

□ □ □

**Observação 3.1.6** É simples verificar que a equação de Euler-Lagrange do cálculo variacional pode ser obtida a partir do princípio do máximo. De fato, o problema variacional de minimização (1.2) pode ser interpretado como

$$\begin{cases} \text{Minimizar } J(z, u) := \int_a^b L(t, z(t), u(t)) dt \\ \text{sujeito a } z' = u(t). \end{cases}$$

Logo, a condição de máximo *iv)* do Teorema 3.1.2 implica em

$$0 = \frac{\partial H}{\partial u}(t, z^*, \lambda^*, u^*) = \frac{\partial}{\partial u}[(\lambda^*, u^*) + \eta L(t, z^*, u^*)] = \lambda^* + \eta \frac{\partial L}{\partial u}(t, z^*, u^*)$$

e, portanto,

$$\lambda^* = -\eta \frac{\partial L}{\partial u}(t, z^*, u^*). \quad (3.4)$$

O sistema Hamiltoniano para as variáveis de estado e adjunta se escreve

$$\begin{cases} \frac{dz^*}{dt} = +\frac{\partial H}{\partial \lambda} = u^* \\ \frac{d\lambda^*}{dt} = -\frac{\partial H}{\partial z} = -\eta \frac{\partial L}{\partial z}(t, z^*, u^*) \end{cases} \quad (3.5)$$

De (3.4) e (3.5) temos

$$-\eta \frac{\partial L}{\partial z}(t, z^*, u^*) = \frac{d}{dt} \left( -\eta \frac{\partial L}{\partial u}(t, z^*, u^*) \right)$$

ou

$$\frac{\partial L}{\partial z}(t, z^*, (z^*)') - \frac{d}{dt} \left( \frac{\partial L}{\partial z'}(t, z^*, (z^*)') \right) = 0,$$

que é a equação de Euler-Lagrange. □ □ □

## 3.2 Problemas com Horizonte Infinito

Analizamos neste parágrafo condições necessárias para problemas de controle ótimo com horizonte infinito. Os problemas de controle desta natureza tem ganho importância nas últimas décadas devido aos modelos de problemas oriundos das ciências econômicas e biológicas que os utilizam.

Os primeiros trabalhos a tratar de problemas de controle com horizonte infinito são devidos aos economistas. Uma referência clássica é o artigo escrito em 1928 por F. Ramsey (veja [Ra]), que discute modelos para problemas do tipo *consumo*  $\times$  *investimento* (veja Aplicação 3.4.3). A primeira extensão do princípio do máximo para problemas com horizonte infinito foi apresentada por H. Halkin em 1964 (veja [Ha]); todavia, a caracterização

de otimalidade proposta neste trabalho é deveras restritiva. Um detalhado resumo do desenvolvimento da teoria pode ser encontrado em [CaHa]. Uma abordagem mais recente é encontrada em [BaCa]; entretanto, somente problemas autônomos são considerados. Na abordagem aqui apresentada, seguimos os passos descritos em [Leit], que trata problemas com horizonte infinito não autônomos e utiliza um conceito de otimalidade (diferente dos encontrados em [Ha] e [CaHa]) que estende de forma natural o conceito usado em problemas com horizonte finito.

Inúmeros modelos econômicos (de horizonte finito e infinito) são tratados, sob a ótica do controle ótimo e programação dinâmica, em [SeSy]. Entre outros, são analisados problemas de exploração de recursos naturais. O leitor interessado em aplicações desta natureza deve consultar ainda [SeZh]. Aplicações a modelos biológicos podem ser encontradas em [Cl]. Um interessante problema relacionado com a exploração ótima de recursos naturais renováveis (pescaria ótima) é analisado em [CCM] e [BaLe] (veja Aplicação 3.4.4).

Antes de discutirmos os problemas com horizonte infinito, analisamos um problema auxiliar com horizonte finito. Trata-se de um caso particular do problema abordado no Parágrafo 3.1, a saber, problemas de controle ótimo com tempo final fixo. As condições necessárias para este problema auxiliar são utilizadas na demonstração do princípio do máximo para os problemas com horizonte infinito.

Suponha que no problema  $P(t_0, z_0)$ , o tempo final  $t_1 = T$  e o estado final  $z_1 = z_T$  são dados. Temos assim o seguinte problema de controle ótimo:

$$P_T(z_0) \quad \begin{cases} \text{Minimizar } J(z, u) := \int_0^T L(t, z(t), u(t)) dt \\ \text{sujeito a} \\ z(t) = z_0 + \int_0^t f(s, z(s), u(s)) ds, \quad t \in [0, T], \quad z(T) = z_T; \\ u \in L^1([0, T]; \mathbb{R}^m), \quad u(t) \in \Omega \text{ q.s. em } [0, T]; \end{cases}$$

Argumentando como na Observação 3.1.4, obtemos o seguinte conjunto de condições necessárias para otimalidade de uma solução do problema  $P_T(z_0)$ :

**Corolário 3.2.1** *Suponha que  $L, f$  são aplicações  $C^2$ . Se  $(z^*, u^*)$  é uma solução do problema  $P_T(z_0)$ , então existe uma função  $\lambda: [0, T] \rightarrow \mathbb{R}^n$  e  $\eta = 1$  ou  $\eta = 0$  que satisfazem*

i) *Sistema Hamiltoniano*

$$\begin{cases} (z^*)'(t) = H_\lambda(t, z^*, \lambda, u^*), \quad \text{q.s. em } [0, T], \\ \lambda'(t) = H_z(t, z^*, \lambda, u^*), \quad \text{q.s. em } [0, T], \\ z^*(t_0) = z_0, \quad z^*(T) = z_T; \end{cases}$$

ii) *Condição de otimalidade*

$$H(t, z^*(t), \lambda(t), u^*(t)) = \min_{u \in \Omega} H(t, z^*(t), \lambda(t), u), \quad \text{q.s. em } [0, T];$$

iii) *Condição de não acoplamento  $\eta + \|\lambda\|_\infty \neq 0$ .*

**Demonstração:** Segue imediatamente do Teorema 3.1.2 e da Observação 3.1.4. ■

Analisamos agora os problemas com horizonte infinito. Considere o seguinte problema de controle ótimo:

$$P_{\infty}(z_0) \quad \begin{cases} \text{Minimizar } J(z, u) := \int_0^{\infty} e^{-\delta t} L(z(t), u(t)) dt \\ \text{sujeito a} \\ z(t) = z_0 + \int_0^t f(z(s), u(s)) ds, \quad t \in [0, \infty); \\ u \in L^1_{loc}([0, \infty); \mathbb{R}^m), \quad u(t) \in \Omega \text{ q.s. em } [0, \infty); \end{cases}$$

onde as funções  $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ , o conjunto  $\Omega \in \mathbb{R}^m$  e a constante  $\delta > 0$  são dados. O resultado a seguir fornece condições necessárias para otimalidade de uma solução de  $P_{\infty}(z_0)$ .

**Teorema 3.2.2** *Suponha que  $L, f$  são aplicações  $C^2$ . Se  $(z^*, u^*)$  é uma solução de  $P_{\infty}(z_0)$ , então existe uma aplicação  $\lambda : [0, \infty) \rightarrow \mathbb{R}^n$  e constantes  $\eta = 0$  ou  $\eta = 1$ ,  $\lambda_0 \in \mathbb{R}^n$  que satisfazem:*

i) *Equação de estado*

$$z^*(t) = z_0 + \int_0^t f(s, z^*(s), u^*(s)) ds, \quad t \in [0, \infty);$$

ii) *Equação adjunta*

$$\lambda(t) = \lambda_0 + \int_0^t \frac{\partial H}{\partial z}(s, z^*(s), \lambda(s), u^*(s)) ds, \quad t \in [0, \infty);$$

iii) *Condição de otimalidade*

$$H(t, z^*(t), \lambda(t), u^*(t)) = \min_{u \in \Omega} H(t, z^*(t), \lambda(t), u), \quad \text{q.s. em } [0, \infty).$$

**Demonstração:** Seja  $(z^*, u^*)$  um processo ótimo para  $P_{\infty}(z_0)$ . Dado  $T > 0$ , as funções  $e^{-\delta t} L : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$  e  $f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$  satisfazem as condições do Corolário 3.2.1. Logo, obtemos deste corolário condições necessárias para otimalidade de cada um dos problemas

$$P_{T_k}(z_0) \quad \begin{cases} \text{Minimizar } J(z, u) := \int_0^{T_k} e^{-\delta t} L(t, z(t), u(t)) dt \\ \text{sujeito a} \\ z(t) = z_0 + \int_0^t f(s, z(s), u(s)) ds, \quad t \in [0, T_k], \quad z(T_k) = z_k := z^*(T_k); \\ u \in L^1([0, T_k]; \mathbb{R}^m), \quad u(t) \in \Omega \text{ q.s. em } [0, T_k]; \end{cases}$$

onde  $T_k \rightarrow \infty$ . Como consequência do princípio de otimalidade de Bellman (veja [Leit]), temos que  $(z^*, u^*)|_{[0, T_k]}$  é uma solução de  $P_{T_k}(z_0)$ . Juntando os fatos, podemos garantir a existência de  $\eta_k \geq 0$ ,  $\lambda_k : [0, T_k] \rightarrow \mathbb{R}^n$  tal que

- $\|\lambda_k\|_\infty + \eta_k > 0$ ;
- $(z^*, \lambda_k)$  é solução do sistema Hamiltoniano<sup>2</sup>

$$\begin{aligned} dz^*(t) &= H_\lambda(t, z^*(t), \lambda_k(t), u^*(t)) dt, \quad t \in [0, T_k] \\ d\lambda_k(t) &= -H_x(t, z^*(t), \lambda_k(t), u^*(t)) dt, \quad t \in [0, T_k] \\ z^*(0) &= z_0, \quad z^*(T_k) = z_k; \end{aligned}$$

- $H(t, z^*(t), \lambda_k(t), u^*(t)) = \max_{u \in \Omega} \{H(t, z^*(t), \lambda_k(t), u)\}$ , q.s. in  $[0, T_k]$ .

Considere agora a sequência  $\{\lambda_k(0), \eta_k\}_{k \in \mathbb{N}}$ . Normalizando os multiplicadores de Lagrange, podemos supor que  $|\lambda_k(0)| + \eta_k = 1$ ,  $k \in \mathbb{N}$ . Tomando subsequências (se necessário) podemos garantir a existência de  $\lambda_0 \in \mathbb{R}^n$  e  $\eta \geq 0$  satisfazendo

$$|\lambda_0| + \eta = 1, \quad \lim_k \lambda_k(0) = \lambda_0, \quad \lim_k \eta_k = \eta. \quad (3.6)$$

Seja agora  $T > 0$  fixo. Logo  $T_k > T$ , para  $k > k_0$  e como o sistema Hamiltoniano desfruta da propriedade de dependência contínua das condições iniciais, podemos garantir que existe  $\lambda : [0, T] \rightarrow \mathbb{R}^n$ , tal que  $\lambda_k$  converge uniformemente para  $\lambda$  em  $[0, T]$ . Essa convergência implica nas desejadas condições de otimalidade para o problema  $P_\infty(x_0)$ , uma vez que  $T > 0$  é arbitrário. ■

Duas diferenças básicas devem ser observadas na formulação do Teorema 3.2.2 em relação ao Teorema 3.1.2: *i)* Falta uma condição de contorno final para a variável adjunta  $\lambda$  (eventualmente uma condição de decaimento do tipo  $\lim_{t \rightarrow \infty} \lambda(t) = \theta$ ); *ii)* Falta a condição de não acoplamento (neste caso  $\eta + |\lambda_0| \neq 0$ ).

**Observação 3.2.3** Uma análise dos problemas (com horizonte infinito) com dinâmica linear e custo quadrático pode ser feita, alternativamente, utilizando-se programação dinâmica. Através de um processo de limite, é possível obter a função valor resolvendo-se a equação algébrica de Riccati (veja [So, Capítulo 7]). Uma vez conhecida a função valor, a variável adjunta<sup>3</sup> pode ser calculada ( $\lambda(t) = \frac{\partial}{\partial x} V(t, x)$ ) e o controle ótimo é obtido através da condição de otimalidade. □ □ □

### 3.3 Problemas Impulsivos

Apresentamos neste parágrafo uma variante do princípio do máximo para problemas com controle impulsivos. Neste tipo de problema, são permitidos saltos (discontinuidades) nas trajetórias admissíveis. Considere o problema impulsivo de controle ótimo

<sup>2</sup>Note que  $H(t, z, \lambda_k, u) = \langle \lambda_k, F(t, z, u) \rangle + \eta_k L(t, z, u)$ .

<sup>3</sup>A variável adjunta corresponde à derivada parcial da função valor em relação a variável de estado.

$$PI_{\infty}(x_0) \begin{cases} \text{Minimizar } J(z, u, \mu) := \int_0^{\infty} e^{-\delta t} f_0(t, z(t), u(t)) dt + \int_{[0, \infty)} e^{-\delta t} g_0(t) \mu(dt) \\ \text{sujeito a} \\ dz(t) = f(t, z(t), u(t)) dt + g(t) \mu(dt), \quad t \in [0, \infty); \\ z(0) = z_0; \quad u(t) \in \Omega \text{ q.s. em } [0, \infty); \quad \mu \geq 0; \end{cases}$$

onde as funções  $f_0 : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $f : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $g_0, g : [0, \infty) \rightarrow \mathbb{R}$ , o conjunto  $\Omega \subset \mathbb{R}^n$  e a taxa de desconto  $\delta > 0$  são dados. Os controles admissíveis são pares do tipo  $(u, \mu)$ . A componente convencional do controle,  $u$ , é uma função mensurável satisfazendo a restrição  $u(t) \in \Omega$  q.s. com respeito a medida de Lebesgue. A componente impulsiva do controle,  $\mu$ , corresponde a uma medida de Borel não negativa.

Consideramos o problema impulsivo  $PI_{\infty}(x_0)$  sob as seguintes hipóteses:

H1) Existe  $K_1 > 0$  tal que para todo  $t \in [0, \infty)$ ,  $u \in \Omega$

$$|f_0(t, x, u) - f_0(t, y, u)| + |f(t, x, u) - f(t, y, u)| \leq K_1 |x - y|, \forall x, y \in \mathbb{R}^n;$$

H2)  $f(\cdot, x, \cdot)$ ,  $f_0(\cdot, x, \cdot)$  são  $\mathcal{L} \times \mathcal{B}$ -mensuráveis,  $g(\cdot)$  e  $g_0(\cdot)$  são  $\mathcal{L}$ -mensuráveis;

H3) Existe  $K_2 > 0$  tal que para todo  $t \in [0, T]$  e  $u \in \Omega$

$$|f(t, x, u)| + |f_0(t, x, u)| \leq K_3(1 + |x|), \forall x \in \mathbb{R}^n;$$

H4) As funções  $f(t, \cdot, u)$ ,  $f_0(t, \cdot, u)$  são continuamente diferenciáveis para todo  $t \in [0, \infty)$ ,  $u \in \Omega$ . A aplicação

$$(t, x, u) \mapsto (f_{0,x}(t, x, u), f_x(t, x, u))$$

é contínua em  $[0, \infty) \times \mathbb{R}^n \times \Omega$ ;

H5)  $\Omega$  é fechado.

O teorema a seguir deriva do resultado discutido em [BLS1]. Este, por sua vez, é a generalização para horizonte infinito do resultado obtido por R. Rishel para horizonte finito em 1965 (veja [Ri, Teorema 4]). A verificação do teorema a seguir segue essencialmente a linha de demonstração do Teorema 3.2.2, utilizando como argumentos principais: princípio de otimalidade de Bellman (para problemas impulsivos); dependência contínua de condições iniciais para soluções do sistema Hamiltoniano.

É importante ressaltar que em [BLS1] são considerados problemas  $PI_{\infty}(x_0)$  com  $g_0 = g_0(t, x)$  e  $g = g(t, x)$ , enquanto que em [Ri] são tratados problemas impulsivos com horizonte finito e  $g_0 = g_0(t)$ ,  $g = g(t)$ .

**Teorema 3.3.1** *Suponha que  $(\bar{z}, \bar{u}, \bar{\mu})$  é um processo ótimo para  $PI_{\infty}(z_0)$ . Então existem constantes  $\eta = 1$  ou  $\eta = 0$ ,  $\lambda_0 \in \mathbb{R}^n$  e uma função  $\lambda \in BV_{loc}^+([0, \infty); \mathbb{R}^n)$  tais que*

i) *O par  $(\bar{z}, \lambda)$  é solução do sistema Hamiltoniano*

$$\begin{aligned} d\bar{z}(t) &= f(t, \bar{z}(t), \bar{u}(t)) dt + \bar{\mu}(dt) \\ d\lambda(t) &= -f_x(t, \bar{z}(t), \bar{u}(t)) \lambda(t) dt - \eta e^{-\delta t} f_{0,x}(t, \bar{z}(t), \bar{u}(t)) dt \\ \bar{z}(0) &= z_0, \quad \lambda(0) = \lambda_0; \end{aligned}$$



ii) É verificada a condição de optimalidade

$$\langle f(t, \bar{z}(t), \bar{u}(t)), \lambda(t) \rangle + \eta e^{-\delta t} f_0(t, \bar{z}(t), \bar{u}(t)) = \max_{u \in \Omega} \{ \langle f(t, \bar{z}(t), u), \lambda(t) \rangle + \eta e^{-\delta t} f_0(t, \bar{z}(t), u) \} \quad q.s.;$$

iii) É verificada a condição de salto

$$\begin{aligned} \langle g(t), \lambda(t) \rangle + \eta e^{-\delta t} g_0(t) &\leq 0, \quad q.s. \\ \langle g(t), \lambda(t) \rangle + \eta e^{-\delta t} g_0(t) &= 0, \quad \bar{\mu}-q.s. \end{aligned}$$

**Demonstração:** Veja [BLS1, Teorema 6]. ■

Uma análise detalhada dos problemas impulsivos  $PI_{\infty}(x_0)$ , assim como a demonstração do Teorema 3.3.1, requer boa dose de conhecimentos sobre teoria de medida de Lebesgue e análise convexa, que estão além do escopo deste manuscrito. Entretanto, os argumentos principais utilizados na demonstração do princípio do máximo se assemelham aos utilizados na demonstração do resultado análogo para problemas com controles regulares (mensuráveis). Nosso interesse pelos problemas impulsivos (com horizonte infinito) é justificado pela importância das aplicações destes a modelos econômicos e biológicos, em que as trajetórias ótimas são sabidamente descontínuas. Um exemplo da utilização do Teorema 3.3.1 na detecção de trajetórias ótimas é apresentado na Aplicação 3.4.4 (veja ainda [BaLe]).

Maiores detalhes sobre condições necessárias para problemas impulsivos com horizonte finito podem ser encontrados em [Mu], [ViPe], [SiVi]. Para problemas impulsivos com horizonte infinito citamos ainda [BLS1], [BLS2].

### 3.4 Aplicações do Princípio do Máximo

Neste parágrafo analisamos, sob a ótica do princípio do máximo, alguns problemas de controle ótimo. Na Aplicação 3.4.1 são discutidos os *problemas de tempo mínimo até a origem*. O modelo aqui discutido é simples, porém possui diversas variantes relacionadas com aplicações tecnológicas (piloto automático, acoplamento de satélites, ...).

Utilizamos na Aplicação 3.4.2 o princípio do máximo para verificar que uma estratégia do tipo bang-bang é a única estratégia ótima admissível para o *problema da alunissagem*.

Na Aplicação 3.4.3, consideramos um modelo econômico clássico, formulado por F. Ramsey em 1928 (veja [Ra]). Utilizando a equação de Euler-Lagrange, obtemos a política ótima para um problema do tipo *consumo × investimento* com horizonte infinito.

Um problema de exploração ótima de recursos naturais renováveis é considerado na Aplicação 3.4.4. A abordagem aqui apresentada segue a linha de [BaLe]. O modelo utilizado foi proposto em [CCM] e analisado por diversos autores. São levados em conta controles do tipo impulsivos e processos definidos em horizonte infinito de tempo.

Em [BMS] podem ser encontradas diversas aplicações do princípio do máximo a problemas aeroespaciais, dentre as quais citamos: Desenho ótimo de uma missão a Netuno; Ascensão ótima de um veículo espacial hipersônico; Alcance máximo de voo para uma

asa delta atravessando uma térmica. Em [Ho] são discutidas (entre outras) as seguintes aplicações: Oscilador harmônico com custo de combustível; Controle de epidemias; Pescaria ótima; Contração do ventrículo esquerdo do coração; Compra e venda de ações.

**Aplicação 3.4.1 (Tempo mínimo)** Suponha que no tempo  $t = 0$  um carro se encontra na posição  $a \in \mathbb{R}$  com velocidade  $b \in \mathbb{R}$ . Nosso objetivo é encontrar uma estratégia  $u^*$  (aceleração e frenagem) que permita levá-lo até a origem no menor tempo possível. É exigido ainda que, ao chegar ao destino, o carro tenha velocidade nula.<sup>4</sup> Temos assim o seguinte problema de controle ótimo

$$\begin{cases} \text{Minimizar } \int_0^T 1 \, dt \\ \text{sujeito a} \\ z' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} z + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u, \quad z(0) = \begin{pmatrix} a \\ b \end{pmatrix}; \\ \psi(T, z(T)) := z(T) = \theta \end{cases}$$

onde  $T \geq 0$ ,  $u \in L^1[0, T]$  e  $u(t) \in \Omega := [-1, 1]$  q.s. em  $[0, T]$ . As condições necessárias fornecidas pelo princípio do máximo são

$$\begin{aligned} z' &= \begin{pmatrix} z_2 \\ u \end{pmatrix}, \quad z(0) = \begin{pmatrix} a \\ b \end{pmatrix}, \quad z(T) = \theta; \\ \lambda' &= - \begin{pmatrix} 0 \\ \lambda_1 \end{pmatrix}, \quad \lambda(T) = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}; \\ H(t, z, \lambda, u) &= z_2 \lambda_1 + u \lambda_2 + \eta; \\ z_2 \lambda_1 + U_*(z, \lambda) \lambda_2 + \eta &= \min_{u \in [-1, 1]} \{z_2 \lambda_1 + u \lambda_2 + \eta\}, \quad \text{q.s. em } [0, T]; \\ \eta + |\mu_1| + |\mu_2| &\neq 0. \end{aligned}$$

Da condição de otimalidade obtemos

$$U_*(z, \lambda) = \begin{cases} -\text{sign } \lambda_2 & , \lambda_2 \neq 0 \\ ? & , \lambda_2 = 0 \end{cases} \quad (3.7)$$

Calculando agora  $\lambda(t)$  para  $t \in [0, T]$  temos:

$$\lambda_1(t) = \mu_1, \quad \lambda_2(t) = (T - t)\mu_1 + \mu_2, \quad t \in [0, T]. \quad (3.8)$$

Portanto,  $\lambda_2$  é linear e muda de sinal no máximo uma vez em  $[0, T]$ . Sendo assim, basta estudar os seguintes casos:

1º Caso:  $u^*$  não muda de sinal em  $[0, T]$ .

- Se  $u^* \equiv 1$  temos

$$z_2(t) = b + t, \quad z_1(t) = a + bt + \frac{t^2}{2}, \quad t \in [0, T].$$

<sup>4</sup>Tal problema é apresentado por alguns autores como *problema da acoplagem*. Para uma descrição do problema veja [Tr, Capítulo 11].

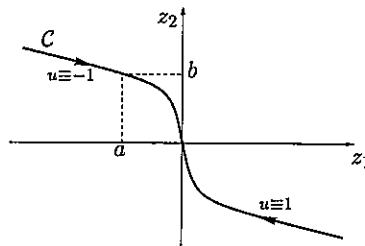


Figura 3.2: Trajetórias ótimas para os controles  $u^* \equiv 1$  e  $u^* \equiv -1$

Como  $z(T) = \theta$ , tais estratégias são admissíveis apenas para condições iniciais do tipo  $(a, b) = (T^2/2, -T)$ , com  $T > 0$ .

- Se  $u^* \equiv -1$  temos

$$z_2(t) = b - t, \quad z_1(t) = a + bt - \frac{t^2}{2}, \quad t \in [0, T].$$

Como  $z(T) = \theta$ , tais estratégias são admissíveis apenas para condições iniciais do tipo  $(a, b) = (-T^2/2, T)$ , com  $T > 0$ .

A curva  $C$  na Figura 3.2 é composta pelas condições iniciais  $(a, b)$ , para as quais as estratégias  $u^* \equiv 1$  ou  $u^* \equiv -1$  são ótimas. As respectivas trajetórias correspondem a parte da curva  $C$  limitada por  $(a, b)$  e pela origem.

2º Caso:  $u^*$  muda de sinal em  $\tau \in (0, T)$ .

No caso anterior vimos que se  $u^* \equiv 1$ , então  $z_2(t)^2 = 2z_1(t) + const$ ; enquanto que  $u^* \equiv -1$  implica em  $z_2(t)^2 = -2z_1(t) + const$ . Portanto, as trajetórias correspondentes a tais controles são necessariamente paralelas a um dos arcos de parábola mostrados na Figura 3.3. De onde concluímos que a trajetória ótima é necessariamente composta por dois arcos: cada um pertencente a uma das famílias na Figura 3.3 (lembre que  $u^*$  muda de sinal uma única vez no intervalo  $[0, T]$ ).

Note que a parte final da trajetória ótima é necessariamente como na Figura 3.2 (caso contrário a trajetória não seria admissível). Para determinar a parte inicial da trajetória,

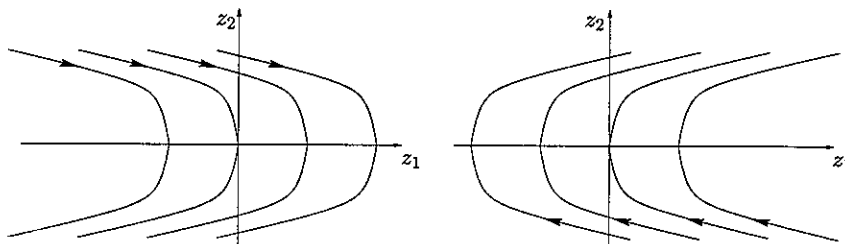


Figura 3.3: Trajetórias correspondentes a controles constantes

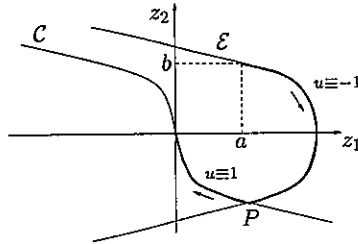


Figura 3.4: Trajetórias ótimas correspondentes aos controles  $u^* \equiv 1$  e  $u^* \equiv -1$

observe que, dada uma condição inicial  $(a, b) \in \mathbb{R}^2$ , existe uma única curva pertencente as famílias mostradas na Figura 3.3, que intercepta tanto o ponto  $(a, b)$  quanto a curva  $C$  (o caso  $a > 0, b > 0$  é mostrado na Figura 3.4). O princípio do máximo nos permite concluir que existe uma única trajetória associada a controles do tipo

$$u^*(t) = \begin{cases} 1, & t < \tau \\ -1, & t > \tau \end{cases} \quad \text{ou} \quad u^*(t) = \begin{cases} -1, & t < \tau \\ 1, & t > \tau \end{cases}$$

que é admissível para a condição inicial  $(a, b)$ . Tal trajetória é composta por dois arcos: um da curva  $\mathcal{E}$  limitado por  $(a, b)$  e pelo ponto  $P$  e outro da curva  $C$  limitado por  $P$  e pela origem. Para calcular  $\tau$  (instante em que trocamos o controle de  $-1$  para  $1$ ) não é necessário calcular as constantes  $\mu_1, \mu_2$  na equação (3.8). No caso  $a > 0, b > 0$ , basta descobrir para qual  $\tau > 0$  a curva  $(z_1(t), z_2(t)) = (a + bt - t^2/2, b - t)$  satisfaz a condição

$$z_2(\tau) < 0, \quad z_2(\tau)^2 = 2z_1(\tau).$$

Um cálculo simples mostra que  $\tau$  é dado por uma das raízes  $b \pm \sqrt{b^2/2 - a}$ . □ □ □

**Aplicação 3.4.2 (Alunissagem)** Considere o problema de controlar a descida de uma espaçonave na Lua, utilizando para isso a menor quantidade possível de combustível. Em um modelo simplificado temos<sup>5</sup>

- $t$  : tempo;
- $h(t)$  : altura da espaçonave;
- $v(t)$  : velocidade da espaçonave;
- $m(t)$  : massa da espaçonave + combustível;
- $u(t)$  : empuxo dos motores da espaçonave.

Seja  $M$  a massa da espaçonave sem combustível,  $F$  a quantidade inicial de combustível,  $h_0$  a altura inicial,  $v_0$  a velocidade inicial,  $u_{max}$  o empuxo máximo dos motores da nave ( $0 \leq u(t) \leq u_{max}, t \geq 0$ ),  $g$  a constante gravitacional da Lua (considerada constante) e  $k$

<sup>5</sup>Este modelo é também discutido em [FIRi], [Ho] e [Know].

a constante de proporcionalidade entre o empuxo e a taxa de queima do combustível. As variáveis de estado  $(h, v, m)$  satisfazem a seguinte dinâmica:

$$\begin{cases} h' = v(t) \\ v' = -g + u(t)/m(t) \\ m' = -ku(t) \end{cases}$$

Definindo  $z(t) = (h(t), v(t), m(t))^*$  temos o sistema não linear

$$\begin{cases} z' = \begin{pmatrix} z_2 \\ -g + u/z_3 \\ -ku \end{pmatrix} =: f(t, z, u) \\ z(0) = (h_0, v_0, M + F)^*, \quad z(T) = (0, 0, ?)^*. \end{cases} \quad (3.9)$$

A condição final segue da hipótese que um pouso suave ocorre quando  $h(T) = v(T) = 0$ , sendo para  $m$  somente relevante que  $m(T) \geq M$ . Como o custo a ser minimizado corresponde ao gasto de combustível vel, temos que maximizar

$$m(T) = M + F - k \int_0^T u(t) dt.$$

O problema de controle ótimo pode ser escrito como:

$$\begin{cases} \text{Minimizar } J(T, z, u) = \int_0^T u(t) dt \\ \text{sujeito a} \\ u \in \{L^1[0, T] \mid u(t) \in \Omega := [0, u_{max}] \text{ q.s. em } [0, T]\}, \\ z' = f(z, u), \quad z(0) = (h_0, v_0, M + F)^* \in \mathbb{R}^3, \\ \psi(T, z(T)) = (z_1(T), z_2(T))^* = \theta \in \mathbb{R}^2 \end{cases}$$

A função de Hamilton é

$$H(t, z, \lambda, u) = \langle \lambda, u \rangle + \eta L(t, z, u) = \lambda_1 z_2 + \lambda_2 (-g + u/z_3) - \lambda_3 k u + \eta u.$$

Minimizando a função de Hamilton em relação a  $u$  obtemos

$$U_*(z, \lambda) = \begin{cases} 0 & , \eta + \lambda_2/z_3 - k\lambda_3 > 0 \\ ? & , \eta + \lambda_2/z_3 - k\lambda_3 = 0 \\ u_{max} & , \eta + \lambda_2/z_3 - k\lambda_3 < 0 \end{cases} \quad (3.10)$$

Tomamos por simplicidade  $u_{max} = 1$ . A fim de tornar o problema fisicamente coerente supomos ainda

$$1 = \text{empuxo máximo} > \text{força gravitacional} = (M + F)g,$$

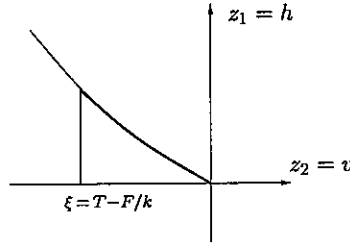


Figura 3.5: Condições iniciais  $(h, v)$  que são levadas pelo controle  $u^* \equiv 1$  à condição final  $(0, 0, m(T))^*$  com  $m(T) \geq M$ .

isto é  $1/(M + F) > g$ . É razoável considerar que existe uma estratégia ótima do tipo bang-bang, i.e. da forma

$$u^*(t) = \begin{cases} 0 & , t \in [0, \xi] \\ 1 & , t \in [\xi, T] \end{cases} \quad (3.11)$$

Cálculamos inicialmente a trajetória associada a estratégia  $u^*$ . Como  $u^* \equiv 1$  em  $[\xi, T]$ , usamos o sistema (3.9) e as condições de contorno  $z_1(T) = z_2(T) = 0$  e  $z_3(\xi) = M + T$  a fim de determinar  $z$  no instante  $t = \xi$ . Obtemos assim

$$\begin{cases} z_1(\xi) = -\frac{1}{2}g(T - \xi)^2 - \frac{M+F}{k^2} \ln\left(\frac{M+F-k(T-\xi)}{M+F}\right) - \frac{T-\xi}{k} \\ z_2(\xi) = g(T - \xi) + \frac{1}{k} \ln\left(\frac{M+F-k(T-\xi)}{M+F}\right) \\ z_3(\xi) = M + F \end{cases}$$

Traçando o gráfico de  $z_1(\xi)$  por  $z_2(\xi)$  obtemos a curva da Figura 3.5, que é formada pelos estados da forma  $z(\xi) = (h(\xi) \ v(\xi) \ M + F)^*$  que são levados pelo controle  $u^*(t) = 1$ ,  $t \in [\xi, T]$  no estado final  $z(T) = (0 \ 0 \ m(T))^*$  com  $m(T) \geq M$ . Note que o comprimento dessa curva é limitado, pois como  $u^* \equiv 1$  temos  $m' = -k$  e o combustível se esgotará após  $F/k$  unidades de tempo. Temos assim a limitação  $T - \xi \leq F/k$  (além obviamente de  $T - \xi \geq 0$ ).

Como inicialmente  $u^* \equiv 0$ , a nave se encontra em *queda livre* durante o intervalo de tempo  $[0, \xi]$ . A trajetória correspondente é

$$\begin{cases} z_1(t) = -\frac{1}{2}gt^2 + v_0t + h_0 \\ z_2(t) = -gt + v_0 \\ z_3(t) = M + F \end{cases} \quad t \in [0, \xi].$$

Explicitando  $z_1 (= h)$  em função de  $z_2 (= v)$  obtemos:

$$h(t) = h_0 - \frac{1}{2g}[v^2(t) - v_0^2], \quad t \in [0, \xi].$$

A curva  $(v(t), h(t))$  é uma parábola no plano de fase  $v \times h$ . Unindo os dois trechos da trajetória correspondente a  $u^*$  obtemos a curva mostrada na Figura 3.6. Segundo essa trajetória a nave cai em queda livre até que o estado  $(v, h)$  alcance a curva da Figura 3.5. Nesse momento os motores são acionados na potência máxima até o estado final admissível ser atingido  $(\psi(T, z(T)) = \theta)$ .

Observe que se a intersecção das duas curvas na Figura 3.6 ocorre num ponto  $(v(\xi), h(\xi))$  com  $\xi < T - F/k$ , a quantidade de combustível não é suficiente para realizar um pouso suave. Enquanto que se a condição inicial  $(v_0, h_0)$  se encontra abaixo da curva na Figura 3.5, mesmo empregando empuxo máximo  $u(t) = 1, t \in [0, T]$ , o solo lunar é atingido com  $v(T) < 0$ .

Através do princípio do máximo verificamos agora que a estratégia de controle definida em (3.11) é um candidato a controle ótimo. Suponha  $\lambda(0) = (l_1 \ l_2 \ l_3)^*$ . Substituindo na equação adjunta

$$\begin{cases} \lambda_1' = 0 \\ \lambda_2' = -\lambda_1 \\ \lambda_3' = \lambda_2 u / z_3^2 \end{cases}$$

temos:

$$\lambda_1(t) = l_1, \ t \in [0, T]; \quad \lambda_2(t) = l_2 - l_1 t, \ t \in [0, T]; \quad \lambda_3(t) = l_3, \ t \in [0, \xi].$$

Como  $z_3(t) = M + F - k(t - \xi), t \in [\xi, T]$ , podemos calcular  $\lambda_3$  no intervalo final de tempo, obtendo

$$\lambda_3(t) = l_3 + \int_{\xi}^t \frac{l_2 - l_1 s}{[k(\xi - s) + M + F]^2} ds, \ t \in [\xi, T].$$

Defina agora  $r(t) := \eta + \lambda_2(t)/z_3(t) - k\lambda_3(t), t \in [0, T]$ . De (3.10) sabemos que a escolha do controle  $u^*$  no tempo  $t$  depende de  $\text{sign}(r(t))$ . Portanto, como a estratégia de controle  $u^*$  *salta* de 0 para 1 em  $t = \xi$  temos obrigatoriamente

$$r(\xi) = \eta + \frac{\lambda_2(\xi)}{z_3(\xi)} - k\lambda_3(\xi) = 0.$$

Escolhendo  $\eta = 1$  (que satisfaz condição de transversalidade) reescrevemos a equação acima como

$$1 + \frac{l_2 - l_1 \xi}{M + F} - k l_3 = 0.$$

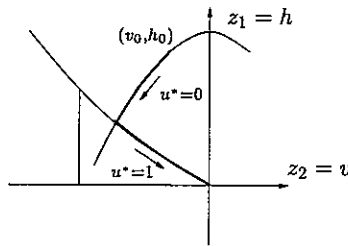


Figura 3.6: Trajetória correspondente à estratégia bang-bang  $u^*$ .

A escolha de  $u^*$  em (3.11) implica em  $r(t) > 0$ ,  $t \in [0, \xi]$ . Portanto  $l_1 > 0$  necessariamente.

O princípio do máximo nos fornece ainda uma condição inicial para a equação adjunta:

$$\lambda(T) = -\frac{\partial \psi}{\partial z}(T, z(T))^* \mu = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^* \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = -\begin{pmatrix} \mu_1 \\ \mu_2 \\ 0 \end{pmatrix},$$

de onde concluímos que  $\lambda_3(T) = 0$ . Obtemos assim para  $l_1, l_2, l_3$  o sistema sub-determinado de equações lineares

$$\begin{cases} 1 + (M + F)^{-1}(l_2 - l_1 \xi) - k l_3 = 0 \\ l_3 + \int_{\xi}^T \frac{l_2 - l_1 s}{[k(\xi - s) + M + F]^2} ds = 0 \end{cases}$$

Considerando  $\lambda_2$  como parâmetro o sistema se reescreve como

$$\begin{cases} (M + F)^{-1} \xi l_1 + k l_3 = 1 + l_2 (M + F)^{-1} \\ \mathcal{P} l_1 - l_3 = \mathcal{Q} l_2 \end{cases}$$

onde  $\mathcal{P} = \int_{\xi}^T s [k(\xi - s) + M + F]^{-2} ds$  e  $\mathcal{Q} = \int_{\xi}^T [k(\xi - s) + M + F]^{-2} ds$  são constantes positivas. Resolvendo o novo sistema obtemos:

$$\begin{pmatrix} l_1 \\ l_3 \end{pmatrix} = \begin{pmatrix} [1 + ((M + F)^{-1} + k \mathcal{Q}) l_2] / [\xi (M + F)^{-1} + k \mathcal{P}] \\ \mathcal{P} [1 + ((M + F)^{-1} + k \mathcal{Q}) l_2] / [\xi (M + F)^{-1} + k \mathcal{P}] - \mathcal{Q} l_2 \end{pmatrix} \quad (3.12)$$

Note que para  $t \in [\xi, T]$  temos

$$\begin{aligned} r(t) &= 1 + (l_2 - l_1 t) / z_3(t) - k \lambda_3(t) \\ &< 1 + l_2 / M - k l_3 + (\mathcal{P} - \xi / (M + F)) l_1. \end{aligned} \quad (3.13)$$

Substituindo em (3.13) as expressões encontradas em (3.12) para  $l_1$  e  $l_3$ , obtemos uma restrição linear para escolha de  $l_2$ . Outra restrição (também linear) para  $l_2$  é dada por  $l_1 > 0$  e (3.12). Como o problema assim colocado possui solução não única, é possível encontrar condição inicial  $(l_1 \ l_2 \ l_3)^*$  de forma que a função  $r$  satisfaça

$$\begin{cases} r(t) > 0, & t \in [0, \xi] \\ r(t) < 0, & t \in (\xi, T] \end{cases}$$

provando que  $u^*$  satisfaz as condições do princípio do máximo.

Verificamos agora que  $u^*$  é o único candidato fornecido pelo princípio de Pontryagin. A função  $r(t)$  obtida de (3.10) determina quando ocorrem saltos na estratégia de controle. Note ainda que como  $\lambda_1 \equiv l_1$  então  $\lambda'_2 \equiv -l_1$  e temos

$$\begin{aligned} r'(t) &= (\lambda'_2 z_3 - \lambda_2 z'_3) z_3^{-2} - k \lambda'_3 \\ &= \lambda'_2 / z_3 - \lambda_2 (-k u) z_3^{-2} - k \lambda_2 u z_3^{-2} \\ &= -l_1 / z_3(t), \quad t \in [0, T]. \end{aligned}$$

Analisamos separadamente as situações possíveis:



- $l_1 \neq 0$ : Como  $z_3(t) = m(t) > 0$  então  $r$  é monótona. Se  $l_1 > 0$  obtemos um controle do tipo  $u^*$ . Se  $l_1 < 0$  obtemos uma estratégia oposta, i.e. inicialmente  $u = 1$  e depois  $u = 0$ . Com essa estratégia não é possível obter um pouso suave. De fato, ou  $(v_0, h_0)$  se situa abaixo ou acima do gráfico na Figura 3.5. No primeiro caso já vimos que  $v(T) < 0$ . No segundo caso, como  $u$  é da forma

$$u(t) = \begin{cases} 1, & t \in [0, \tau) \\ 0, & t \in [\tau, T] \end{cases}$$

obtemos do sistema adjunto

$$v(T) - v(\tau) = \int_{\tau}^T v'(t) dt = \int_{\tau}^T -g dt = -g(T - \tau).$$

Logo uma condição necessária para  $v(T) = 0$  é que  $T = v(\tau)/g + \tau$ . Novamente do sistema adjunto obtemos

$$-h(\tau) = h(T) - h(\tau) = \int_{\tau}^T h'(t) dt = \int_{\tau}^T v(t) dt = \int_{\tau}^T g(T - t) dt = \frac{g}{2}(T - \tau)^2,$$

isto é  $h(\tau) = -v^2(\tau)/2g < 0$ . Portanto a transição de 1 para 0 na estratégia de controle ocorre abaixo da superfície da Lua, e o pouso obviamente não é suave.

- $l_1 = 0$ : Neste caso  $r' = 0$  e  $r$  é constante. Se  $r \neq 0$  os possíveis candidatos são  $u \equiv 0$  e  $u \equiv 1$ . O primeiro controle obviamente não permite pouso suave. Já o segundo será ótimo somente se  $(v_0, h_0)$  pertence à curva na Figura 3.5, quando a estratégia se torna idêntica a  $u^*$ . Por fim, se  $r = 0$  temos

$$1 + l_2 \frac{1}{z_3(t)} + k \lambda_3(t) = 0, \quad t \in [0, T],$$

isto é, as funções  $\{1, z_3^{-1}, \lambda_3\}$  são linearmente dependentes. Mas isto é uma contradição pois

$$z_3(t) = M + F - k \int_0^t u(s) ds, \quad \lambda_3(t) = l_2 \int_0^t u(s) z_3(s)^{-2} ds.$$

Portanto o único controle admissível que satisfaz as condições do princípio do máximo é  $u^*$  definido em (3.11). □ □ □

**Aplicação 3.4.3 (Consumo  $\times$  Investimento)** Tratamos a seguir um problema clássico da economia, que foi um dos primeiros a ser tratado sob a ótica do cálculo variacional. Consideramos o seguinte problema macroeconômico: Como equacionar a relação entre consumo e investimento, a fim de otimizar o desenvolvimento econômico?

Suponha que a economia de uma nação é representada pelas variáveis

- $K(t)$  : Capital;  
 $C(t)$  : Consumo;  
 $Y(t)$  : Produção (produto interno);  
 $K'(t)$  : Investimento (variação do Capital);

ao longo do intervalo de tempo  $t \in [0, \infty)$ . Considere ainda o seguinte modelo simplificado:

- i)  $Y = g(K)$ , onde  $g' > 0$  e  $g'' \leq 0$ ;  
 ii)  $C = Y - K'$  (parte da produção é consumida e o restante é reinvestido);  
 iii)  $K(0) = K_0$  (o capital inicial é conhecido);  
 iv)  $U = U(C)$  é a utilidade do capital, onde  $U' > 0$ ,  $U'' < 0$ ;  
 v)  $\delta > 0$  é o fator de desconto.

A tarefa a ser realizada é a de encontrar uma política ótima de investimento para o problema:

$$\left\{ \begin{array}{l} \text{Maximizar } \int_0^{\infty} e^{-\delta t} U(C(t)) dt \\ \text{sujeito a } K' = g(K) - C, K(0) = K_0. \end{array} \right.$$

Este problema foi originalmente formulado e resolvido por Ramsey em 1928 (veja [Ra]). A hipótese  $C = g(K) - K'$  nos permite analisar este problema utilizando cálculo variacional (veja Parágrafo 2.2). Note que a equação de Euler-Lagrange é dada por

$$K'' - g'(K)K' + \frac{U'(g(K) - K')}{U''(g(K) - K')}(\delta - g'(K)) = 0.$$

No caso geral esta equação não pode ser resolvida analiticamente. Fazemos aqui a hipótese simplificadora:

$$U(r) = \frac{1}{1-q} r^{1-q}, \quad g(r) = br,$$

onde  $b > 0$ ,  $q \in (0, 1)$ . Neste caso particular a equação de Euler-Lagrange se simplifica para uma equação que sabemos resolver, a saber:

$$qK'' + (\delta - b - qb)K' + b(b - \delta)K = 0.$$

Calculando as raízes do polinômio característico, temos  $\lambda_1 = b$ ,  $\lambda_2 = a := q^{-1}(b - \delta)$ . Portanto, as soluções que satisfazem a condição inicial  $K(0) = K_0$  são da forma:

$$K^*(t) = (K_0 - A)e^{at} + Ae^{bt}, \quad t \geq 0,$$

onde  $A$  é um parâmetro livre. Suponha agora que  $b > a$ , i.e.  $\delta > (1 - q)b$ . Neste caso as hipóteses do modelo:

$$C(t) = g(K(t)) - K'(t) > 0, \quad t \geq 0 \quad \text{e} \quad \lim_{t \rightarrow \infty} K(t) \geq 0$$

são satisfeitas respectivamente para  $A \geq 0$  e  $A < K_0$ . Para determinar o parâmetro  $A \in [0, K_0)$  é necessária uma condição de contorno para a equação da dinâmica - por exemplo

$K'(0)$  ou  $K(\infty)$ . Como tal condição não é explicitamente fornecida, é preciso analisar a equação de Hamilton–Jacobi–Bellman, que para este problema autônomo se escreve como

$$-\delta V(x) + \min_{u \in \Omega} \left\{ \frac{\partial V}{\partial x} g(x) - u + \frac{1}{1-q} u^{1-q} \right\} = 0, \quad x \in \mathbb{R}^n$$

(note que  $\delta = b - qa$ ). É fácil verificar que  $V(x) := (1 - q)/(b - a)^q x^{1-q}$ ,  $x \geq 0$  é solução da equação acima. Note ainda que se  $A = 0$  a condição

$$\liminf_{t \rightarrow \infty} e^{-\delta t} V(K^*(t)) = 0$$

é satisfeita, pois  $a(1 - q) < 0$  por hipótese. Portanto  $V(x)$  e a trajetória  $K^*(t) = K_0 e^{at}$  satisfazem as condições do teorema, de onde concluímos que uma estratégia ótima de consumo é dada por

$$C^*(t) = (b - a)K_0 e^{at}, \quad t \geq 0.$$

□ □ □

**Aplicação 3.4.4 (Pescaria ótima)** Consideramos a seguir um modelo bio-econômico para pescaria comercial controlada por monopólio. O modelo em questão é representado pelas seguintes quantidades:

- $t$  : tempo;
- $x(t)$  : população de peixes (biomassa);
- $h(t)$  : taxa de captura;
- $E(t)$  : esforço de pesca;
- $K(t)$  : capital investido na atividade pesqueira;
- $I(t)$  : taxa de investimento.

O modelo é baseado nas seguintes hipóteses:

- $h(t) = qE(t)x(t)$ ;  $q$  é um coeficiente de captura;
- $x'(t) = F(x(t)) - qE(t)x(t)$ ,  $t \geq 0$ ,  $x(0) = x_0$ ;  $F$  é a função de crescimento natural;
- $K'(t) = -\gamma K(t) + I(t)$ ,  $t \geq 0$ ,  $K(0) = K_0$ ;  $\gamma \geq 0$  é a taxa de depreciação do capital;
- Restrições:  $0 \leq x(t), K(t), E(t)$ ;  $E(t) \leq K(t)$ ;
- Não-maleabilidade:  $0 \leq I(t) \leq \infty$ ,  $t \geq 0$ ;
- Existência de dois pontos de equilíbrio biológico:  $F(0) = F(\bar{x}) = 0$ ,  $\bar{x} > 0$ ;
- Propriedades da função  $F$ :  
 $F \in C^2[0, \infty)$ ,  $F(x) > 0$ ,  $0 < x < \bar{x}$ ,  $F''(x) < 0$ ,  $0 \leq x \leq \bar{x}$ ;
- Função objetivo (fluxo descontado de dinheiro):  $\int_0^\infty e^{-\delta t} \{ph(t) - cE(t) - rI(t)\} dt$ ;
- $\delta > 0$  é a taxa instantânea de desconto;  $p \geq 0$  é o preço de mercado do peixe;  $c \geq 0$  é o custo de operação por unidade de esforço de pesca;  $r \geq 0$  é o preço do capital.

Um exemplo concreto para a função de crescimento natural satisfazendo as hipóteses acima é dado pela função logística:  $F(x) := ax(1 - \frac{x}{k})$  ( $a > 0$ ,  $k > 0$ ). Para construção das figuras na seqüência do texto foi utilizada esta função.

**O problema de controle ótimo**

Definimos  $E = uK$ , com  $u \in [0, 1]$  e obtemos assim uma segunda variável de controle:  $u$ . Podemos supor sem perda de generalidade  $q = 1$ . Assim sendo, escrevemos o problema de controle ótimo da seguinte forma:

$$Q(x_0, K_0) \left\{ \begin{array}{l} \text{Minimizar } J(x_0, K_0; I, u) := \int_0^{\infty} e^{-\delta t} \{rI(t) + cu(t)K(t) - pu(t)K(t)x(t)\} dt \\ \text{sujeito a} \\ x' = F(x) - u(t)K(t)x, t \geq 0, x(0) = x_0, \\ K' = -\gamma K + I(t), t \geq 0, K(0) = K_0, \\ x(t) \geq 0, K(t) \geq 0, 0 \leq u(t) \leq 1, I(t) \geq 0, t \in [0, \infty). \end{array} \right.$$

Este problema foi formulado em [CCM], onde foi considerado através de uma abordagem variacional. Entretanto a análise apresentada não é rigorosa nos detalhes. Em [Mu] o modelo acima é utilizado para ilustrar os problemas considerados no artigo, entretanto os resultados lá obtidos não se aplicam a este problema (prova de existência). Em [ViPe] também pode ser encontrada uma referência a este problema. Finalmente em [Cl] o problema  $Q(x_0, K_0)$  é apresentado como uma aplicação do princípio do máximo, mas o tratamento novamente não é rigoroso. Para uma melhor descrição do modelo bio-econômico o leitor deve consultar [An] ou [SeSy].

É um fato conhecido (veja e.g [Mu]) que problemas de controle podem não possuir solução caso a variável de controle não seja limitada e tanto a função objetivo quanto a dinâmica sejam lineares no controle. Este é exatamente o caso acima em relação ao controle  $I$ . A fim de evitar problemas de não-existência, temos que substituir o controle convencional  $I$  por um controle impulsivo. Uma consequência imediata desta escolha é que a trajetória da variável de estado correspondente ao capital se torna descontínua. Sendo assim, consideramos  $I$  como uma medida de Borel e  $K$  como uma função de variação limitada. O problema  $Q(x_0, K_0)$  se reescreve da seguinte forma:

$$P(x_0, K_0) \left\{ \begin{array}{l} \text{Minimizar } J(x_0, K_0; \mu, u) := \int_0^{\infty} e^{-\delta t} r \mu(dt) + \int_0^{\infty} e^{-\delta t} \{c - pqx(t)\} u(t) K(t) dt \\ \text{sujeito a } (u, \mu) \in U_{ad} \times C^*, \\ x' = F(x) - u(t)Kx, x(0) = x_0, \\ dK = -\gamma K dt + \mu(dt), K(0) = K_0, \end{array} \right.$$

onde

$$\begin{aligned} U_{ad} &:= \{v \in L^\infty[0, \infty) \mid 0 \leq v(t) \leq 1 \text{ q.s. em } [0, \infty)\}, \\ C^* &:= \{\mu \mid \mu \text{ medida de Borel não-negativa em } [0, \infty)\}. \end{aligned}$$

Observe que as restrições  $x(t) \geq 0$ ,  $K(t) \geq 0$  em  $[0, \infty)$  são satisfeitas (devido as hipóteses acima), caso  $x_0 \geq 0$ ,  $K_0 \geq 0$ . O problema de valor inicial

$$dK = -\gamma K dt + \mu(dt), K(0) = K_0 \quad (3.14)$$

tem que ser considerado como uma equação diferencial com medida. Uma função  $K : [0, t_1] \rightarrow \mathbb{R}$  ( $t_1 \in (0, \infty)$ ) é solução de (3.14) quando

$$K(t) = K_0 - \int_0^t \gamma K(s) ds + \int_{[0,t]} \mu(ds), \quad 0 \leq t < t_1. \quad (3.15)$$

Isto implica que  $K$  é contínua pela direita em  $(0, t_1)$  e  $K(0) = K_{0,+} \mu(\{t_0\})$ , onde  $K_{0,+}$  representa  $\lim_{t \downarrow 0} K(t)$ .

**Observação 3.4.5** O conceito de soluções definido acima caracteriza as denominadas *soluções de Young* (veja [Ri]). Podemos aqui usar este conceito pois os coeficientes  $g, g_0$  não dependem das variáveis de estado  $x, K$ . Se fosse este o caso, teríamos que utilizar o conceito de soluções robustas (veja e.g. [SiVi], [BLS1]).  $\square \square \square$

A seguir definimos as constantes  $\kappa := \delta + \gamma$ ,  $r' := r\kappa$ ,  $c_* := c + r'$  e as funções  $g(x) := \delta - F'(x) + x^{-1}F(x)$ ,

$$\psi(x) := (px - c)(\delta - F'(x)) - \frac{cF(x)}{x}, \quad \psi_*(x) := (px - c_*)(\delta - F'(x)) - \frac{c_*F(x)}{x}$$

em  $(0, \bar{x})$ . Fazemos ainda as seguintes hipóteses (notação:  $K^* := F(x^*)/x^*$ ,  $\bar{K} := F(\bar{x})/\bar{x}$ )

$$(V1) \quad F \in C^2[0, \infty) \cap C^3(0, \bar{x}); \quad F(0) = F(\bar{x}) = 0; \quad F(x) > 0, \quad 0 < x < \bar{x}; \quad F''(x) < 0, \quad 0 \leq x \leq \bar{x};$$

$$(V2) \quad \delta > 0, \quad r > 0, \quad c > 0, \quad \gamma > 0;$$

$$(V3) \quad c_* - p\bar{x} < 0;$$

$$(V4) \quad \text{Existem } \tilde{x}, x^* \in (0, \bar{x}) \text{ com}$$

$$\psi(x) < 0, \quad 0 < x < \tilde{x}, \quad \psi(\tilde{x}) = 0, \quad \psi(x) > 0, \quad \tilde{x} < x < \bar{x},$$

$$\psi_*(x) < 0, \quad 0 < x < x^*, \quad \psi_*(x^*) = 0, \quad \psi_*(x) > 0, \quad x^* < x < \bar{x};$$

$$(V5) \quad \psi'(x) > 0, \quad x \in (0, \tilde{x}), \quad \psi'_*(x) > 0, \quad x \in (\tilde{x}, \bar{x});$$

$$(V6) \quad g'(x) > 0, \quad x \in (0, \bar{x}).$$

#### As abordagens em [CCM] e [BaLe]

Em [CCM] a equação de Hamilton-Jacobi-Bellman é considerada em  $(0, \bar{x}) \times (0, \infty)$  e é obtido um candidato a controle ótimo para cada condição inicial nesta faixa. Isto resulta na definição de uma função  $S : (0, \bar{x}) \times (0, \infty) \rightarrow \mathbb{R}$ , tal que para todos  $(x, K) \in (0, \bar{x}) \times (0, \infty)$ ,  $u \in [0, 1]$  e  $I \geq 0$ , tem-se

$$\delta S(x, K) + F(x)S_x(x, K) - \gamma K S_K(x, K) \geq I(S_K(x, K) + r) + uK\{gxS_x(x, K) - pqx + c\}. \quad (3.16)$$

É então afirmado que  $S$  é a função valor para o problema.<sup>6</sup> Implicitamente, são utilizados na argumentação somente controles cujos estados correspondentes satisfazem

$$\lim_{t \rightarrow \infty} e^{-\delta t} S(x(t), K(t)) = 0.$$

A análise apresentada em [CCM] pode ser seguida parcialmente, mas em algumas passagens as hipóteses não são suficientes e a argumentação não está completa. Como  $S$  não é diferenciável em toda a faixa, os autores argumentam que cada problema  $P(x_0, K_0)$  pode ser aproximado por uma problema  $Q(x_0, K_0)$  correspondente. Para tanto é necessário um argumento de densidade rigoroso, que não é fornecido e tão pouco pode ser encontrado na literatura especializada. Portanto, a verificação da otimalidade das estratégias discutidas em [CCM] tem que ser considerada como um problema em aberto. Esta é a justificativa apresentada em [BaLe] para uma nova análise do problema, a qual descrevemos a seguir.

As demonstrações dos resultados discutidos a seguir são, em sua maioria, longas (apesar de utilizarem essencialmente resultados de análise no  $\mathbb{R}^n$  e da teoria de equações diferenciais ordinárias). Devido a este fato, limitamo-nos a enunciar e interpretar os resultados principais, que ilustram a aplicação do princípio do máximo para este problema impulsivo com horizonte infinito. O objetivo principal da discussão aqui apresentada é obter, para cada condição inicial no plano de fase, a trajetória e o controle ótimo correspondente (veja Teorema 3.4.12). O leitor interessado encontra mais detalhes em [BaLe].

### Condições necessárias

A seguir utilizamos o princípio do máximo para obter condições necessárias de primeira ordem para o problema  $P(x_0, K_0)$ . Além disso são definidas, apartir dos multiplicadores de Lagrange, duas funções auxiliares (switches) que desempenham um papel chave na análise das trajetórias ótimas. Começamos definindo a função de Hamilton  $\tilde{H}$  por

$$\tilde{H}(t, \tilde{x}, \tilde{K}, w, \tilde{\lambda}_1, \tilde{\lambda}_2, \eta) := \tilde{\lambda}_1(F(\tilde{x}) - w\tilde{K}\tilde{x}) - \tilde{\lambda}_2\gamma\tilde{K} - \eta e^{-\delta t}(c - p\tilde{x})w\tilde{K}.$$

Seja  $(u, \mu)$  um controle ótimo para  $P(x_0, K_0)$  e  $(x, K)$  os estados correspondentes. O Teorema 3.3.1 garante a existência de constantes  $\lambda_{1,0}, \lambda_{2,0}, \eta \in \mathbb{R}$  e funções adjuntas  $\tilde{\lambda}_1, \tilde{\lambda}_2 : [0, \infty) \rightarrow \mathbb{R}$  satisfazendo o sistema Hamiltoniano, a condição de salto e a condição de otimalidade (com relação a  $\tilde{H}$ ). Defina agora

$$H(t, \tilde{x}, \tilde{K}, w, \lambda, \eta) := (-\lambda\tilde{x} + \eta(p\tilde{x} - c))\tilde{K}w,$$

onde

$$\lambda_1(t) := \tilde{\lambda}_1(t)e^{\delta t}, \quad \lambda_2(t) := \tilde{\lambda}_2(t)e^{\delta t}, \quad \lambda_{1,0} := \lambda_1(0), \quad \lambda_{2,0} := \lambda_2(0).$$

Definimos ainda as funções auxiliares  $z, \lambda : [0, \infty) \rightarrow \mathbb{R}$  por

$$z := -\lambda_1 x + \eta(px - c), \quad \lambda := \lambda_2, \quad z_0 := \lambda_{1,0}, \quad \lambda_0 := \lambda_{2,0},$$

<sup>6</sup>Lembre que a função valor  $V$  é definida por:

$$V(x_0, K_0) := \inf\{J(x_0, K_0; I, u) \mid (I, u) \text{ admissível}\}, \quad (x_0, K_0) \in (0, \bar{x}) \times (0, \infty).$$

As funções  $z$  e  $\lambda$  são denominadas switches, devido ao seu papel na obtenção dos controles ótimos. Note que  $z$  define juntamente com a condição de máximo o valor de  $u(t)$ . A saber,  $u(t) = 0$  se  $z(t) < 0$  e  $u(t) = 1$  se  $z(t) > 0$ . Se  $z(t)$  se anula, o valor de  $u(t)$  tem que ser determinado de outra forma. A função  $\lambda$  define um switch para os saltos de  $K$ , uma vez que a condição  $\mu(\{t\}) > 0$  para algum  $t \in [0, \infty)$  implica em  $\lambda(t) = \eta r$ .

Desta forma podemos reescrever as condições necessárias do princípio do máximo como:  $z_0^2 + \lambda_0^2 + \eta^2 \neq 0$ ,  $\eta \geq 0$ ,

$$\begin{aligned} \begin{pmatrix} x' \\ dK \end{pmatrix} &= \begin{pmatrix} F(x) - u(t)Kx \\ -\gamma K dt + \mu(dt) \end{pmatrix}, & \begin{pmatrix} x(0) \\ K(0) \end{pmatrix} &= \begin{pmatrix} x_0 \\ K_0 \end{pmatrix}, \\ \begin{pmatrix} z' \\ \lambda' \end{pmatrix} &= \begin{pmatrix} zg(x) - \eta\psi(x) \\ \kappa\lambda - zu(t) \end{pmatrix}, & \begin{pmatrix} z(0) \\ \lambda(0) \end{pmatrix} &= \begin{pmatrix} z_0 \\ \lambda_0 \end{pmatrix}. \end{aligned}$$

$$\lambda(t) - \eta r \leq 0 \text{ para todo } t \in [0, \infty),$$

$$\lambda(t) - \eta r \leq 0 \text{ } \mu\text{-q.s. em } [0, \infty),$$

$$z(t)K(t)u(t) \leq \max_{w \in [0,1]} z(t)K(t)w \text{ q.s. em } [0, \infty).$$

Inicialmente excluimos o caso irregular  $\eta = 0$  (de modo que possamos escolher  $\eta = 1$ ). Isto é consequência do

**Lema 3.4.6** [BaLe, Teorema 6] *Seja  $(x, K, u, \mu)$  um processo ótimo e  $\eta, z, \lambda$  as variáveis adjuntas correspondentes. Então  $\eta \neq 0$ .* ■

#### Análise das condições necessárias

O problema central da análise das condições necessárias consiste em descobrir condições iniciais  $\lambda_0 = \lambda(0)$ ,  $z_0 = z(0)$  que estejam de acordo com a condição

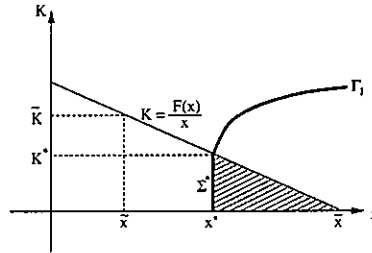
$$(R) \quad \lambda(t) \leq r, \text{ for all } t \in [0, \infty).$$

É de simples verificação o fato de que a igualdade  $\lambda(\tau) = r$  para algum  $\tau > 0$  implica em  $\lambda'(\tau) = 0$  e  $\lambda''(\tau) \leq 0$ . É igualmente claro que  $K_{0,+} \neq K_0$  somente ocorre quando  $\lambda(0) = r$ .

A seguir apresentamos um lema que discute propriedades de trajetórias ótimas seus correspondentes controles e variáveis adjuntas. Tal resultado é utilizado na seqüência para obter a solução do problema de controle ótimo (para demonstração veja [BaLe, Parágrafo 3]).

**Lema 3.4.7** *Seja  $(x, K, u, \mu)$  um processo ótimo com variáveis adjuntas  $z, \lambda$ . São verdadeiras as afirmações:*

- Para cada  $\tau > 0$  existe  $t > \tau$  com  $\lambda(t) = r$ ;
- Não existe  $\tau > 0$  com  $x(\tau) > x^*$  e  $F(x(\tau)) \geq K(\tau)x(\tau)$ ;
- Para  $x_0 = x^*$ ,  $K_0 > K^*$  as condições iniciais  $\lambda_0 = r$ ,  $z_0 \geq r'$  não podem ser assumidas;
- Se  $x_0 = x^*$ ,  $K_0 = K^*$ , então  $\mu = \gamma K^* dt$  e ainda  $x(t) = x^*$ ,  $K(t) = K^*$ ,  $\lambda(t) = r$ ,  $z(t) = r'$ ,  $u(t) = 1$ , para  $t \geq 0$ ;
- Para  $x_0 = x^*$ ,  $K_0 \leq K^*$ , temos necessariamente os valores iniciais  $\lambda_0 = r$ ,  $z_0 = r'$ ,  $\lambda'(0) = 0$ ,  $z'(0) = 0$ . Além disso temos  $K_{0,+} = K^*$ .

Figura 3.7: Curvas  $\Gamma_1$  e  $\Sigma^*$ 

f) Se  $x_0 > x^*$  e  $K_0 < F(x_0)/x_0$ , temos necessariamente:  $\lambda_0 = r$ ,  $z_0 > r'$ ,  $K_{0,+} > K_0$ .

O Lemma 3.4.7 possui algumas conseqüências interessantes:

— Do item b) concluímos que uma trajetória ótima não pode entrar na região hachurada da Figura 3.7.

— O item f) garante ainda que caso a condição inicial  $(x_0, K_0)$  pertença a região hachurada da Figura 3.7, então  $K_{0,+}$  é tal que  $(x_0, K_{0,+})$  está acima desta região, i.e. uma trajetória ótima com tal condição inicial salta no tempo  $t = 0$  para algum estado fora desta região.

A seguir definimos algumas curvas no plano de fase. Seja  $(x, K)$  a solução do sistema

$$\begin{pmatrix} x' \\ K' \end{pmatrix} = \begin{pmatrix} -F(x) + Kx \\ \gamma K \end{pmatrix}, \quad \begin{pmatrix} x(0) \\ K(0) \end{pmatrix} = \begin{pmatrix} x^* \\ K^* \end{pmatrix}.$$

Denotamos por  $\Gamma_1$  a curva definida por  $[0, \bar{t}] \ni t \mapsto (x(t), K(t)) \in [x^*, \bar{x}] \times [K^*, \infty)$ . A curva  $[0, K^*] \ni K \mapsto (x^*, K) \in [0, \infty) \times [0, K^*]$  é denotada por  $\Sigma^*$  (veja Figura 3.7).

Ainda uma conseqüência do Lemma 3.4.7: Caso a condição inicial  $(x_0, K_0)$  pertença a curva  $\Sigma^*$ , segue dos itens d) e e) que devemos saltar no instante  $t = 0$  para o estado  $(x^*, K^*)$  e lá permanecer (pescando com  $u = 1$ ) até o fim dos tempos.

No próximo resultado usamos as curvas definidas acima para definir uma região de saltos no plano de fase.

**Lema 3.4.8** *Seja  $(x, K, u, \mu)$  um processo ótimo com variáveis adjuntas  $z, \lambda$ . Se  $x_0 > x^*$  e a condição inicial  $(x_0, K_0)$  está abaixo da curva  $\Gamma_1$ , então as condições iniciais satisfazem:  $\lambda_0 = r$ ,  $z_0 > r'$ ,  $K_{0,+} \geq h_1(x_0)$ .<sup>7</sup>*

**Demonstração:** Veja [BaLe, Lema 14]. ■

Note que o Lemma 3.4.8 engloba o resultado do Lemma 3.4.7 f). A seguir verificamos que se uma trajetória ótima encontra a curva  $\Sigma^*$  (i.e.  $(x(\tau), K(\tau)) \in \Sigma^*$  para algum  $\tau > 0$ ), certas condições precisam ser satisfeitas pelas variáveis adjuntas.

<sup>7</sup>Notação:  $\{x^*, \bar{x}\} \ni x \mapsto (x, h_1(x))$  é uma parameterização de  $\Gamma_1$ .



**Lema 3.4.9** *Seja  $(x, K, u, \mu)$  um processo ótimo com variáveis adjuntas  $z, \lambda$ . Suponha que  $x(\sigma) = x^*, K(\sigma) \in (0, K^*)$  para algum  $\sigma > 0$ . Temos então  $\lambda(\sigma) = r, \lambda'(\sigma) = 0, z(\sigma) = r', z'(\sigma) = 0$ .*

**Demonstração:** Veja [BaLe, Lema 15]. ■

O Lemma 3.4.9 garante que caso um processo ótimo encontre a curva  $\Sigma^*$ , ele o faz com  $\lambda(\sigma) = r$ . Combinando este resultado com os itens *d)* e *e)* do Lemma 3.4.7, obtemos o comportamento de uma trajetória ótima após esta encontrar a curva  $\Sigma^*$ . O próximo e natural passo é analisar o comportamento das trajetórias ótimas que encontram a curva  $\Sigma^*$  (antes do encontro obviamente). Para tanto é necessário acompanhar a dinâmica da evolução das variáveis de estado *para atrás* no tempo.

**Lema 3.4.10** *Para cada  $K_1 \in [0, K^*)$ , seja  $(x, z)$  a solução do problema de valor inicial  $\begin{pmatrix} x' \\ z' \end{pmatrix} = \begin{pmatrix} -F(x) + K_1 e^{rx} \\ -(z - r')g(x) + \psi_*(x) \end{pmatrix}, \begin{pmatrix} x(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} x^* \\ r' \end{pmatrix}$ . Então, para cada  $K_1 \in (0, K^*)$  existe  $\tau := \tau_{K_1} > 0$  com  $x(\tau) = x^*$ . Além disso, existe  $\tilde{K}_1 \in (0, K^*)$  satisfazendo:*

- a) *Se  $K_1 \in (0, \tilde{K}_1)$ , então  $z$  possui um único zero  $\sigma \in (0, \tau)$ , onde  $z'(\sigma) < 0$  e  $x(\sigma) \in (0, \tilde{x})$ ;*
- b) *Se  $K_1 \in (\tilde{K}_1, K^*)$ , então  $z(t) > 0$  para todo  $t \in [0, \tau]$ ;*
- c) *Se  $K_1 = \tilde{K}_1$ , então existe um único  $\sigma$  em  $(0, \tau)$  com  $z(\sigma) = z'(\sigma) = 0$ ; além disso  $x(\sigma) = \tilde{x}$  e  $\tilde{K} := K_1 e^{r\sigma} > F(\tilde{x})/\tilde{x} = \tilde{K}$ .*

**Demonstração:** Veja [BaLe, Lema 16]. ■

Argumentando com o Lema 3.4.10 e o teorema da função implícita, construímos uma curva  $\Sigma_0$  tal que as soluções  $(x, K, z)$  do sistema

$$\begin{pmatrix} x' \\ z' \\ K' \end{pmatrix} = \begin{pmatrix} -F(x) + Kx \\ -(z - r')g(x) + \psi_*(x) \\ -\gamma K \end{pmatrix}, \quad \begin{pmatrix} x(0) \\ z(0) \\ K(0) \end{pmatrix} = \begin{pmatrix} x^* \\ r' \\ K_1 \end{pmatrix},$$

onde  $K_1 \in [0, K^*)$ , satisfazem uma das seguintes alternativas:

- (i) *Se  $K_1 \in (\tilde{K}_1, K^*)$ , então  $z(t) > 0$ , para todo  $t \geq 0$  (veja curva  $\gamma_1$  na Figura 3.8);*
- (ii) *Se  $K_1 = \tilde{K}_1$ , então  $z(t) > 0$  exceto em um único instante de tempo, quando  $(x, K) = (\tilde{x}, \tilde{K})$  (veja curva  $\gamma_2$  na Figura 3.8);*
- (iii) *Se  $K_1 \in [0, \tilde{K}_1)$ , então  $z(t) > 0$  antes da trajetória interceptar  $\Sigma_0$  e  $z(t) < 0$  após a interceptação (veja curva  $\gamma_3$  na Figura 3.8).*

Fica assim claro comportamento das trajetórias ótimas antes do encontro com  $\Sigma^*$ , em uma vizinhança de  $\Sigma_0$ . A fim de analisar as demais condições iniciais, precisamos definir ainda algumas curvas no plano de fase. Seja  $(x, K)$  a solução do problema de valor inicial  $\begin{pmatrix} x' \\ K' \end{pmatrix} = \begin{pmatrix} -F(x) + Kx \\ \gamma K \end{pmatrix}, \begin{pmatrix} x(0) \\ K(0) \end{pmatrix} = \begin{pmatrix} \tilde{x} \\ \tilde{K} \end{pmatrix}$ . Denotamos por  $\Gamma_3$  a curva  $[0, \tilde{t}] \ni t \mapsto (x(t), K(t)) \in [\tilde{x}, \tilde{x}] \times [\tilde{K}, \infty)$  (veja Figura 3.9). Seja agora  $(x, K)$  a solução do problema de valor inicial  $\begin{pmatrix} x' \\ K' \end{pmatrix} = \begin{pmatrix} F(x) - Kx \\ -\gamma K \end{pmatrix}, \begin{pmatrix} x(0) \\ K(0) \end{pmatrix} = \begin{pmatrix} \tilde{x} \\ \tilde{K} \end{pmatrix}$ . Esta solução encontra a curva  $\Sigma^*$  em  $(x^*, \tilde{K}_1)$  para algum  $\tau > 0$  (veja Lema 3.4.10). A curva definida por esta trajetória é denominada  $\Gamma_2$  (veja

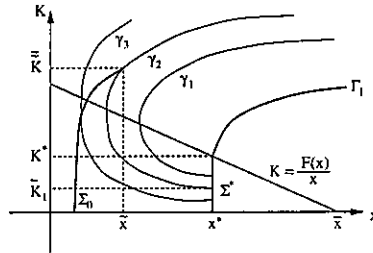
Figura 3.8: Curva  $\Sigma_0$ 

Figura 3.9). Por fim definimos a curva  $\Gamma_4$ , a qual é parametrizada pela solução  $(x, K)$  de  $(x', K') = \begin{pmatrix} -F(x) \\ \gamma K \end{pmatrix}$ ,  $(x(0), K(0)) = \begin{pmatrix} \bar{x} \\ \bar{K} \end{pmatrix}$ .

Analizamos novamente a região  $x > x^*$ , mais especificamente a parte acima da curva  $\Gamma_1$ . O objetivo é determinar uma curva  $\Sigma_s$ , acima de  $\Gamma_1$ , sobre a qual devem *saltar* as trajetórias com condição inicial contida na região de salto. Argumentando como na demonstração do Lema 3.4.10 obtemos o seguinte resultado:

**Lema 3.4.11** *Existe  $\tilde{K}_2 \in (0, \tilde{K}_1)$  e  $a > 0$ , tais que a solução  $(x(\cdot; K_1), z(\cdot; K_1), \lambda(\cdot; K_1))$  do sistema  $\begin{pmatrix} x' \\ z' \\ \lambda' \end{pmatrix} = \begin{pmatrix} -F(x) + K_1 z^{\gamma} x \\ -(z - r')g(x) + \psi_*(z) \\ -(\lambda - r')\alpha + z - r' \end{pmatrix}$ ,  $\begin{pmatrix} x(0) \\ z(0) \\ \lambda(0) \end{pmatrix} = \begin{pmatrix} x^* \\ r' \\ r \end{pmatrix}$  existe em  $[0, a]$  para cada  $K_1 \in (\tilde{K}_2, K^*)$ . Além disso, para cada  $K_1 \in (\tilde{K}_2, K^*)$  existem escalares  $\rho(K_1)$ ,  $\sigma(K_1)$ ,  $\tau(K_1)$  satisfazendo*

- $0 < \rho(K_1) < \sigma(K_1) < \tau(K_1) < a$ ;
- $x(t; K_1) \leq x^*$ ,  $t \in (0, \rho(K_1))$ ,  $x(\rho(K_1); K_1) = x^*$ ,  $x(t; K_1) > x^*$ ,  $t \in (\rho(K_1), a]$ ,  $x(a; K_1) \geq \bar{x}$ ;
- $0 < z(t; K_1) < r'$ ,  $t \in (0, \sigma(K_1))$ ,  $z(\sigma(K_1); K_1) = r'$ ,  $z(t; K_1) > r'$ ,  $t \in (\sigma(K_1), a]$ ;
- $\lambda(t; K_1) < r$ ,  $t \in (0, \tau(K_1))$ ,  $\lambda(\tau(K_1); K_1) = r$ ,  $\lambda'(\tau(K_1); K_1) > 0$ ;
- $\lim_{K_1 \uparrow K^*} \tau(K_1) = 0$ .

**Demonstração:** Veja [BaLe, Lema 19]. ■

Argumentando com o Lema 3.4.11 e o teorema da função implícita, é possível construir uma curva  $\Sigma_s$  no plano  $(x, K)$ , tal que as soluções  $(x, z, \lambda)$  do sistema no Lema 3.4.11 satisfazem:

- A trajetória  $(x(\cdot; K_1), z(\cdot; K_1))$  intercepta a curva  $\Sigma_s$  com  $\lambda = r$ ,  $\lambda' > 0$  e  $z > r'$ , para  $K_1 \in (\tilde{K}_2, K^*)$ ;
- A curva  $\Sigma_s$  começa em  $(x^*, K^*)$  e alcança a reta vertical  $x = x^*$ ;
- A curva  $\Sigma_s$  fica estritamente acima de  $\Gamma_1$ .

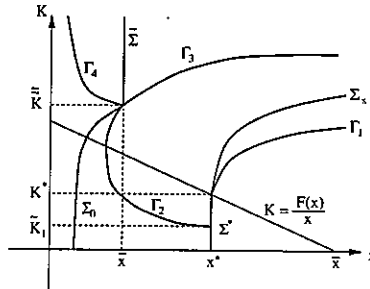


Figura 3.9: Cenário para o problema da pescaria ótima

A curva  $\Sigma_s$  conecta o ponto  $x^*$  ao ponto  $\bar{x}$  e deve ser usada como *curva de salto* na região  $[x^*, \bar{x}] \times [0, \infty)$ . Esta curva de salto nos permite encontrar, dado  $(x_0, K_0)$  com  $x_0 \in (x^*, \bar{x})$  e  $K_0 < h_s(x_0)$ , o valor inicial ótimo  $K_{0,+} := h_s(K_0)$ .<sup>8</sup>

Observe que na vizinhança das curvas  $\Gamma_1, \Sigma^*$ , e  $\Sigma_s$  o comportamento dos extremais já foi determinado. Note ainda que a construção das curvas  $\Sigma_s$  e  $\Gamma_3$  possibilita distinguir dois casos distintos:  $\Sigma_s$  e  $\Gamma_3$  possuem um ponto em comum;  $\Sigma_s$  e  $\Gamma_3$  não se interceptam.

Foram construídas em  $[0, \bar{x}] \times [0, \infty)$  as curvas  $\Sigma^*, \tilde{\Sigma}, \Sigma_s, \Sigma_0$  assim como as trajetórias  $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ . Nosso próximo passo é obter, para cada condição inicial  $(x_0, K_0) \in (0, \bar{x}) \times (0, \infty)$ , a trajetória ótima correspondente.

### Trajетórias ótimas

A seguir determinamos um cenário de todas as trajetórias ótimas no plano de fase. Também são determinados os controles ótimos correspondentes, baseado em informações fornecidas pelos switches  $z$  e  $\lambda$ . Começamos definindo as regiões determinadas em  $[0, \bar{x}] \times [0, \infty)$  por  $\Sigma^*, \tilde{\Sigma}, \Sigma_s, \Sigma_0, \Gamma_1, \Gamma_2, \Gamma_3$  e  $\Gamma_4$  (na Figura 3.10 são esboçadas as cinco regiões principais para o caso da função logística):

**Domínio (R1):** fronteiras  $\Sigma^*, \Sigma_s, \{(x, K) \in [0, \bar{x}] \times [0, \infty) \mid x = \bar{x}\}$ ;

**Domínio (R2):** fronteiras  $\Sigma_0, \Sigma^*, \Sigma_s, \Gamma_3$  e eventualmente  $\{(x, k) \in [0, \bar{x}] \times [0, \infty) \mid x = \bar{x}\}$ ;

**Domínio (R3):** fronteiras  $\{(x, K) \in [0, \bar{x}] \times [0, \infty) \mid x = 0\}, \Sigma_0, \Gamma_4$ ;

**Domínio (R4):** fronteiras  $\{(x, K) \in [0, \bar{x}] \times [0, \infty) \mid x = 0\}, \Gamma_4, \tilde{\Sigma}$ ;

**Domínio (R5):** fronteiras  $\tilde{\Sigma}, \Gamma_3, \{(x, K) \in [0, \bar{x}] \times [0, \infty) \mid x = \bar{x}\}$ .

O próximo teorema fornece uma completa descrição das trajetórias ótimas (e correspondentes controles e switches) para todas as possíveis condições iniciais no plano de fase.

**Teorema 3.4.12** *Seja  $(x, K, u, \mu)$  um processo ótimo com variáveis adjuntas  $z, \lambda$ . São verdadeiras as afirmações:*

<sup>8</sup>Notação:  $x \mapsto (x, h_s(x))$  é uma parametrização local de  $\Sigma_s$ .

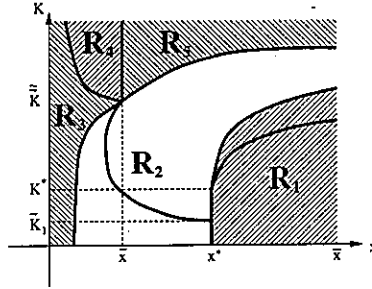


Figura 3.10: Regiões principais do plano de fase

- a) Se  $(x_0, K_0)$  pertence a  $(R1)$ , então  $z_0 > r'$ ,  $\lambda_0 = r$ ,  $(x_0, K_{0,+}) \in \Sigma_s$  e  $\exists \tau > 0$  com  $(x(\tau), K(\tau)) \in \Sigma^*$  e  $u \equiv 1$ ,  $\mu \equiv 0$  em  $(0, \tau)$ ;
- b) Se  $(x_0, K_0)$  pertence a  $(R2)$ , então  $z_0 > 0$ ,  $\lambda_0 < r$  e  $\exists \tau > 0$  com  $(x(\tau), K(\tau)) \in \Sigma^*$  e  $u \equiv 1$ ,  $\mu \equiv 0$  em  $(0, \tau)$ ;
- c) Se  $(x_0, K_0)$  pertence a  $(R3)$ , então  $z_0 < 0$ ,  $\lambda_0 < r$  e  $\exists \tau > 0$  com  $(x(\tau), K(\tau)) \in \Sigma_0$  e  $u \equiv 0$ ,  $\mu \equiv 0$  em  $(0, \tau)$ ;
- d) Se  $(x_0, K_0)$  pertence a  $(R4)$ , então  $z_0 < 0$ ,  $\lambda_0 < r$  e  $\exists \tau > 0$  com  $(x(\tau), K(\tau)) \in \tilde{\Sigma}$  e  $u \equiv 0$ ,  $\mu \equiv 0$  em  $(0, \tau)$ ;
- e) Se  $(x_0, K_0)$  pertence a  $(R5)$ , então  $z_0 > 0$ ,  $\lambda_0 < r$  e  $\exists \tau > 0$  com  $(x(\tau), K(\tau)) \in \tilde{\Sigma}$  e  $u \equiv 1$ ,  $\mu \equiv 0$  em  $(0, \tau)$ .

**Demonstração:** Veja [BaLe, Lema 27]. ■

Se a condição inicial  $(x_0, K_0)$  pertence a uma das curvas  $\Gamma_3, \Gamma_4, \Sigma^*, \Sigma_0, \tilde{\Sigma}$ , o processo ótimo é obtido argumentando-se como no Teorema 3.4.12 (veja [BaLe, Lemas 28 e 29]).

A interpretação do Teorema 3.4.12 é simples. Por exemplo, se  $(x_0, K_0)$  pertence a  $(R1)$ , escolhemos  $K_{0,+}$  tal que  $(x_0, K_{0,+}) \in \Sigma_s$ ; prosseguimos pescando com  $u \equiv 1$  e sem investir ( $\mu \equiv 0$ ) até a trajetória  $(x, K)$  encontrar a curva  $\Sigma^*$ ; saltamos uma segunda (e última) vez para alcançar o estado  $(x^*, K^*)$  e permanecemos neste estado (pescando com  $u \equiv 1$  e investindo  $\mu \equiv \gamma K^* dt$ ) até o fim dos tempos.

Para certas condições iniciais é necessário decretar uma moratória ao longo de determinados trechos da trajetória. Por exemplo, se  $(x_0, K_0)$  pertence a  $(R3)$ , não se pesca nem investe ( $u \equiv 0$ ,  $\mu \equiv 0$ ) até que a trajetória alcance a curva  $\Sigma_0$ ; a partir deste instante pesca-se com  $u \equiv 1$  (ainda sem investir) até que a trajetória alcance a curva  $\Sigma^*$ ; saltamos então para alcançar o estado  $(x^*, K^*)$  e permanecemos neste estado (pescando com  $u \equiv 1$  e investindo  $\mu \equiv \gamma K^* dt$ ) até o fim dos tempos.

Nos dois exemplos anteriores a estratégia ótima correspondente a pesca (controle  $u$ ) é do tipo bang-bang. Entretanto, este não é sempre o caso. Uma interessante situação ocorre quando  $(x_0, K_0)$  pertence a  $(R4)$ . Neste caso é decretada inicialmente moratória ( $u \equiv 0$ ,  $\mu \equiv$

0) até que a trajetória alcance a curva  $\tilde{\Sigma}$ ; a partir daí pesca-se com  $u(t) = K(t)^{-1}F(\tilde{x})\tilde{x}^{-1}$  e não se investe ( $\mu \equiv 0$ ) até que a trajetória alcance o estado  $(\tilde{x}, \tilde{K})$ ; pesca-se então com  $u \equiv 1$  (ainda sem investir) até alcançar a curva  $\Sigma^*$  (este trecho da trajetória corresponde à curva  $\Gamma_2$ ); agimos agora como nos exemplos anteriores.

### 3.5 Exercícios

3.1 Considere o problema de Bolza

$$\begin{cases} \text{Minimizar } \frac{1}{2} \int_0^1 u(t)^2 dt + z_1(1)^2 + z_2(1)^2 \\ \text{sujeito a} \\ u \in L^1[0, 1], z'_1 = z_2, z'_2 = u, z_1(0) = z_2(0) = 0. \end{cases}$$

- a) Obtenha condições necessárias para o problema acima.  
b) Encontre o processo ótimo.

3.2 (Um problema de Investimento) Suponha que um determinado produto é fabricado com a taxa  $z(t)$ . No tempo  $t > 0$  uma fração  $u(t)$  da produção é reinvestida para aumentar a produção, sendo o restante vendido para geração de lucro. O objetivo é determinar uma política de investimento ótima, de forma a maximizar o lucro total no horizonte fixo de tempo  $[0, T]$ . Temos assim o problema

$$\begin{cases} \text{Maximizar } \int_0^T (1 - u(t))z(t)dt \\ \text{sujeito a} \\ z' = uz, z(0) = z_0 > 0, z(t) \geq 0, u \in \tilde{C}[0, T] \end{cases}$$

- a) Reescreva o problema como um problema variacional com restrições lagrangeanas:  $y'(t) \geq 0, y'(t) \leq y(t)$ .  
b) Obtenha condições necessárias para o novo problema.  
c) Encontre a taxa ótima de produção  $\bar{y}$ .

3.3 (Problema do Café) Uma xícara cheia de café a temperatura de  $100^\circ\text{C}$  deve ser esfriada a temperatura de  $0^\circ\text{C}$  por adição de uma quantidade fixa de creme de leite. Uma equação aproximada para evolução da temperatura  $z$  da mistura é dada por

$$z' = -z - 25u - uz/4.$$

As condições de contorno são  $z(0) = 100, z(T) = 0$ .

- a) Obtenha condições necessárias para o problema de tempo ótimo sujeito as restrições  $0 \leq u(t) \leq 1, t \in [0, T], \int_0^T u(t)dt = 1$ , impostas ao fluxo externo de líquido  $u$ .  
b) Use o fato  $z' < 0$  para obter um problema equivalente com intervalo de tempo fixo. O

que se pode afirmar sobre a unicidade da solução obtida no item *a*).

(Dica: A nova variável livre é  $s = z$ .)

## Capítulo 4

# Demonstração do Princípio do Máximo

Apresentamos neste capítulo uma demonstração do princípio de Pontryagin para problemas com horizonte finito e tempo final desconhecido. O desenvolvimento apresentado aqui se baseia nas notas de aula de M. Brokate (veja [Br]) e se divide em três partes principais:

- No Parágrafo 4.1 obtemos um conjunto de condições necessárias para otimalidade de soluções de um problema abstrato de otimização em espaços de dimensão infinita (os detalhes são discutidos no Apêndice A);
- No Parágrafo 4.2 aplicamos os resultados do item anterior a um problema auxiliar, construído a partir do problema de controle ótimo através de uma reparametrização da variável de tempo;
- No Parágrafo 4.3 invertemos esta parametrização, a fim de reinterpretar as condições necessárias obtidas para o problema auxiliar e obter as desejadas condições necessárias para o problema de controle ótimo.

Citamos [FIRi], [LiYo] e [Tr] como literatura alternativa, onde podem ser encontradas outras demonstrações (recentes) do princípio do máximo para problemas com horizonte finito e tempo final desconhecido.

### 4.1 Otimização Infinita

Este parágrafo é dedicado à análise de condições necessárias para otimalidade. Os problemas de otimização estudados são formulados em espaços de dimensão infinita e envolvem funções continuamente diferenciáveis. Detalhes sobre o desenvolvimento apresentado neste parágrafo são discutidos no Apêndice A.

Considere o seguinte problema abstrato de otimização:

$$(P) \quad \begin{cases} \text{Minimizar } J(x) \\ \text{sujeito a } x \in C, F(x) \in K, \end{cases}$$

onde  $X, Y$  são espaços de Banach;  $O \subset X$  aberto;  $C \subset X, K \subset Y$  são fechados e convexos;  $J : O \rightarrow \mathbb{R}, F : O \rightarrow Y$  são continuamente diferenciáveis.

Na teoria de controle ótimo, a condição  $x \in C$  está associada a uma limitação no controle (e.g.  $u(t) \in \Omega$  q.s.) e a restrição  $F(x) \in K$  representa a equação diferencial (ou integral), assim como outras equações envolvendo a variável de estado (e.g.  $\psi(t_1, z(t_1)) = \theta$ ).

A forma desejada do teorema de multiplicadores que procuramos é a seguinte:

**Conjectura: [Lagrange generalizado]** *Se  $\bar{x}$  é um mínimo local do problema (P), então existem  $\eta \in \mathbb{R}, \psi \in X^*, \phi \in Y^*$  tais que*

$$i) \eta dJ(x_*) - \psi \circ dF(x_*) = \phi;$$

$$ii) \eta \geq 0, \eta + \|\psi\| + \|\phi\| \neq 0;$$

onde  $\psi$  e  $\phi$  satisfazem eventualmente outras restrições com relação aos conjuntos  $C$  e  $K$  (de preferência com  $\eta \neq 0$ ).

Para obter o resultado desejado é necessário aproximar localmente o conjunto dos pontos admissíveis  $X_{ad} := C \cap F^{-1}(K) \cap O$  para o problema (P) por conjuntos convexos apropriados e, posteriormente, aplicar um teorema de separação para conjuntos convexos em espaços de Banach.

Antes de apresentarmos os resultados principais deste parágrafo, esclarecemos a notação utilizada no texto subsequente. Dado  $X$  um espaço de Banach,  $C \subset X, x \in C, y \in X_{ad}$ , denotamos por

- $C(x)$  o cone tangencial em  $x$  por  $C$  (somente para  $C$  convexo); (veja Definição A.2.1)
- $T(C, x)$  o cone tangencial a  $C$  por  $x$ ; (veja Definição A.2.2)
- $C^*$  o cone dual a  $C$  em  $X^*$ . (veja Definição A.3.1)
- $L(X_{ad}, y)$  o cone linearizado em  $y$  para o problema (P) (veja Definição A.2.9)

A seguir enunciamos os teoremas de multiplicadores usados na demonstração do princípio do máximo. Demonstrações completas destes resultados são fornecidas no Parágrafo A.3.

**Teorema 4.1.1** *Seja  $x_*$  um mínimo local de (P). Se  $x_*$  é um ponto regular para (P), então existem  $\psi \in C(x_*)^*, \phi \in K(F(x_*))^*$  tais que*

$$dJ(x_*) - \phi \circ dF(x_*) = \psi. \quad (4.1)$$

A equação (4.1) é denominada **condição de Kuhn-Tucker** (veja Definição A.2.5 para o conceito de ponto regular).

**Teorema 4.1.2** *Seja  $x_*$  um mínimo local de (P). Se  $x_*$  é um ponto fracamente regular para (P), então existem  $\eta \geq 0, \psi \in C(x_*)^*, \phi \in K(F(x_*))^*$  tais que*

$$\eta dJ(x_*) - \phi \circ dF(x_*) = \psi, \quad (4.2)$$

onde  $|\eta| + \|\psi\| + \|\phi\| \neq 0$ .



A equação (4.2) é denominada **condição de Fritz–John** (veja Definição A.2.5 para o conceito de ponto fracamente regular).

**Observação 4.1.3** Para problemas de otimização da forma especial

$$\begin{cases} \text{Minimizar } J(x) \\ \text{s.a. } x \in \mathbb{R}^n \\ g_i(x) \leq 0, \quad i = 1, \dots, m \\ h_j(x) = 0, \quad j = 1, \dots, l \end{cases}$$

é possível obter uma condição suficiente (**condição de Slater**) para que uma solução  $x_*$  seja sempre regular (veja [We]). Neste caso obtemos um teorema de multiplicadores com  $\eta > 0$  (compare com a conjectura discutida na página 82).  $\square \square \square$

## 4.2 Um Problema Auxiliar

A fim de simplificar a notação, consideramos o problema  $P(t_0, z_0)$  – veja Parágrafo 3.1 – com tempo inicial  $t_0 = 0$  e tempo final  $t_1 = T$ . Temos assim o seguinte problema de controle ótimo

$$P(0, z_0) \quad \begin{cases} \text{Minimizar } J(z, u) := L_1(T, z(T)) + \int_0^T L(t, z(t), u(t)) dt \\ \text{sujeito a} \\ T \geq 0; \quad u \in L^1([0, T]; \mathbb{R}^m), \quad u(t) \in \Omega \text{ q.s. em } [0, T]; \\ z(t) = z_0 + \int_0^t f(s, z(s), u(s)) ds, \quad t \in [0, T]; \quad \psi(T, z(T)) = \theta \end{cases}$$

onde as funções  $L$ ,  $f$ ,  $L_1$ ,  $\psi$  e  $H$  são definidas como no Parágrafo 3.1. Seja agora  $(z^*, u^*, T^*)$  uma solução qualquer para o problema  $P(0, z_0)$ .

Neste parágrafo, obtemos através de uma mudança de variáveis denominada *transformação no tempo* um problema auxiliar. Tal problema nos permite investigar o problema  $P(0, z_0)$  sob a ótica do princípio do máximo. Começamos por definir o conjunto

$$V_+ := \{v \in L^\infty(0, 1) \mid v(\tau) \geq 0 \text{ q.s.}\}.$$

Dada agora uma função  $v \in V_+$ , definimos a aplicação  $t \in C[0, 1]$  e o escalar  $T \geq 0$  da seguinte forma:

$$t(\tau) := \int_0^\tau v(s) ds, \quad \tau \in [0, 1], \quad T := t(1). \quad (4.3)$$

Observe que para cada trajetória  $z \in C([0, T]; \mathbb{R}^n)$ , é possível associar a função  $x \in C([0, 1]; \mathbb{R}^n)$  definida pela correspondência

$$x(\tau) := z(t(\tau)), \quad \tau \in [0, 1]. \quad (4.4)$$

Seja ainda  $v \in V_+$ . Note que, uma vez fixada  $\bar{w} \in L_{loc}^\infty([0, \infty); \mathbb{R}^m)$  com  $\bar{w}(t) \in \Omega$  q.s., para cada controle  $u \in L^\infty([0, T]; \mathbb{R}^m)$  corresponde o controle  $w \in L^\infty([0, 1]; \mathbb{R}^m)$  da forma

$$w(\tau) := \begin{cases} u(t(\tau)), & \text{se } \tau \in \{s \in [0, 1] \mid v(s) > 0\} \\ \bar{w}(t(\tau)), & \text{se } \tau \in \{s \in [0, 1] \mid v(s) = 0\} \end{cases} \quad (4.5)$$

(a escolha da função  $\bar{w}$  é esclarecida no Parágrafo 4.3). Formulamos a seguir um problema de otimização auxiliar nas variáveis  $t, x, v$ .

$$PA(w_*) \begin{cases} \text{Minimizar } I(t, x, v) := L_1(t(1), x(1)) + \int_0^1 v(\tau) L(t(\tau), x(\tau), w_*(\tau)) d\tau \\ \text{sujeito a} \\ x \in C([0, 1]; \mathbb{R}^n), t \in C([0, 1]; \mathbb{R}), v \in V_+, \psi(t(1), x(1)) = \theta, \\ x(\tau) = z_0 + \int_0^\tau v(s) f(t(s), x(s), w_*(s)) ds, \tau \in [0, 1], \\ t(\tau) = \int_0^\tau v(s) ds, \tau \in [0, 1], \end{cases}$$

onde  $w_* \in L^\infty([0, 1]; \mathbb{R}^m)$  é dado (a escolha do "parâmetro"  $w_*$ , função correspondente ao controle ótimo  $u^*$ , é esclarecida no decorrer do texto). Verificamos inicialmente que toda solução  $(z^*, u^*, T^*)$  do problema  $P(0, z_0)$  induz uma solução  $(t_*, x_*, v_*)$  do problema  $PA(w_*)$ .

**Lema 4.2.1** *Seja  $(z^*, u^*, T^*)$  uma solução de  $P(0, z_0)$ . Sejam ainda  $w_* \in L^\infty$  e funções  $t_*, x_*, v_*$  satisfazendo (4.3), (4.4) e (4.5). Então  $(t_*, x_*, v_*)$  é mínimo local de  $PA(w_*)$ .*

**Demonstração:** Verificamos inicialmente que  $t_*, x_*, v_*$  satisfazem as restrições de  $PA(w_*)$ .

- $$\begin{aligned} x_*(\tau) &= z^*(t(\tau)) = z_0 + \int_0^{t_*(\tau)} f(s, z^*(s), u^*(s)) ds \\ &= z_0 + \int_0^\tau f(t_*(r), z^*(t_*(r)), u^*(t_*(r))) v_*(r) dr \\ &= z_0 + \int_0^\tau v_*(r) f(t_*(r), x_*(r), w_*(r)) dr; \end{aligned}$$
- $\psi(t_*(1), x_*(1)) = \psi(T^*, z^*(T^*)) = \theta;$

Note ainda que  $I(t_*, x_*, v_*) = J(z^*, u^*, T^*)$ . De fato,

$$\begin{aligned} I(t_*, x_*, v_*) &= L_1(t_*(1), x_*(1)) + \int_0^1 v_*(\tau) L(t_*(\tau), x_*(\tau), w_*(\tau)) d\tau \\ &= L_1(T^*, z^*(T^*)) + \int_0^{T^*} L(s, z^*(s), u^*(s)) ds \\ &= J(z^*, u^*, T^*). \end{aligned}$$

Definimos a vizinhança  $O$  de  $(t_*, x_*)$  da seguinte forma:

$$O := \{(t, x) \in C([0, 1]; \mathbb{R}^n) \times C([0, 1]; \mathbb{R}) \mid \|t - t_*\|_\infty < 1, \|x - x_*\|_\infty < 1\}.$$

Seja agora  $(t, x) \in O$  satisfazendo as restrições de  $PA(w_*)$ . Definimos as funções  $z \in L^\infty([0, t(1)]; \mathbb{R}^n)$  e  $u \in L^\infty([0, t(1)]; \mathbb{R}^m)$  por

$$z(s) := x(\tau(s)), \quad u(s) := w_*(\tau(s)),$$

onde  $\tau(s) := \inf\{\tau \in [0, 1] \mid t(\tau) = s\}$ .<sup>1</sup> Temos assim

$$z(t(\tau)) = x(\tau), \quad \forall \tau \in [0, 1], \quad u(t(\tau)) = w_*(\tau), \quad \forall \tau \text{ com } v(\tau) > 0.$$

Com uma simples mudança de variável nas integrais que surgem em (4.3) – (4.5) verificamos que  $(z, u, t(1))$  satisfaz as restrições do problema  $P(0, z_0)$ . Logo, a trajetória  $z$  é contínua, de onde concluímos que

$$J(z, u, t(1)) = I(t, x, v).$$

A otimalidade de  $(z^*, u^*, T^*)$  implica agora em

$$I(t, x, v) = J(z, u, t(1)) \geq J(z^*, u^*, T^*) = I(t_*, x_*, v_*),$$

ficando assim provado o teorema. (Usamos tacitamente na demonstração o fato de que, caso  $\Omega$  seja ilimitado, então  $L$  e  $f$  satisfazem

$$\sup\{|L(t, x, w_*(\cdot))|; t \in [0, t^* + 1], x \in \mathbb{R}^n\} \leq m(\cdot),$$

$$\sup\{|f(t, x, w_*(\cdot))|; t \in [0, t^* + 1], x \in \mathbb{R}^n\} \leq m(\cdot),$$

para alguma função  $m \in L^1[0, 1]$ .) ■

Antes de prosseguirmos com a análise dos extremais do problema  $PA(w_*)$ , é necessário esclarecer a seguinte questão: A equação integral para  $x$  em  $PA(w_*)$  define um operador continuamente diferenciável. Este resultado é verificado no lema a seguir, que trata a forma geral do operador em questão.

**Lema 4.2.2** *Seja  $g : [0, 1] \times \mathbb{R}^l \rightarrow \mathbb{R}^l$  mensurável no primeiro argumento e duas vezes continuamente diferenciável no segundo. Seja ainda  $\hat{x} \in C([0, 1]; \mathbb{R}^l)$ . Suponha que existe  $m \in L^1[0, 1]$  tal que para todo  $\tau \in [0, 1]$  e  $y \in \mathbb{R}^n$  com  $|y| \leq \|\hat{x}\|_\infty + 1$  tenhamos*

$$\max\{|g(\tau, y)|, |D_x g(\tau, y)|, |D_{xx} g(\tau, y)|\} \leq m(\tau).$$

*Seja  $O$  uma vizinhança aberta de  $\hat{x}$ . A aplicação  $G : O \subset L^\infty([0, 1]; \mathbb{R}^l) \rightarrow C([0, 1]; \mathbb{R}^l)$  definida por*

$$(Gx)(\tau) := \int_0^\tau g(s, x(s)) ds \tag{4.6}$$

<sup>1</sup>O conjunto  $\{\tau \in [0, 1] \mid t(\tau) = s\}$  é um intervalo fechado para todo  $s \in [0, 1]$ .

é continuamente diferenciável em  $O$  e sua derivada (de Fréchet) em  $x \in O$  é dada por

$$(DG(x)(h))(\tau) = \int_0^\tau D_x g(s, x(s))h(s) ds, \quad \tau \in [0, 1]. \quad (4.7)$$

**Demonstração:** As hipóteses do lema garantem que o operador  $G$  em (4.6) está bem definido em uma vizinhança  $O \subset L^\infty([0, 1]; \mathbb{R}^l)$  de  $\hat{x}$ . Temos ainda que o operador no lado direito de (4.7)

$$L^\infty([0, 1]; \mathbb{R}^l) \ni h \mapsto \int_0^\tau D_x g(s, x(s))h(s) ds \in C([0, 1]; \mathbb{R}^l)$$

é linear e contínuo se  $x \in O$ . Portanto, para provar o lema basta verificar (4.7) em um único ponto  $x \in O$ , por exemplo em  $x = \hat{x}$ .

Seja  $h \in L^\infty([0, 1]; \mathbb{R}^l)$  com  $\|h\|_\infty < 1$ . Então, para todo  $\tau \in [0, 1]$  temos

$$\begin{aligned} |G(\hat{x} + h)(\tau) - G(\hat{x})(\tau) - \int_0^\tau D_x g(s, \hat{x}(s))h(s) ds| \\ \leq \int_0^1 |g(s, \hat{x}(s) + h(s)) - g(s, \hat{x}(s)) - D_x g(s, \hat{x}(s))h(s)| ds \\ \leq \int_0^1 |h(s)|^2 \sup\{|D_{xx}g(s, \xi)|; |\xi - \hat{x}(s)| \leq |h(s)|\} ds \\ \leq \|h\|_\infty^2 \|m\|_1. \end{aligned}$$

Portanto,

$$\lim_{\|h\|_\infty \rightarrow 0} \left\{ \frac{\|G(\hat{x} + h) - G(\hat{x}) - \int_0^\tau D_x g(s, \hat{x}(s))h(s) ds\|_\infty}{\|h\|_\infty} \right\} = 0,$$

provando assim que  $(DG(\hat{x})(h))(\tau) = \int_0^\tau D_x g(s, \hat{x}(s))h(s) ds$ . ■

Uma vez esclarecida esta questão sobre a equação integral para  $x$  em  $PA(w_*)$ , podemos enfim aplicar o teorema de multiplicadores 4.1.1 ao problema  $PA(w_*)$ .

**Lema 4.2.3** *Seja  $(t_*, x_*, v_*)$  um mínimo local do problema auxiliar  $PA(w_*)$  e  $T^* = t_*(1)$ . Suponha que para  $t \in [0, T^* + 1]$  e  $x \in \mathbb{R}^n$  com  $|x| \leq \|x_*\|_\infty + 1$  tenhamos*

$$\begin{aligned} |L(t, x, w_*(\cdot))|, |D_t L(t, x, w_*(\cdot))|, |D_x L(t, x, w_*(\cdot))| &\leq m(\cdot), \\ |D_{tt} L(t, x, w_*(\cdot))|, |D_{tx} L(t, x, w_*(\cdot))|, |D_{xx} L(t, x, w_*(\cdot))| &\leq m(\cdot), \\ |f(t, x, w_*(\cdot))|, |D_t f(t, x, w_*(\cdot))|, |D_x f(t, x, w_*(\cdot))| &\leq m(\cdot), \\ |D_{tt} f(t, x, w_*(\cdot))|, |D_{tx} f(t, x, w_*(\cdot))|, |D_{xx} f(t, x, w_*(\cdot))| &\leq m(\cdot), \end{aligned} \quad (4.8)$$

com  $m \in L^1[0, 1]$ . Então existem  $\eta \geq 0$ ,  $\mu \in \mathbb{R}^p$  e funções absolutamente contínuas  $q : [0, 1] \rightarrow \mathbb{R}^n$ ,  $r : [0, 1] \rightarrow \mathbb{R}$ , tais que  $(\eta, \mu, r, q) \neq \theta$ ,

$$\begin{cases} \frac{dr}{d\tau}(\tau) = -v_*(\tau) [D_t f(t_*(\tau), x_*(\tau), w_*(\tau))^* q(\tau) + \eta D_t L(t_*(\tau), x_*(\tau), w_*(\tau))], \\ \frac{dq}{d\tau}(\tau) = -v_*(\tau) [D_x f(t_*(\tau), x_*(\tau), w_*(\tau))^* q(\tau) + \eta D_x L(t_*(\tau), x_*(\tau), w_*(\tau))], \end{cases} \quad (4.9)$$

para todo  $\tau \in [0, 1]$ ,

$$\begin{cases} r(1) = \eta D_t L_1(t_*(1), x_*(1)) - D_t \psi(t_*(1), x_*(1))^* \mu, \\ q(1) = \eta D_x L_1(t_*(1), x_*(1)) - D_x \psi(t_*(1), x_*(1))^* \mu, \end{cases} \quad (4.10)$$

e ainda

$$\int_0^1 [r(\tau) + f(t_*(\tau), x_*(\tau), w_*(\tau))^* q(\tau) + \eta L(t_*(\tau), x_*(\tau), w_*(\tau))] (v(\tau) - v_*(\tau)) d\tau \geq 0, \quad (4.11)$$

para todo  $v \in V_+$ .

**Demonstração:** A fim de escrever o problema  $PA(w_*)$  na forma do problema abstrato de otimização  $(P)$  introduzido no Parágrafo 4.1, definimos

$$\begin{aligned} X &:= C([0, 1]; \mathbb{R}) \times C([0, 1]; \mathbb{R}^n) \times L^\infty[0, 1], \quad J := I, \\ Y &:= C([0, 1]; \mathbb{R}) \times C([0, 1]; \mathbb{R}^n) \times \mathbb{R}^p, \quad K := \{(\theta, \theta)\} \subset Y, \\ C &:= \{(t, x, v) \in X \mid v(t) \geq 0 \text{ q.s.}\}, \\ O &:= \{(t, x, v) \in X \mid \|t - t_*\|_\infty < 1, \|x - x_*\|_\infty < 1\}. \end{aligned}$$

A aplicação  $F : O \rightarrow Y$  é definida por

$$F(t, x, v) := \begin{pmatrix} t(\cdot) - \int_0^\cdot v(s) ds \\ x(\cdot) - z_0 - \int_0^\cdot v(s) f(t(s), x(s), w_*(s)) ds \\ \psi(t(1), x(1)) \end{pmatrix}.$$

A hipótese sobre  $\psi$  garante que a aplicação  $(t, x) \mapsto \psi(t(1), x(1))$  é continuamente diferenciável, enquanto que o Lema 4.2.2 garante que  $F$  é continuamente (Fréchet-) diferenciável com derivada<sup>2</sup>

$$DF(t_*, x_*, v_*)(t, x, v) = \begin{pmatrix} t(\cdot) - \int_0^\cdot v(s) ds \\ x(\cdot) - \int_0^\cdot (v_*(s)[D_t f(s)t(s) + D_x f(s)x(s)] + v(s)f(s)) ds \\ D_t \psi(t_*(1), x_*(1))t(1) + D_x \psi(t_*(1), x_*(1))^* x(1) \end{pmatrix}.$$

A aplicação  $J : O \rightarrow \mathbb{R}$  também é continuamente diferenciável (pelo mesmo argumento) e sua derivada é dada por<sup>3</sup>

$$DJ(t_*, x_*, v_*)(t, x, v) = \int_0^1 (v_*(s)[D_t L(s)t(s) + D_x L(s)^* x(s)] + v(s)L(s)) ds + D_t L_1(1)t(1) + D_x L_1(1)^* x(1)$$

<sup>2</sup>Para simplificar a notação escrevemos  $f(s)$  no lugar de  $f(t_*(s), x_*(s), w_*(s))$ .

<sup>3</sup>Novamente adotamos uma notação simplificada:  $L(s) := L(t_*(s), x_*(s), w_*(s))$  e  $L_1(1) := L_1(1, x_*(1))$ .

(1) Provamos inicialmente o lema no caso em que  $DF(t_*, x_*, v_*) : X \rightarrow Y$  não é sobrejetiva. É suficiente provar que

$$\exists (\mu, q, r) \neq 0 \text{ com } r(\tau) + f(\tau)^*q(\tau) = 0 \text{ q.s.} \quad (4.12)$$

e que a equação adjunta (4.9) vale com  $\eta = 0$ .

Considere o problema de valor inicial

$$\frac{dy}{d\tau} = A(\tau)y + b(\tau)v, \quad y(0) = 0, \quad y_1 = Cy(1) \quad (4.13)$$

onde

$$y = \begin{pmatrix} t \\ x \end{pmatrix}, \quad A(\tau) = \begin{pmatrix} 0 & 0 \\ v_* D_t f & v_* D_x f \end{pmatrix}, \quad b(\tau) = \begin{pmatrix} 1 \\ f \end{pmatrix}, \quad C = (D_t \psi, D_x \psi^*).$$

Note que o fato de todo  $y_1 \in \mathbb{R}^p$  poder ser alcançado como valor final em (4.13) por uma escolha adequada de  $v \in L^\infty[0, 1]$  implica na sobrejetividade de  $DF(t_*, x_*, v_*)$ . De fato, seja  $\Phi(\cdot, \cdot)$  a matriz de transição de (4.13), i.e. as colunas de  $\Phi(\cdot, \cdot)$  geram soluções linearmente independentes de  $\frac{dy}{d\tau} = A(\tau)y$ . Seja  $(\bar{y}, \bar{y}_1) \in Y$  qualquer ( $\bar{y} = (\bar{t}, \bar{x})$ ). Logo, existe  $\tilde{y} = (\tilde{t}, \tilde{x})$  tal que

$$\tilde{y}(\tau) - \int_0^\tau A(s)\tilde{y}(s) ds = \bar{y}(\tau), \quad \tau \in [0, 1].$$

(Aplicação do teorema de ponto fixo de Banach para operadores integrais de Volterra; veja [Kre, Parágrafo 5.4].)

Escolha  $v$  de modo que (4.13) possua solução  $y = (t, x)$  com  $y_1 (= Cy(1)) = \bar{y}_1 - C\tilde{y}(1)$ . É fácil ver que  $(\tilde{y} + y, v) = (\tilde{t} + t, \tilde{x} + x, v) \in X$  satisfaz

$$DF(t_*, x_*, v_*)(\tilde{y} + y, v) = (\tilde{t}, \tilde{x}, \bar{y}_1) = (\bar{y}, \bar{y}_1),$$

ficando assim verificada a sobrejetividade de  $DF(t_*, x_*, v_*)$ .

Como supomos que  $DF(t_*, x_*, v_*)$  não é sobrejetiva, necessariamente existe (ao menos um)  $\mu \in \mathbb{R}^p \setminus \{\theta\}$  ortogonal ao subespaço

$$\{y_1 \in \mathbb{R}^p \mid y_1 \text{ é atingível por algum } v \in L^\infty[0, 1]\}.$$

Note que

$$\theta = \mu^* Cy(1) = \int_0^1 \mu^* C \Phi(1, \tau) b(\tau) v(\tau) d\tau, \quad \forall v \in L^\infty[0, 1].$$

Logo,  $-\mu^* C \Phi(1, \tau) b(\tau) = 0$  q.s. em  $[0, 1]$ . Defina  $\bar{q} = (r, q) \in C([0, 1]; \mathbb{R}^{n+1})$  por  $\bar{q}(\tau) := -\Phi(1, \tau)^* C^* \mu$ . Temos então

$$\begin{cases} \frac{d\bar{q}}{d\tau}(\tau) = -A(\tau)^* \bar{q}(\tau), & \tau \in [0, 1] & (\text{provando (4.9) com } \eta = 0) \\ \bar{q}(1) = -C^* \mu & & (\text{provando (4.10)}) \end{cases}$$

e ainda  $0 = b(\tau)^* \bar{q}(\tau) = r(\tau) + f(\tau)^* q(\tau)$ , q.s. em  $[0, 1]$ , provando assim (4.12) e, por conseguinte, (4.11).

(2) Consideramos agora o caso em que  $DF(t_*, x_*, v_*) : X \rightarrow Y$  é sobrejetiva. Neste caso  $(t_*, x_*, v_*)$  é um ponto fracamente regular para  $PA(w_*)$ . De fato, como  $\text{int}(C) \neq \emptyset$  temos que o cone tangencial  $C(t_*, x_*, v_*)$  também possui interior não vazio. O teorema da aplicação aberta (Teorema A.2.6 e Observação A.2.8) garante que  $\text{int}\{DF(t_*, x_*, v_*) C(t_*, x_*, v_*)\} \neq \emptyset$ .

O Teorema 4.1.2 (veja também Observação 4.1.3) garante a existência de  $\eta \geq 0$ ,  $\rho^* \in C(t_*, x_*, v_*)^*$  e  $y^* \in Y^*$  com  $\eta + \|\rho^*\| + \|y^*\| \neq 0$  satisfazendo

$$\rho^*(t, x, v) = \eta DJ(t_*, x_*, v_*)(t, x, v) - y^*(DF(t_*, x_*, v_*)(t, x, v)), \tag{4.14}$$

para todo  $(t, x, v) \in X$ . Note que  $\rho^*(t, x, 0) \geq 0$  para todo par  $t, x$ . De fato, isto segue da identidade  $\rho^*(t, x, 0) = \rho^*((t + t_*, x + x_*, v_*) - (t_*, x_*, v_*))$  e do fato  $(t + t_*, x + x_*, v_*) \in C(t_*, x_*, v_*)^*$ . Logo, como  $\rho^*$  é linear em todos os argumentos, segue que  $\rho^*(t, x, 0) = 0$  para todo par  $t, x$ . Desta forma, podemos concluir que  $\rho^*(t, x, v) = \rho^*(v)$  e da definição do cone tangencial  $C(t_*, x_*, v_*)^*$  temos

$$\rho^*(v - v_*) \geq 0, \forall v \in V_+. \tag{4.15}$$

Escrevendo agora  $y^* = (\zeta^*, \gamma^*, \mu) \in Y^*$ , obtemos para todo par  $t, x$

$$\begin{aligned} \rho^*(v) &= \eta \int_0^1 [v_*(s)(D_t L(s)t(s) + D_x L(s)^* x(s)) + v(s)L(s)] ds \\ &\quad + \eta [D_t L_1(1)t(1) + D_x L_1(1)^* x(1)] - \zeta^* \left( t(\cdot) - \int_0^1 v(s) ds \right) \\ &\quad - \gamma^* \left( x(\cdot) - \int_0^1 v_*(s)[D_t f(s)t(s) + D_x f(s)x(s)] + v(s)f(s) ds \right) \\ &\quad - \mu^*(D_t \psi(1)t(1) + D_x \psi(1)^* x(1)), \end{aligned} \tag{4.16}$$

com  $(\eta, \zeta^*, \gamma^*, \mu) \neq \theta$ .<sup>4</sup>

Defina  $r$  e  $q$  como solução do problema de valor final (retrocedendo no tempo) definido em (4.9), (4.10). Temos que provar que  $(\eta, \mu, r, q)$  satisfazem (4.11) e  $(\eta, \mu, r, q) \neq \theta$ .

Escolha  $\binom{t}{x} = y$  como solução do problema de valor inicial (4.13). Substituindo em (4.16), os termos  $\zeta^*(\dots)$  e  $\gamma^*(\dots)$  desaparecem. Para o primeiro termo em  $\eta$  obtemos

$$\begin{aligned} \eta \int_0^1 v_*(s) D_x L(s)^* x(s) ds &= \\ &= \int_0^1 \left( -\frac{dq}{d\tau}(s) - v_*(s) D_x f(s)^* q(s) \right)^* x(s) ds \\ &= \int_0^1 q(s)^* \frac{dx}{d\tau}(s) ds - [q(s)^* x(s)]_0^1 - \int_0^1 (v_*(s) D_x f(s)^* q(s))^* x(s) ds \\ &= \int_0^1 q(s)^* [v_*(s) D_t f(s)t(s) + v_*(s) D_x f(s)x(s) + v(s)f(s)] ds \\ &\quad - q(1)^* x(1) - \int_0^1 v_*(s) q(s)^* D_x f(s)x(s) ds \\ &= \int_0^1 q(s)^* [v_*(s) D_t f(s)t(s) + v(s)f(s)] ds - q(1)^* x(1) \end{aligned}$$

<sup>4</sup>Notação simplificada:  $\psi(1) = \psi(t_*(1), x_*(1))$ .

e ainda

$$\begin{aligned} \eta \int_0^1 v_*(s) D_t L(s) t(s) ds &= \int_0^1 \left[ -\frac{r}{d\tau}(s) - v_*(s) q(s)^* D_t f(s) \right] t(s) ds \\ &= \int_0^1 r(s) v(s) ds - r(1)t(1) - \int_0^1 v_*(s) q(s)^* D_t f(s) t(s) ds. \end{aligned}$$

Substituindo as duas últimas expressões em (4.16) e observando que

$$\begin{pmatrix} r(1)t(1) \\ q(1)^* x(1) \end{pmatrix} = \begin{pmatrix} [\eta D_t L_1(1) - D_t \psi(1)^* \mu] t(1) \\ [\eta D_x L_1(1) - D_x \psi(1)^* \mu] x(1) \end{pmatrix},$$

obtemos

$$\begin{aligned} \rho^*(v) &= \eta \int_0^1 v(s) L(s) ds + \int_0^1 v(s) q(s)^* f(s) ds + \int_0^1 q(s)^* v_*(s) D_t f(s) t(s) ds \\ &\quad - q(1)^* x(1) + \int_0^1 v(s) r(s) ds - \int_0^1 q(s)^* v_*(s) D_t f(s) t(s) ds - r(1)t(1) \\ &\quad + \eta [D_t L_1(1) t(1) + D_x L_1(1)^* x(1)] - \mu^* (D_t \psi(1) t(1) + D_x \psi(1)^* x(1)) \\ &= \int_0^1 v(s) [r(s) + f(s)^* q(s) + \lambda L(s)] ds. \end{aligned}$$

De (4.15) segue agora

$$\int_0^1 [r(s) + f(s)^* q(s) + \lambda L(s)] (v(s) - v_*(s)) ds = \rho^*(v - v_*) \geq 0, \quad \forall v \in V_+,$$

provando (4.11). Se  $(\lambda, \mu, r, q) = \theta$ , teríamos  $\rho = 0$  e, como consequência de (4.14),  $y^* = 0$  (supomos  $DF(t_*, x_*, v_*)$  sobrejetora). Isto entretanto contradiz a desigualdade  $\eta + \|\rho^*\| + \|y^*\| \neq 0$ , garantida pelo Teorema 4.1.2. Fica assim provado o lema. ■

### 4.3 Condições Necessárias de Otimalidade

A demonstração do princípio do máximo (ver Teorema 3.1.2) é constituída basicamente da transformação das equações e inequações do Lema 4.2.3 ao intervalo  $[0, T^*]$ . A condição de otimalidade (veja ítem *iv*) do Teorema 3.1.2) é, em particular, obtida através da escolha adequada de  $w_*$  e  $v_*$ . A demonstração está dividida em seis passos principais:

(1) Obtenção da equação adjunta.

Usando a notação do Lema 4.2.3 definimos

$$\tau(s) := \inf\{\tau \mid t(\tau) = s\}, \quad \Lambda(s) := r(\tau(s)), \quad \lambda(s) := q(\tau(s)),$$

para  $s \in [0, T^*]$ . Logo  $\Lambda(t(\tau)) = r(\tau)$ ,  $\lambda(t(\tau)) = q(\tau)$ . De (4.9) segue que

$$\frac{d\lambda}{dt} \frac{dt}{d\tau}(\tau) = -v_*(\tau) [D_x f(\tau)^* q(\tau) - \eta D_x L(\tau)].$$



Como  $\frac{d\lambda}{d\tau}(\tau) = v_*(\tau)$ , obtemos a equação adjunta<sup>5</sup>

$$\frac{d\lambda}{dt}(t) = -D_x f[t]^* \lambda(t) - \eta D_x L[t] = -D_x H[t]$$

com a condição de contorno

$$\lambda(T^*) = q(1) = \eta D_x L_1[T^*] - D_x \psi[T^*] \mu.$$

Analogamente, obtemos para  $\Lambda$  o problema de valor inicial

$$\begin{aligned} \frac{d\Lambda}{dt}(t) &= -D_t H[t], \\ \Lambda(T^*) &= \eta D_t L_1[T^*] - D_t \psi[T^*] \mu. \end{aligned}$$

(2) A escolha de  $v_* \in V_+$ .

Definimos a seguir  $v_* : [0, 1] \rightarrow \mathbb{R}$ , de modo que  $\{\tau \in [0, 1] \mid v_*(\tau) = 0\}$  seja uma união enumerável de intervalos disjuntos  $B_k$  e que  $\{t_*(B_k) \mid k \in \mathbb{N}\}$  seja um subconjunto denso de  $[0, T^*]$ .

Tome  $\{t_k\}_{k \in \mathbb{N}}$  um subconjunto denso de  $[0, T^*]$ ,  $\beta_k > 0$  com  $\sum_{k=1}^{\infty} \beta_k = \frac{1}{2}$  e defina

$$\tau_k := \frac{t_k}{2T^*} + \sum_{t_i < t_k} \beta_i, \quad B_k := [\tau_k, \tau_k + \beta_k], \quad B := \bigcup_{k \in \mathbb{N}} B_k.$$

Definimos agora a função  $v_*$  da seguinte forma:

$$v_*(\tau) := \begin{cases} 0, & \tau \in B \\ 2T^*, & \text{senão.} \end{cases}$$

Note que  $t_*(1) = \int_0^1 v_*(\tau) d\tau = T^*$ , pois  $\mu([0, 1] \setminus B) = \frac{1}{2}$ . Note ainda que para todo  $\tau \in B_k$ , temos

$$t_*(\tau) = \int_0^\tau v_*(\tau) d\tau = 2T^* \mu\left([0, \tau_k] \setminus \bigcup_{t_i < t_k} B_i\right) = 2T^* \left(\tau_k - \sum_{t_i < t_k} \beta_i\right) = t_k.$$

Portanto  $\{t_*(B_k) \mid k \in \mathbb{N}\}$  é denso em  $[0, T^*]$ , completando assim a construção.

(3) Provamos que

$$\Lambda(t) + H(t, z^*(t), \lambda(t), u^*(t)) = 0, \quad \text{q.s. em } [0, T^*].$$

Suponha por contradição que  $\Lambda(t) + f(t, z^*(t), u^*(t))^* \lambda(t) + \eta L(t, z^*(t), u^*(t)) > 0$  em algum subconjunto com medida positiva de  $[0, T^*]$ . Como  $t_*$  é absolutamente contínua, a mudança de variáveis  $t = t_*(\tau)$  nos permite concluir que

$$r(\tau) + f(t_*(\tau), x_*(\tau), w_*(\tau))^* q(\tau) + \eta L(t_*(\tau), x_*(\tau), w_*(\tau)) > 0$$

<sup>5</sup>Notação:  $f[t] = f(t, z^*(t), u^*(t))$ ,  $L[t] = L(t, z^*(t), u^*(t))$ ,  $L_1[t] = L_1(t, z^*(t))$ ,  $\psi[t] = \psi(t, z^*(t))$ . Analogamente, escrevemos para a função hamiltoniana (veja Definição 3.1.1)  $H[t] = H(t, z^*(t), \lambda(t), u^*(t)) = (\lambda(t), f(t, z^*(t), u^*(t))) + \eta L(t, z^*(t), u^*(t))$ .

em um subconjunto  $A$  de  $\{\tau \in [0, 1] \mid v_*(\tau) > 0\}$  com  $\mu(A) > 0$ . Definindo agora

$$v(\tau) := \begin{cases} v_*(\tau), & \tau \notin A \\ 0, & \tau \in A \end{cases}$$

obtemos

$$\int_0^1 [r(\tau) + f(\tau)^*q(\tau) + \eta L(\tau)] (v(\tau) - v_*(\tau)) d\tau < 0,$$

contradizendo (4.11).

(4) Obtenção da equação de evolução da função de Hamilton.

Note que de (1) e (3) segue

$$H(t, z^*(t), \lambda(t), u^*(t)) = -\Lambda(t) = -\int_t^{T^*} D_t H[s] ds - \eta D_t L_1[T^*] + D_t \psi[T^*]^* \mu.$$

(5) Escolha de  $w_*$  no subconjunto  $\{\tau \in [0, 1] \mid v_*(\tau) = 0\}$ .

Sejam  $c_j \in \mathbb{R}$  para  $j \in \mathbb{N}$ , tais que para todo  $|t| \leq T^* + 1$ ,  $|x| \leq \|x_*\|_\infty + 1$ ,  $|u| < j$  tenhamos

$$\begin{aligned} |L(t, x, u)|, |f(t, x, u)|, |D_x L(t, x, u)|, |D_x f(t, x, u)|, |D_t L(t, x, u)|, |D_t f(t, x, u)| &\leq c_j, \\ |D_{xx} L(t, x, u)|, |D_{xx} f(t, x, u)|, |D_{xt} L(t, x, u)|, |D_{xt} f(t, x, u)| &\leq c_j, \\ |D_{tt} L(t, x, u)|, |D_{tt} f(t, x, u)|, |D_{tx} L(t, x, u)|, |D_{tx} f(t, x, u)| &\leq c_j. \end{aligned}$$

Escreva agora  $B_k = \bigcup_{j \in \mathbb{N}} B_{k,j}$ ,  $k \in \mathbb{N}$ , onde  $\mu(B_{k,j}) \leq c_0/(2^k c_j)$  para alguma constante  $c_0$ . (Este passo é desnecessário quando  $\Omega$  é limitado.)

Para cada  $j \in \mathbb{N}$  seja  $\{u_i^j\}_{i \in \mathbb{N}}$  um subconjunto denso de  $\Omega \cap \{u \in \mathbb{R}^m \mid j-1 \leq |u| < j\}$ . Escreva  $B_{k,j} = \bigcup_{i \in \mathbb{N}} B_{k,j,i}$ , onde  $B_{k,j,i}$  são intervalos disjuntos. Defina agora

$$w_*(\tau) := u_i^j, \text{ se } \tau \in B_{k,j,i}.$$

Note que a condição (4.8) do Lema 4.2.3 é satisfeita para essa escolha de  $w_*$ . De fato, para todo  $|t| \leq T^* + 1$ ,  $|x| \leq \|x_*\|_\infty + 1$  temos

$$\begin{aligned} \int_0^1 |L(t, x, w_*(\tau))| d\tau &= \int_{[0,1] \setminus B} |L(t, x, u^*(t_*(\tau)))| d\tau + \int_B |L(t, x, u^*(t_*(\tau)))| d\tau \\ &\leq M + \int_{\bigcup B_{k,j,i}} |L(t, x, u_i^j)| d\tau \\ &\leq M + \sum_{k,j} \mu(B_{k,j}) c_j \\ &\leq M + \sum_{k,j} \frac{c_0}{2^k c_j} c_j \leq \infty. \end{aligned}$$

Provando assim que  $L$ ,  $f$  e suas derivadas parciais de ordem  $\leq 2$  são majoradas por uma função  $L^1$ .

(6) Obtenção da condição de otimalidade.

Tome  $v \in V_+$  com  $v = v_*$  q.s. fora de  $B_{k,j,i}$ . De (4.11) segue que

$$\int_{B_{k,j,i}} [r(\tau) + f(t_k, z^*(t_k), u_i^j)^* q(\tau) + \eta L(t_k, z^*(t_k), u_i^j)] v(\tau) d\tau \geq 0.$$

Como  $t_*(\tau) = t_k$  para  $\tau \in B_{k,j,i}$ , temos  $r(\tau) = \Lambda(t_k)$ ,  $q(\tau) = \lambda(t_k)$  em  $B_{k,j,i}$ . Logo,

$$\Lambda(t_k) + f(t_k, z^*(t_k), u_i^j)^* \lambda(t_k) + \eta L(t_k, z^*(t_k), u_i^j) \geq 0,$$

isto é

$$\Lambda(t_k) + H(t_k, z^*(t_k), u_i^j, \lambda(t_k)) \geq 0.$$

Como  $\{t_k\}_{k \in \mathbb{N}}$  é denso em  $[0, T^*]$ , obtemos da continuidade de  $\Lambda$  e  $H$

$$\Lambda(t) + H(t, z^*(t), u_i^j, \lambda(t)) \geq 0, \quad \forall t \in [0, T^*].$$

A densidade de  $\{u_i^j\}_{i,j}$  em  $\Omega$  nos permite concluir, por argumento semelhante, que

$$\Lambda(t) + H(t, z^*(t), u, \lambda(t)) \geq 0, \quad \forall u \in \Omega, t \in [0, T^*].$$

Esta desigualdade, juntamente com (3), implica na condição de otimalidade

$$H(t, z^*(t), u, \lambda(t)) \geq H(t, z^*(t), u^*(t), \lambda(t)), \quad \text{q.s. em } [0, T^*],$$

para todo  $u \in \Omega$ . ■

## 4.4 Exercícios

4.1 Considere o problema de otimização (P) do Parágrafo 4.1. Mostre que  $x \in O$  é ponto regular para (P) se e somente se

$$Y = dF(x)C(x) - K(F(x)).$$

4.2 De exemplo de um problema de otimização (P) que possua um ponto fracamente regular que não seja regular.

4.3 (Teorema da aplicação aberta) Sejam  $X, Y$  espaços de Banach. Mostre que se  $T \in B(X, Y)$  é sobrejetora, então existe  $r > 0$  tal que  $B_r \subset T(B_1)$ .



## Apêndice A

# Otimização Infinita

Neste apêndice discutimos os detalhes do desenvolvimento apresentado no Parágrafo 4.1. As definições pertinentes são fornecidas assim como as demonstrações completas dos resultados daquele parágrafo.

### A.1 Um Problema Abstrato de Otimização

A título de motivação, suponha que  $J : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma aplicação diferenciável. Se  $\bar{x} \in \mathbb{R}^n$  é tal que

$$J(\bar{x}) = \min_{x \in \mathbb{R}^n} J(x),$$

então o gradiente de  $J$  se anula em  $\bar{x}$ , i.e.  $\nabla J(\bar{x}) = \theta$ . A recíproca é verdadeira se  $J$  for convexo. Considere agora o problema de otimização com restrições diferenciáveis:

$$\left\{ \begin{array}{l} \text{Minimizar } J(x) \\ \text{sujeto a} \\ g_i(x) \leq 0, \quad 1 \leq i \leq m, \\ h_i(x) = 0, \quad 1 \leq i \leq l \end{array} \right.$$

onde  $J : U \rightarrow \mathbb{R}$ ,  $G = (g_1, \dots, g_m) : U \rightarrow \mathbb{R}^m$ ,  $H = (h_1, \dots, h_l) : U \rightarrow \mathbb{R}^l$  são aplicações diferenciáveis no aberto  $U \subset \mathbb{R}^n$ . Representamos as restrições de forma simplificada por  $G(x) \leq \theta$ ,  $H(x) = \theta$ .

As condições necessárias para uma solução  $\bar{x}$  deste problema são fornecidas pelo teorema dos multiplicadores de Lagrange.

**Teorema A.1.1** *Sejam  $J, G, H$  funções satisfazendo as condições acima. Se a função  $J$  sujeita às restrições  $G(x) \leq \theta$ ,  $H(x) = \theta$  possui um extremo local (max/min) em  $\bar{x}$ , então existem  $\eta \in \mathbb{R}$ ,  $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ ,  $\mu = (\mu_1, \dots, \mu_l) \in \mathbb{R}^l$  tais que:*

$$i) \quad \eta \nabla J(x_*) + \sum_{i=1}^m \lambda_i \nabla g_i(x_*) + \sum_{j=1}^l \mu_j \nabla h_j(x_*) = \theta;$$

$$ii) \quad \lambda_i \geq 0, \quad \lambda_i g_i(x_*) = 0, \quad 1 \leq i \leq m;$$

iii)  $\eta \geq 0$ ,  $|\eta| + |\lambda| + |\mu| \neq 0$ .

Neste parágrafo desenvolvemos um resultado análogo em espaços de Banach. Tal resultado é usado na demonstração do princípio do máximo, apresentada no Parágrafo 4.2.

Antes de formular o problema abstrato de otimização em espaços de Banach necessitamos de alguns conceitos. Sejam  $X, Y$  espaços normados,  $M \subset X$ ,  $z \in X$ , usamos a notação:

- $B_r(z) := \{x \mid \|x - z\| \leq r\}$ ;  $B_r := B_r(\theta)$ ,  $r > 0$ ;
- $\text{int}(M) := \{x \in X \mid \exists r > 0 \text{ com } B_r(x) \subset M\}$ ;  $\text{cl}(M) := X \setminus \text{int}(X \setminus M)$ ;
- $B(X, Y) := \{T \mid T : X \rightarrow Y \text{ linear, contínua}\}$ ;  $X^* := B(X, \mathbb{R})$ .

Note que  $B(X, Y)$  é também um espaço normado que é completo quando  $Y$  for completo. O espaço  $X^*$  é o dual topológico de  $X$ . O último conceito que necessitamos diz respeito à regularidade de aplicações entre espaços de Banach.

**Definição A.1.2** Sejam  $X, Y$  Espaços de Banach e  $O \subset X$  aberto. Uma aplicação  $F : O \rightarrow Y$  é dita (**Frechet-**) diferenciável em  $x \in O$ , quando existe  $T \in B(X, Y)$  satisfazendo

$$\lim_{h \rightarrow \theta} \frac{\|F(x+h) - F(x) - Th\|}{\|h\|} = 0.$$

$dF(x) := T$  é denominada **derivada de F em x** (unicamente determinada).  $F$  é dita **continuamente (Frechet-) diferenciável** em  $O$ , quando for diferenciável em todo  $x \in O$  e a aplicação  $dF : O \rightarrow B(X, Y)$  é contínua. □ □ □

Formulamos novamente o problema abstrato de otimização introduzido no Parágrafo 4.1:

$$(P) \quad \begin{cases} \text{Minimizar } J(x) \\ \text{sujeito a } x \in C, F(x) \in K \end{cases}$$

onde  $X, Y$  são espaços de Banach;  $O \subset X$  aberto;  $C \subset X$ ,  $K \subset Y$  são fechados e convexos;  $J : O \rightarrow \mathbb{R}$ ,  $F : O \rightarrow Y$  são continuamente diferenciáveis.

A seguir obtemos o teorema de multiplicadores desejado (veja conjectura na página 82) aproximando localmente o conjunto dos pontos admissíveis  $X_{ad} := C \cap F^{-1}(K) \cap O$  para  $(P)$  por conjuntos convexos apropriados e utilizando um teorema de separação para conjuntos convexos em espaços de Banach.

## A.2 Linearização do Problema de Otimização

**Definição A.2.1** Seja  $X$  um espaço de Banach,  $C \subset X$  convexo,  $x \in C$ . O conjunto

$$C(x) := \{a(c-x) \mid a \geq 0, c \in C\}$$

é denominado **cone tangencial em x** por  $C$  (veja Figura A.1). □ □ □

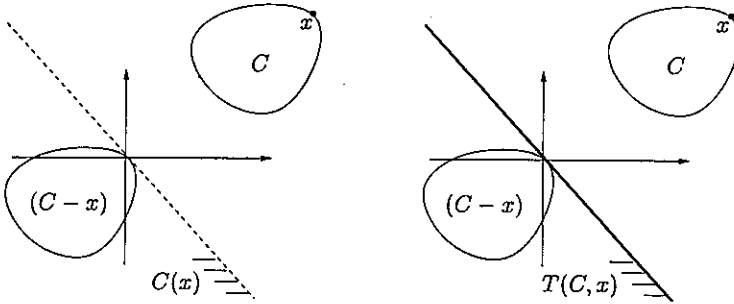


Figura A.1: Cones tangenciais  $C(x)$  e  $T(C, x)$ .

**Definição A.2.2** Seja  $X$  um espaço de Banach,  $M \subset X$ ,  $x \in M$ . O conjunto

$$T(M, x) := \{h \in X \mid \exists t_0 > 0 \text{ e } r : [0, t_0] \rightarrow X \text{ tal que} \\ x + th + r(t) \in M \text{ para } t \in [0, t_0], \lim_{t \downarrow 0} r(t)/t = 0\}$$

é denominado **cone tangencial a M por x** (veja Figura A.1). □ □ □

**Lema A.2.3** *Seja  $X$  um espaço de Banach,  $C \subset X$  fechado e convexo,  $x \in C$ . São verdadeiras as afirmações:*

- i)  $C - x$ ,  $(C - x)_1$ ,  $C(x)$  são convexos;
- ii)  $C - x$ ,  $(C - x)_1$  são fechados;
- iii)  $C(x) \subset T(C, x)$ ;

onde  $(M)_1 := M \cap B_1$ , para todo subconjunto  $M$  de um espaço normado  $Z$ .

**Demonstração:** Segue imediatamente das Definições A.2.1 e A.2.2. ■

**Teorema A.2.4** *Se  $\bar{x} \in X_{ad} = C \cap F^{-1}(K) \cap O$  é um mínimo local de  $(P)$ , então para todo  $h \in T(X_{ad}, \bar{x})$  temos*

$$dJ(\bar{x})h \geq 0.$$

**Demonstração:** Sejam  $t_0 \in \mathbb{R}$  e  $r : [0, t_0] \rightarrow X$  escolhidos como na Definição A.2.2 para o cone tangencial  $T(X_{ad}, \bar{x})$ . Logo, para todo  $t \in (0, t_0]$  temos

$$\begin{aligned} 0 &\leq t^{-1}(J(\bar{x} + th + r(t)) - J(\bar{x})) \\ &\leq t^{-1}(J(\bar{x} + th) - J(\bar{x})) + t^{-1}(J(\bar{x} + th + r(t)) - J(\bar{x} + th)) \\ &\leq t^{-1}(J(\bar{x} + th) - J(\bar{x})) + t^{-1}\|dJ(\xi_t)\| \|r(t)\|, \end{aligned}$$

onde a existência de  $\xi_t \in X$  é garantida pelo teorema do valor médio. O teorema segue agora tomando o limite  $t \downarrow 0$ , uma vez que  $\lim_{t \downarrow 0} r(t)t^{-1} = 0$  e  $\lim_{t \downarrow 0} \xi_t = \bar{x}$ . ■

**Definição A.2.5** Seja  $x \in O$ . Dizemos que  $x$  é um **ponto regular** para  $(P)$  quando  $\theta \in \text{int}[dF(x)(C-x) - (K-F(x))]$ . Caso  $\text{int}[dF(x)(C-x) - (K-F(x))] \neq \emptyset$ ,  $x$  é dito **ponto fracamente regular** para  $(P)$ .  $\square \square \square$

O resultado a seguir representa um passo fundamental para a demonstração do desejado teorema de condições necessárias. O leitor atento observa que este resultado pode ser interpretado como uma generalização do teorema da aplicação aberta. Para detalhes veja [Zoku].

**Teorema A.2.6** *Sejam  $X, Y$  espaços de Banach,  $T \in B(X, Y)$  e  $C \subset X, K \subset Y$  fechados e convexos. Para  $x \in C, y \in K$  são equivalentes as afirmações:*

- a)  $Y = T(C(x)) - K(y)$ ;  
 b) Existe  $r > 0$ , tal que  $B_r \subset T((C-x)_1) - (K-y)_1$ .

**Demonstração:** (a)  $\Rightarrow$  (b)

(1) Provamos que  $C(x) = \bigcup_{n \in \mathbb{N}} n(C-x)_1$ . A inclusão  $C(x) \supset \bigcup_{n \in \mathbb{N}} n(C-x)_1$  segue diretamente da definição de  $C(x)$ . Seja agora  $z = a(c-x) \in C(x)$ . Escolha  $b \in (0, 1)$  tal que  $b(c-x) \in B_1$ . Como  $B_1$  e  $C-x$  são conjuntos convexos, temos que

$$r(c-x) \in (C-x)_1 \text{ para todo } r \in [0, b].$$

Logo  $z = a(c-x) = \frac{na}{n}(c-x) \in n(C-x)_1$  para  $n$  suficientemente grande. Como  $z \in C(x)$  é arbitrário, temos  $C(x) \subset \bigcup_{n \in \mathbb{N}} n(C-x)_1$ .

(2) Analogamente prova-se que  $K(y) = \bigcup_{n \in \mathbb{N}} n(K-y)_1$ .

(3) Defina  $A_s := s[T((C-x)_1) - (K-y)_1]$ ,  $s \in \mathbb{R}$ . Afiramos que  $Y = \bigcup_{n \in \mathbb{N}} A_n$ . De fato, basta observar que

$$\begin{aligned} Y &= T(C(x)) - K(y) = T\left(\bigcup_{n \in \mathbb{N}} n(C-x)_1\right) - \bigcup_{m \in \mathbb{N}} m(K-y)_1 \\ &= \bigcup_{m, n \in \mathbb{N}} [nT((C-x)_1) - m(K-y)_1] = \bigcup_{n \in \mathbb{N}} n[T((C-x)_1) - (K-y)_1]. \end{aligned}$$

(Na última igualdade usamos a inclusão  $m(K-y)_1 \subset n(K-y)_1$ , para  $m \leq n$ , a qual se deve a  $\theta \in (K-y)_1$  e a convexidade de  $(K-y)_1$ .)

(4) Do teorema de Baire segue que existe pelo menos um  $m \in \mathbb{N}$  com  $\text{int}(cl(A_m)) \neq \emptyset$ .

(Teorema de Baire: Se um espaço métrico completo é escrito como união enumerável de subconjuntos, então pelo menos um destes contém um aberto; veja [Kre] ou [Ru1].)

(5) Provamos que  $\theta \in \text{int}(cl(A_1))$ . De fato, (4) implica que  $\exists a \in \text{int}(cl(A_m))$  e de (3) temos que  $\exists k \in \mathbb{N}$  com  $-a \in cl(A_k)$ . Logo, pela definição de  $A_s$ , temos que  $-mk^{-1}a \in cl(A_m)$ . Como  $a \in \text{int}(cl(A_m))$ , então  $\exists \delta > 0$  com  $B_\delta(a) \subset \text{int}(cl(A_m))$ . Mas  $-mk^{-1}a \in cl(A_m)$  e  $cl(A_m)$  é convexo, então  $\exists \varepsilon > 0$  tal que  $\theta \in B_\varepsilon \subset cl(A_m)$ . Portanto,  $B_{\varepsilon m^{-1}} \subset cl(A_1)$ .

(6) Seja  $r > 0$  com  $B_{2r} \subset cl(A_1)$ . Provamos que  $B_r \subset A_1$ , obtendo assim (b). Note que

$$B_r \subset cl(A_{1/2}) \subset \{\bar{y} \in Y \mid \text{dist}(\bar{y}, A_{1/2}) \leq \frac{r}{2}\} = A_{2^{-1}} + B_{2^{-1}r}.$$



Reduzindo o diâmetro dos conjuntos pelo fator  $2^i$  temos

$$B_{2^{-i}r} \subset A_{2^{-(i+1)}} + B_{2^{-(i+1)}r}, \forall i \in \mathbb{N}.$$

Dado  $\bar{y} \in B_r$ , definimos as seqüências  $\{x_i\}_{i \in \mathbb{N}} \subset X$  e  $\{y_i\}_{i \in \mathbb{N}}, \{r_i\}_{i \in \mathbb{N}} \subset Y$  indutivamente

$$i = 1: \quad \bar{y} = T(2^{-1}x_1) - 2^{-1}y_1 + r_1 \\ \text{com } x_1 \in (C-x)_1, y_1 \in (K-y)_1, \|r_1\| \leq 2^{-1}r;$$

$$i > 1: \quad r_i = T(2^{-(i+1)}x_{i+1}) - 2^{-(i+1)}y_{i+1} + r_{i+1} \\ \text{com } x_{i+1} \in (C-x)_1, y_{i+1} \in (K-y)_1, \|r_{i+1}\| \leq 2^{-(i+1)}r.$$

Definindo agora  $u_n := \sum_{i=1}^n 2^{-i}x_i$  e  $v_n := \sum_{i=1}^n 2^{-i}y_i$ , temos

$$\bar{y} = Tu_n - v_n + r_n, \quad n \in \mathbb{N}. \tag{A.1}$$

Note que  $\lim r_n = \theta$  e ainda que  $u_n$  é de Cauchy em  $X$  pois  $\|x_i\| \leq 1, i \in \mathbb{N}$  e

$$\|u_{n+m} - u_n\| \leq \left\| \sum_{i=n+1}^{n+m} 2^{-i}x_i \right\| \leq 2^{-n}.$$

Como  $X$  é Banach, existe  $u \in X$  com  $\lim u_n = u$  e como  $T$  é contínua então  $Tu = \lim Tu_n$ . Logo (A.1) garante que  $v_n$  converge para algum  $v \in Y$  e temos  $\bar{y} = Tu - v$ . Verificamos por fim que  $u \in (C-x)_1$  e  $v \in (k-y)_1$ .

Note que  $u_n = \sum_{i=1}^n 2^{-i}x_i + 2^{-n}\theta, n \in \mathbb{N}$ . Logo  $u_n$  é combinação linear convexa dos elementos  $x_1, x_2, \dots, x_n, \theta$  de  $(C-x)_1$ . Como  $(C-x)_1$  é convexo, temos  $u \in (C-x)_1$ . A inclusão  $v \in (k-y)_1$  é demonstrada de forma análoga.

(b)  $\Rightarrow$  (a) Obviamente  $\theta \in T(C(x)) - K(y)$ . Se  $\bar{y} \in Y, \bar{y} \neq \theta$  então  $r\|\bar{y}\|^{-1}\bar{y} \in B_r$ . Logo

$$r\|\bar{y}\|^{-1}\bar{y} = T(a(c-x)) - b(k-y),$$

para  $a, b \geq 0, c \in C$  e  $k \in K$  apropriados. Isto prova que

$$\bar{y} = T(r^{-1}\|\bar{y}\|a(c-x)) - (r^{-1}\|\bar{y}\|b(k-y)) \in T(C(x)) - K(y).$$

■

**Corolário A.2.7** *Seja  $x$  admissível para (P), i.e.  $x \in X_{ad} = C \cap F^{-1}(K) \cap O$ . São equivalentes as afirmativas:*

a)  $x$  é um ponto regular para (P);

b)  $Y = dF(x)C(x) - K(F(x))$ .

**Demonstração:** (b)  $\Rightarrow$  (a): Segue da aplicação do Teorema A.2.6 a  $T = dF(x)$ ,  $y = F(x)$ .  
 (a)  $\Rightarrow$  (b): Seja  $\bar{y} \in Y$ . Como  $x$  é regular, existem  $\varepsilon > 0$  e  $\delta > 0$  tais que

$$\varepsilon \bar{y} \in B_\delta \subset dF(x)(C - x) - (K - F(x)).$$

Portanto,

$$\bar{y} \in dF(x)\varepsilon^{-1}(C - x) - \varepsilon^{-1}(K - F(x)) \subset dF(x)C(x) - (K(F(x)))$$

e o teorema fica provado. ■

**Observação A.2.8** Para  $C = X$ ,  $K = \{\theta\}$  obtemos como corolário do Teorema A.2.6 o conhecido teorema da aplicação aberta da análise funcional:

Se  $T \in B(X, Y)$  é sobrejetiva, então existe  $r > 0$  com  $B_r \subset T(B_1)$ .

(Veja [Kre, Parágrafo 4.12].) □ □ □

**Definição A.2.9** Seja  $x \in X_{ad} = C \cap F^{-1}(K) \cap O$ . O conjunto

$$L(X_{ad}, x) := \{h \in X \mid h \in C(x), dF(x)h \in K(F(x))\}.$$

é denominado **cone linearizado em x para (P)**. □ □ □

O teorema a seguir nos permite acoplar a desigualdade  $dJ(x_0)h \geq 0$ ,  $\forall h \in T(X_{ad}, x_*)$ , obtida no Teorema A.2.4 com a derivada de  $F$ . Este é exatamente o fato que fornece as condições necessárias para o problema abstrato de otimização (veja Teoremas A.3.3 e A.3.4).

**Teorema A.2.10** Se  $x$  é um ponto regular para (P) então

$$L(X_{ad}, x) \subset T(X_{ad}, x).$$

**Demonstração:**<sup>1</sup> Seja  $h \in C(x)$  com  $h = a(c - x)$ ,  $dF(x)h = b(k - F(x))$ , onde  $a, b \geq 0$ ,  $c \in C$  e  $k \in K$ . Temos que provar que  $\exists t_0 > 0$  e  $r : [0, t_0] \rightarrow X$  satisfazendo para todo  $t \in [0, t_0]$ :

$$x + th + r(t) \in C, F(x + th + r(t)) \in K$$

e ainda que  $\lim_{t \downarrow 0} r(t)/t = 0$ .

Como  $x$  é ponto regular, então  $x + 2th \in C$  e  $F(x) + 2t dF(x)h \in K$  para  $t$  suficientemente pequeno. Como  $C$  e  $K$  são ambos convexos, basta provar que  $\exists t_0 > 0$  e  $r : [0, t_0] \rightarrow X$  satisfazendo

$$x + 2r(t) \in C, F(x) + 2z(t) \in K, \forall t \in [0, t_0] \text{ e } \lim_{t \downarrow 0} r(t)/t = 0, \quad (\text{A.2})$$

com  $z : [0, t_0] \rightarrow Y$  definida por  $z(t) := F(x + th + r(t)) - F(x) - t dF(x)h$ .

Para cada  $t \in [0, t_0]$ , construímos os vetores  $r(t) \in X$  e  $z(t) \in Y$  tomando limites de

<sup>1</sup>Conforme [A].

sequências que são construídas por uma variante do método de Newton. Vamos à construção:

Se  $h = 0$  tome simplesmente  $r(t) \equiv \theta$ . Seja então  $h \in X \setminus \{\theta\}$  com  $h = a(c - x)$ ,  $dF(x)h = b(k - F(x))$ , onde  $a, b \geq 0$ ,  $c \in C$  e  $k \in K$ . O Teorema A.2.6 e o Corolário A.2.7 garantem a existência de  $s > 0$  tal que

$$B_s \subset dF(x)(C - x)_1 - (K - F(x))_1. \quad (\text{A.3})$$

Escolha  $\delta \in [0, \frac{1}{4}]$  com  $B_{4\delta} \subset O$  e  $\|dF(\xi) - dF(x)\| \leq s/2$ , para todo  $\xi \in B_{4\delta}$ . Defina agora  $t_0 := 2\delta\|h\|^{-1}$  e tome  $t \in [0, t_0]$  qualquer. Nos passos a seguir definimos os valores de  $r(t)$  e  $z(t)$  e provamos que satisfazem (A.2).

(1) Provamos inicialmente que  $\|F(\bar{x}) - F(\hat{x}) - dF(x)(\bar{x} - \hat{x})\| \leq \frac{s}{2}\|\bar{x} - \hat{x}\|$  para todo  $\bar{x}, \hat{x} \in B_{4\delta}$ .

De fato, definindo  $g(\nu) := F(\nu\bar{x} + (1 - \nu)\hat{x}) - \nu dF(x)(\bar{x} - \hat{x})$  para  $\nu \in [0, 1]$  temos

$$\begin{aligned} \|g(1) - g(0)\| &\leq \sup_{\nu \in [0, 1]} \|g'(\nu)\| \\ &= \sup\{\|(dF(\xi) - dF(x))(\bar{x} - \hat{x})\| \mid \xi \text{ entre } \bar{x} \text{ e } \hat{x}\} \\ &\leq \frac{s}{2}\|\bar{x} - \hat{x}\|. \end{aligned}$$

(2) Definimos indutivamente sete sequências que nos permitirão construir  $r(t)$  e  $z(t)$ . Para que a construção faça sentido definimos  $r_{-1} = z_{-1} = x_{-1} = y_{-1} = \theta$ . Para  $k \in \mathbb{N} \cup \{0\}$  tome

$$\begin{aligned} r_k &:= r_{k-1} + x_{k-1}, \\ z_k &:= z_{k-1} + y_{k-1}, \\ d_k &:= z_k - F(x + th + r_k) + F(x) + t dF(x)h, \\ u_k, v_k &: \begin{cases} \text{se } d_k = \theta, & u_k = v_k = \theta, \\ \text{se } d_k \neq \theta, & u_k \in (C - x)_1, v_k \in (K - F(x))_1 \text{ com} \\ & dF(x)u_k - v_k = s\|d_k\|^{-1}d_k, \end{cases} \\ x_k &:= s^{-1}\|d_k\|u_k, \\ y_k &:= s^{-1}\|d_k\|v_k. \end{aligned}$$

(3) Provamos usando argumento indutivo que as sequências em (2) estão bem definidas e que para todo  $k \in \mathbb{N}$  vale

$$r_k \in 2s^{-1}\|d_0\|(C - x)_1, \quad z_k \in 2s^{-1}\|d_0\|(K - F(x))_1, \quad \|d_k\| \leq 2^{-k}\|d_0\| \leq 2^{-k}s\delta.$$

• Para  $k = 0$  temos  $r_0 = z_0 = 0$  e  $\|x + th + r_0 - x\| \leq t\|h\| \leq 2\delta$ . Logo, de (1) segue

$$\|d_0\| = \|F(x + th) - F(x) - dF(x)th\| \leq \frac{s}{2}t\|h\| \leq s\delta.$$

Note ainda que  $u_0$  e  $v_0$  estão bem definidos por (A.3).

• Para  $k \geq 1$  temos

$$r_k = \sum_{i=0}^{k-1} x_i = \sum_{i=0}^{k-1} s^{-1}\|d_i\|u_i = 2s^{-1}\|d_0\| \sum_{i=0}^{k-1} \frac{\|d_i\|}{2\|d_0\|} u_i,$$

com  $u_i \in (C-x)_1$ . Note que  $0 \leq \|d_i\|(2\|d_0\|)^{-1} \leq 1$ ,  $0 \leq i \leq k-1$  e  $\sum_{i=0}^{k-1} \|d_i\|(2\|d_0\|)^{-1} \leq 1$ . Logo  $s(2\|d_0\|)^{-1}r_k$  é combinação linear convexa de  $u_0, \dots, u_{k-1}, \theta \in (C-x)_1$ , i.e.  $r_k \in 2s^{-1}\|d_0\|(C-x)_1$ . A estimativa para  $\|z_k\|$  é obtida de forma análoga. Como  $\|d_0\| \leq s\delta$ , temos

$$\|x + th + r_k - x\| \leq t\|h\| + \|r_k\| \leq 4\delta, \quad \|x + th + r_{k-1} - x\| \leq 4\delta$$

e

$$\begin{aligned} -d_k &= F(x + th + r_k) - F(x) - t dF(x)h - z_k \\ &= F(x + th + r_k) - F(x) - t dF(x)h - z_{k-1} - y_{k-1} \\ &= F(x + th + r_k) - F(x) - t dF(x)h - z_{k-1} + d_{k-1} - dF(x)_{x_{k-1}} \\ &= F(x + th + r_k) - F(x + th + r_{k-1}) - dF(x)_{x_{k-1}}. \end{aligned}$$

De (1) obtemos agora

$$\|d_k\| \leq \frac{s}{2}\|x_{k-1}\| \leq \frac{1}{2}\|d_{k-1}\|\|u_{k-1}\| \leq \frac{1}{2}\|d_{k-1}\| \leq 2^{-k}\|d_0\|,$$

completando assim a prova por indução.

(4) Sobre a convergência das sequências  $r_k$  e  $z_k$ :

•  $\{r_k\}_{k \in \mathbb{N}}$  é de Cauchy pois  $\sum_{k=0}^{\infty} \|x_k\| \leq 2\delta < \infty$ . Definindo  $r(t) := \lim_{k \rightarrow \infty} r_k$ , temos que  $r(t) \in 2s^{-1}\|d_0\|(C-x)_1$ , pois  $(C-x)_1$  é fechado.

•  $\{z_k\}_{k \in \mathbb{N}}$  é de Cauchy pois  $\sum_{k=0}^{\infty} \|y_k\| \leq 2\delta < \infty$ . Definindo  $z(t) := \lim_{k \rightarrow \infty} z_k$ , temos que  $z(t) \in 2s^{-1}\|d_0\|(K-F(x))_1$ , pois  $(K-F(x))_1$  é fechado.

Como  $t \in [0, t_0]$  é arbitrário,  $r$  e  $z$  estão bem definidas no intervalo  $[0, t_0]$ .

(5) Verificação de (A.2).

Em (3) vimos que  $s^{-1}\|d_0\| \leq \delta \leq \frac{1}{4}$ . Logo  $r(t) \in \frac{1}{2}(C-x)_1$  e  $z(t) \in \frac{1}{2}(K-y)_1$ , i.e.

$$x + 2r(t) \in C, \quad F(x) + 2z(t) \in K, \quad \forall t \in [0, t_0].$$

Para completar note que

$$0 \leq t^{-1}\|r(t)\| \leq t^{-1}2s^{-1}\|d_0\| \leq 2s^{-1}t^{-1}\|F(x+th) - F(x) - t dF(x)(h)\|,$$

provando assim que  $\lim_{t \downarrow 0} r(t)t^{-1} = \theta$ . ■

### A.3 Condições Necessárias para o Problema Abstrato

Neste parágrafo provamos uma versão do teorema de multiplicadores de lagrange, que nos permite obter condições necessárias para a otimalidade do problema (P).

**Definição A.3.1** Seja  $X$  um espaço de Banach e  $M \subset X$ . O conjunto

$$M^* := \{\lambda \in X^* \mid \langle \lambda, m \rangle \geq 0 \forall m \in M\}$$

é denominado **cone dual a M em  $X^*$** .<sup>2</sup>

□ □ □

<sup>2</sup>A aplicação de um funcional  $\lambda \in X^*$  ao elemento  $x \in X$  é aqui denotada por  $\langle \lambda, x \rangle$ .

**Observação A.3.2** Um subconjunto  $M$  de um espaço vetorial é denominado *cone*, quando  $x \in M$  implica em  $ax \in X$  para todo  $a \in (0, \infty)$ . É fácil ver que o conjunto  $M^*$  na Definição A.3.1 satisfaz essa propriedade. Outra propriedade de fácil verificação do cone dual é que

$$(M - x)^* = M(x)^*$$

para  $x \in M \subset X$ .

□ □ □

**Teorema A.3.3** *Seja  $x_*$  um mínimo local de  $(P)$ . Se  $x_*$  é um ponto regular para  $(P)$ , então existem  $\psi \in C(x_*)^*$ ,  $\phi \in K(F(x_*))^*$  satisfazendo a condição de Kuhn-Tucker:*

$$dJ(x_*) - \phi \circ dF(x_*) = \psi. \tag{A.4}$$

**Demonstração:** Dos Teoremas A.2.4 e A.2.10 temos  $dJ(x_*)h \geq 0$ , para todo  $h \in L(X_{ud}, x_*)$ , isto é

$$dJ(x_*)h \geq 0, \forall h \in C(x_*) \text{ com } dF(x_*)h \in K(F(x_*)).$$

Fazemos agora uma construção que nos permite utilizar um teorema de separação para conjuntos convexos para provar o teorema.

(1) Definimos o conjunto  $A \subset Y \times \mathbb{R}$  por

$$A := \{(y - dF(x_*)x, dJ(x_*)x + a) \mid x \in C(x_*), y \in K(F(x_*)), a \geq 0\}.$$

$A$  é obviamente convexo e ainda  $(\theta, 0) \in A$  (tome  $x = \theta, y = \theta, a = 0$ ). Note ainda que  $(\theta, 0) \notin \text{int}(A)$  pois  $(\theta, b) \notin A$  para  $b < 0$ .

(2) Provamos que  $\text{int}(A) \neq \emptyset$ .

Como  $x_*$  é um ponto regular para  $(P)$ , o Corolário A.2.7 e o Teorema A.2.6 garantem que  $\exists r > 0$  tal que

$$B_r \subset -dF(x_*)(C - x_*)_1 + (K - F(x_*))_1.$$

Defina  $\gamma := \sup\{dJ(x_*)x \mid x \in (C - x_*)_1\}$ . Basta provarmos que  $B_r \times \{b \in \mathbb{R} \mid b \geq \gamma\} \subset A$ . Sejam  $\bar{y} \in B_r$  e  $b \geq \gamma$ , então existem  $x \in (C - x_*)_1 \subset C(x_*)$ ,  $y \in (K - F(x_*))_1 \subset K(F(x_*))$  com  $\bar{y} = y - dF(x_*)x$ . Tomando  $a := b - dJ(x_*)x$ , temos  $a \geq b - \gamma \geq 0$  e  $b = dJ(x_*)x + a$ . Logo  $(\bar{y}, b) \in A$ .

(3) Provamos que

$$\langle \rho, y - dF(x_*)x \rangle + \eta(dJ(x_*)x + a) \geq 0, \forall x \in C(x_*), y \in K(F(x_*)), a \geq 0.$$

De (1) e (2) podemos aplicar um teorema de separação a  $(\theta, 0) \in \partial A$  (veja [We]) que, neste caso particular, nos garante a existência de  $\lambda \in (Y \times \mathbb{R})^*$ ,  $\lambda \neq \theta$  satisfazendo

$$\langle \lambda, z \rangle \geq \langle \lambda, (\theta, 0) \rangle, \forall z \in A.$$

Podemos decompor  $\lambda$  em  $\lambda = (\rho, \eta) \in Y^* \times \mathbb{R}$ , onde  $\|\rho\| + |\eta| \neq 0$ . Temos então para  $z = (\bar{y}, b) \in A$

$$0 \leq \langle \lambda, (\bar{y}, b) \rangle = \langle \rho, \bar{y} \rangle + \eta b.$$

A desigualdade desejada segue agora da definição de  $A$ .

(4) Provamos que  $\eta > 0$ .

Tomando  $x = y = \theta$  em (3) obtemos  $\eta a \geq 0$  para todo  $a \geq 0$ ; logo  $\eta \geq 0$ . Se  $\eta = 0$ , então

$$\langle \rho, dF(x_*)x - y \rangle \leq 0, \quad \forall x \in C(x_*), y \in K(F(x_*)).$$

Entretanto, do Corolário A.2.7 temos  $dF(x_*)C(x_*) - K(F(x_*)) = Y$ , o que implica em

$$\langle \rho, \bar{y} \rangle \leq 0, \quad \forall \bar{y} \in Y.$$

De onde concluímos que  $\rho = \theta$ , o que contradiz o fato de  $\|\rho\| + |\eta| \neq 0$ .

(5) Construção de  $\psi \in C(x_*)^*$ ,  $\phi \in K(F(x_*))^*$ .

Podemos supor sem perda de generalidade que  $\eta = 1$ . Tomando  $x = \theta$ ,  $a = 0$  em (3) temos

$$\langle \rho, y \rangle \geq 0, \quad \forall y \in K(F(x_*)),$$

provando que  $\rho \in K(F(x_*))^*$ . Tomando agora  $y = \theta$ ,  $a = 0$  em (3) obtemos

$$dJ(x_*)x - \langle \rho, dF(x_*)x \rangle \geq 0, \quad \forall x \in C(x_*),$$

provando que  $dJ(x_*) - \rho \circ dF(x_*) \in C(x_*)^*$ .

A equação (A.4) segue agora tomando  $\psi := dJ(x_*) - \rho \circ dF(x_*)$  e  $\phi := \rho$ , completando assim a demonstração. ■

**Teorema A.3.4** *Seja  $x_*$  um mínimo local de  $(P)$ . Se  $x_*$  é um ponto fracamente regular para  $(P)$ , então existem  $\eta \geq 0$ ,  $\psi \in C(x_*)^*$ ,  $\phi \in K(F(x_*))^*$  satisfazendo a condição de Fritz-John:*

$$\eta dJ(x_*) - \phi \circ dF(x_*) = \psi, \tag{A.5}$$

onde  $|\eta| + \|\psi\| + \|\phi\| \neq 0$ .

**Demonstração:** Se  $x_*$  for ponto regular para  $(P)$ , basta aplicar o Teorema A.3.3 e obtemos (A.5) com  $\eta = 1$ .

Suponha agora que  $x_*$  é ponto fracamente regular para  $(P)$ , mas não é regular. Definindo o conjunto  $A := dF(x_*)(C - x_*) - (K - F(x_*))$ , temos que  $\text{int}(A) \neq \emptyset$ ,  $A$  é convexo,  $\theta \notin \text{int}(A)$ . Portanto, o mesmo teorema de separação usado na demonstração do Teorema A.3.3 (veja [We]) nos garante a existência de  $\phi \in Y^*$ ,  $\phi \neq \theta$  satisfazendo

$$\langle \phi, a \rangle \leq 0 \quad \forall a \in A.$$

Da definição de  $A$  segue

$$\langle \phi, dF(x_*)x - y \rangle \leq 0, \quad \forall x \in (C - x_*), y \in (K - F(x_*)).$$

Tomando  $x = \theta$ , temos  $\phi \in (K - F(x_*))^*$ . Tomando  $y = \theta$ , segue que  $-\phi \circ dF(x_*) \in (C - x_*)^*$ . Defina agora  $\psi := -\phi \circ dF(x_*)$ . Note que a Observação A.3.2 garante que

$$C(x_*)^* = (C - x_*)^*, \quad (K - F(x_*))^* = K(F(x_*))^*.$$

Logo,  $\psi \in C(x_*)^*$ ,  $\phi \in K(F(x_*))^*$  e a equação (A.5) segue com  $\eta = 0$ . ■

# Bibliografia

- [Al] ALT, W., *Stabilität mengenwertiger Abbildungen mit Anwendungen auf nichtlineare Optimierungsprobleme*, Bayreuther Mathematische Schriften, 3, 1979
- [An] ANAND, D.K., *Introduction to Control Systems. 2nd ed.*, Pergamon Press, Oxford, 1984
- [AtFa] ATHANS, M. E FALB, P.L., *Optimal Control*, McGraw-Hill Book Company, New York, 1966
- [BaCa] BARDI, M. E CAPUZZO-DOLCETTA, I., *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston, 1997
- [BaNe] BALAKRISHNAN, A.V. E NEUSTADT, L.W., EDS., *Computing Methods in Optimization Problems*, Proceedings of a conference held at University of California, Los Angeles, January 1964, Academic Press, London, 1964
- [Bar] BARNETT, S., *Matrices in Control Theory with Applications to Linear Programming*, Van Nostrand Reinhold Company, London, 1971
- [BaLe] BAUMEISTER, J. E LEITÃO, A., *Optimal exploitation of renewable resource stocks: Necessary conditions*, Optimal Control Applications and Methods, to appear
- [BLS1] BAUMEISTER, J., LEITÃO, A. E SILVA, G.N., *Optimal control problems with infinite horizon whose control laws contain measures I: Maximum principle*, Proceedings of the 3rd International Conference on Nonlinear Dynamics, Campos do Jordão, Brazil, July 2000, to appear
- [BLS2] BAUMEISTER, J., LEITÃO, A. E SILVA, G.N., *Optimal control problems with infinite horizon whose control laws contain measures II: Dynamic programming*, to appear
- [Bc] BELLMAN, R.E., *Introduction to the Mathematical Theory of Control Process, Vol I: Linear equations and quadratic criteria*, Academic Press, New York, 1967
- [BeDr] BELLMAN, R.E. E DREYFUS, S.E., *Applied Dynamic Programming*, Princeton University Press, Princeton, 1962
- [Ber] BERKOVITZ, L.D., *Optimal Control Theory*, Springer-Verlag, New York, 1974

- [Bo] BOLTJANSKI, W.G., *Mathematische Methoden der optimalen Steuerung*, Carl Hanser Verlag, München, 1972
- [Br] BROKATE, M., *Kontrolltheorie*, Vorlesungskript, Universität Hamburg, 1984
- [BMS] BULIRSCH, R., MIELE, A., STOER, J. E WELL, K.H., EDS., *Optimal control: Calculus of Variations, Optimal Control Theory and Numerical Methods*, International Series of Numerical Mathematics, 111, Birkhäuser, Basel, 1993
- [CaHa] CARLSON, D.A. E HAURIE, A.B., *Infinite Horizon Optimal Control: Theory and Applications*, Springer-Verlag, Heidelberg, 1987
- [Cl] CLARK, C.W., *Mathematical Bioeconomics: The Optimal Management of Renewable Resources. 2nd ed.*, John Wiley and Sons, New York, 1990
- [CCM] CLARK, C.W., CLARKE, F.H. E MUNRO, G.R. *The optimal exploitation of renewable resource stocks: Problems of irreversible investment*, *Econometrica*, 47 (1979), 25 - 47
- [CrLi] CRANDALL, M.G. E LIONS, P.L., *Viscosity solutions of Hamilton-Jakobi Equations*, *Trans. of the American Mathematical Society*, 277 (1983), 1 - 42
- [FlRi] FLEMING, W.H. E RISHEL, R.W., *Deterministic and Stochastic Optimal Control*, Springer-Verlag, Berlin, 1975
- [FISo] FLEMING, W.H. E SONER, H.M., *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York, 1993
- [Fö1] FÖLLINGER, O., *Nichtlineare Regelungen I: Grundlagen und harmonische Balance*, Oldenbourg-Verlag, München, 1969
- [Fö2] FÖLLINGER, O., *Nichtlineare Regelungen II: Anwendung der Zustandsebene*, Oldenbourg-Verlag, München, 1970
- [Fö3] FÖLLINGER, O., *Nichtlineare Regelungen III: Ljapunow-Theorie und Popow-Kriterium*, Oldenbourg-Verlag, München, 1970
- [Gan] GANTMACHER, F.R., *Applications of the Theory of Matrices*, Intersciences Publishers, New York, 1959
- [Gol] GOLDENSTINE, H., *A History of the Calculus of Variations, from 17th through the 19th Century*, Springer-Verlag, New York, 1980
- [Gr] GROETSCH, C.W., *Generalized Inverses of Linear Operators*, Dekker, New York, 1977
- [Ha] HALKIN, H., *Necessary conditions for optimal control problems with infinite horizons*, *Econometrica* 42 (1974), 267 - 272



- [Hel] HESTENES, M.R., *Calculus of Variations and Optimal Control Theory*, John Wiley and Sons, London, 1966
- [He2] HESTENES, M.R., *Variational theory and optimal control theory*, in Balakrishnan, A.V. e Neustadt, L.W., eds., *Computing Methods in Optimization Problems*, 1-22, Proceedings of a conference held at University of California, Los Angeles, January 1964, Academic Press, London, 1964
- [Heu] HEUSER, H., *Funktionalanalysis. 3. Aufl.*, B.G. Teubner, Stuttgart, 1992
- [HeLa] HERMES, H.R. E LASALLE, J.P., *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969
- [Ho] HOCKING, L.M., *Optimal Control: An Introduction to the Theory with Applications*, Clarendon Press, Oxford, 1991
- [Is] ISIDORI, A., *Nonlinear Control Systems: An Introduction. 2nd ed.*, Springer-Verlag, Berlin, 1989
- [Je] JETSCHKE, G., *Mathematik der Selbstorganisation: Qualitative Theorie nichtlinearer dynamischer Systeme und gleichgewichtsferner Strukturen in Physik, Chemie und Biologie*, Vieweg-Verlag, Braunschweig, 1989
- [KIF] KNOBLOCH, H.W., ISIDORI, A. E FLOCKERZI, D., *Topics in Control Theory*, DMV Seminar, 22, Birkhäuser, Basel, 1993
- [KnKw] KNOBLOCH, H.W. E KWAKERNAAK, H., *Lineare Kontrolltheorie*, Springer-Verlag, Berlin, 1985
- [Know] KNOWLES, G., *An Introduction to Applied Optimal Control*, Academic Press, San Francisco, 1981
- [Kra] KRABS, W., *Einführung in die Kontrolltheorie*, Wissenschaftliche Buchgesellschaft, Darmstadt, 1978
- [Kre] KREYZIG, R., *Introductory Functional Analysis with Applications*, John Wiley and Sons, New York, 1978
- [LeMa] LEE, E.B. E MARKUS, L., *Foundations of Optimal Control Theory*, John Wiley and Sons, London, 1967
- [Leig] LEIGH, J.R., *Functional Analysis and Linear Control Theory*, Academic Press, London, 1980
- [Leit] LEITÃO, A., *On optimal control problems with infinite horizon*, Proceedings of the 3rd International Conference on Nonlinear Dynamics, Campos do Jordão, Brazil, July 2000, to appear

- [Lei] LEITMANN, G., *The Calculus of Variations and Optimal Control: An Introduction*, Plenum Press, London, 1981
- [Li] LIONS, J.L., *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Heidelberg, 1971
- [LiYo] LI, X. E YONG, J., *Optimal control theory for infinite dimensional systems*, Birkhäuser, Basel, 1994
- [MaSt] MACKI, J. E STRAUSS, A., *Introduction to Optimal Control Theory*, Springer-Verlag, Heidelberg, 1981
- [Mo] MORGAN, F., *Geometric Measure Theory: A Beginner's guide*, Academic Press, Boston, 1988
- [Mu] MURRAY, J.M., Existence theorems for optimal control and calculus of variations problems where the state can jump, *SIAM J. Control and Optimization*, **24** (1986), 412 - 438
- [Pi] PIEHLER, J., *Einführung in die dynamische Optimierung*, B.G. Teubner, Leipzig, 1966
- [PBG] PONTRYAGIN, L.S., BOLTYANSKIJ, V.G., GAMKRELIDZE, R.V. E MISHCHENKO, E.F., *The Mathematical Theory of Optimal Processes*, Interscience Publishers, New York, 1962 (edição original em russo, 1961)
- [Ra] RAMSEY, F.P., *A mathematical theory of saving*, *Economic Journal*, **38** (1928), 543 - 549
- [Ri] RISHEL, R.W., *An extended Pontryagin principle for control systems whose control laws contain measures*, *SIAM J. Control*, **3** (1965), 191 - 205
- [Ru1] RUDIN, W., *Functional Analysis*, McGraw-Hill, New York, 1973
- [Ru2] RUDIN, W., *Real and Complex Analysis*, McGraw-Hill, New York, 1966
- [Sc] SCHECK, F., *Mechanics: From Newton's Laws to Deterministic Chaos. 3rd ed.*, Springer-Verlag, Berlin, 1999
- [SeSy] SEIERSTAD, A. E SYDSÆTER, K., *Optimal Control Theory with Economic Applications*, North-Holland, Amsterdam, 1987
- [SeZh] SETHI, S.P. E THOMPSON, G.L., *Optimal Control Theory: Applications to Management Science*, Martinus Nijhoff Publishing, Boston, 1981
- [SHB] SCHMIDT, W.H., HEIER, K., BITTNER, L. E BULIRSCH, R., EDS., *Variational Calculus, Optimal Control and Applications*, Proceedings of the 12th International Conference in Honour of L. Bittner and R. Kloetzler, Trassenheide, Germany, September, 1996, International Series of Numerical Mathematics, 124, Birkhäuser, Basel, 1998

- [SFG] SKOWRONSKI, J.M., FLASHNER, H. E GUTTALU, R.S., EDS., *Mechanics and control*, Proceedings of the 3rd Workshop on Control Mechanics, in Honor of the 65th Birthday of George Leitmann, January, 1990, Lecture Notes in Control and Information Sciences, 151, Springer-Verlag, Berlin, 1991
- [SiVi] SILVA, G.N. E VINTER, R.B., *Necessary conditions for optimal impulsive control problems*, SIAM J. Control Optimization, 35 (1997), 1829 - 1846
- [So] SONTAG, E.D., *Mathematical Control Theory: Deterministic Finite Dimensional Systems. 2nd ed.*, Springer-Verlag, New York, 1998
- [Sot] SOTOMAYOR, J., *Lições de Equações Diferenciais Ordinárias*, Projeto Euclides, 11, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1979
- [Sn] SNIEDOVICH, M., *Dynamic Programming*, Marcel Dekker, New York, 1992
- [Tr] TROUTMAN, J.L., *Variational Calculus and Optimal Control: Optimization with Elementary Convexity. 2nd ed.*, Springer-Verlag, New York, 1996
- [ViPe] VINTER, R.B. E PEREIRA, F.M.F.L., *A maximum principle for optimal processes with discontinuous trajectories*, SIAM J. Control and Optimization, 26 (1988), 205 - 229
- [Wa1] WALTER, W., *Analysis II. 2. Aufl.*, Springer-Verlag, Berlin, 1991
- [Wa2] WALTER, W., *Gewöhnliche Differentialgleichungen. 6. Aufl.*, Springer-Verlag, Berlin, 1996
- [Wei] WEINSTOCK, R., *Calculus of Variations*, Dover, New York, 1974
- [We] WERNER, J., *Optimization: Theory and Applications*, Vieweg-Verlag, Wiesbaden, 1984
- [Za] ZABCZYK, J., *Mathematical Control Theory: An Introduction*, Birkhäuser, Boston, 1992
- [ZoKu] ZOWE, J. E KURCYSZ, S., *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optimization, 5 (1979), 49 - 62

# Índice

- Bernoulli, Jakob, 16
- Bernoulli, Johann, 16
- cálculo variacional
  - condição (de contorno) natural, 12
  - condição necessária, 11, 22
  - condição suficiente, 8, 11
  - problemas de fronteira livre, 11, 31
  - problemas de horizonte livre, 31
- ciclóide, 17
- condição
  - de contorno, 4
  - de contorno natural, 12, 30, 31, 41, 44, 52
  - de contorno transversal, 4, 29
  - de Fritz–John, 83, 104
  - de Kuhn–Tucker, 82, 103
  - de mínimo, 45, 52
  - de não acoplamento, 51, 57
  - de otimalidade, 45, 51, 56
  - de salto, 59
  - de Slater, 83
  - de transversalidade, 31, 44
  - de Weierstraß–Erdmann, 24
- cone, 103
  - dual, 82, 102
  - linearizado, 100
  - tangencial
    - a um conjunto por um ponto, 82, 97
    - em um ponto por um conjunto, 82, 96
- conjunto de nível, 35
- controle
  - ótimo
    - condição de mínimo, 51, 56, 59
    - condição de salto, 59
    - condição necessária, 51, 56, 59
    - admissível, 42
    - Bang–Singular–Bang, 41
    - Bang–Bang, 37
    - impulsivo, 58
- derivada de Frechet, 96
- du Bois–Reymond, Paul, 9
- equação
  - adjunta, 51, 56
  - de estado, 51, 56
  - de Euler–Lagrange, 8, 11, 23, 54
  - primeira integral, 14
- espaço
  - $C[a, b]$ , 4, 5
  - $C^1[a, b]$ , 4, 5
  - $C_0^1[a, b]$ , 9
  - $\hat{C}[a, b]$ , 18
  - $\hat{C}^1[a, b]$ , 18
  - de funções teste, 9
- Euler, Leonhard, 9
- exemplo
  - alunissagem, 62
  - Braquistócrona, 16
  - consumo  $\times$  investimento, 67
  - controle ótimo a uma curva alvo, 31
  - geodésicas na esfera, 14
  - geodésicas no plano, 14
  - lançamento de um foguete, 36
  - pescaria ótima, 69
  - tempo mínimo, 60
- extremal, 13
  - singular, 39
- função

- admissível, 4
- de Hamilton, 42, 50
- de identificação, 39
- estacionária, 13
- teste, 9
- funcional
  - convexo, 6
  - estritamente convexo, 6
- Galilee, Galileo, 16
- grau de liberdade, 53
- Lagrange, Joseph Louis, 9
- Leibniz, Gottfried Wilhelm, 16
- lema
  - du Bois-Reymond, 9
  - Lagrange, 9
- mínimo
  - global, 4
  - local
    - forte, 5
    - fraco, 5, 21
- Newton, Isaac, 16
- Pareto otimalidade, 25
- ponto fracamente regular, 98
- ponto regular, 98
- princípio
  - do máximo
    - controle impulsivo, 58
    - hamiltoniana convexa, 46
    - horizonte finito, 51
    - horizonte infinito, 56
- problema de Bolza, 50
- problema de Lagrange, 50
- problema de Mayer, 50
- problema variacional, 3
- problemas anormais, 52
- processo
  - admissível, 42
- restrição
  - isoperimétrica, 4, 35
  - lagrangeana, 4, 33
- sistema
  - Hamiltoniano, 55, 58
- teorema
  - da aplicação aberta, 100
  - fundamental do cálculo, 18
  - multiplicadores de Lagrange, 29, 95
- trajetória
  - admissível, 42
- transformação no tempo, 83
- variação de Gâteaux, 5
- Weierstraß, Karl, 9

Impresso na Gráfica do



pelo Sistema Xerox / 5390