



**INSTITUTO DE BIOCIÊNCIAS, LETRAS E CIÊNCIAS EXATAS
IBILCE / UNESP - CÂMPUS DE SÃO JOSÉ DO RIO PRETO**

Pedro Otávio de Souza Mussatto

Princípio de Pontrjagin e Formalismo de Dubovickij-Miljutin

São José do Rio Preto

2024

Pedro Otávio de Souza Mussatto

Princípio de Pontrjagin e Formalismo de Dubovickij- -Miljutin

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática junto ao Programa de Pós-Graduação em Matemática do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, *campus* de São José do Rio Preto

Financiador: CNPq — Processo nº 165080/2021-3

Orientador: Prof. Dr. Geraldo Nunes Silva

São José do Rio Preto

2024

M989p Mussatto, Pedro Otávio de Souza
Princípio de Pontrjagin e Formalismo de Dubovickij-Miljutin /
Pedro Otávio de Souza Mussatto. -- São José do Rio Preto, 2024
122 p.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP),
Instituto de Biociências Letras e Ciências Exatas, São José do Rio
Preto
Orientador: Geraldo Nunes Silva

1. Matemática. 2. Análise matemática. 3. Otimização matemática. I.
Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Universidade
Estadual Paulista (UNESP), Instituto de Biociências Letras e Ciências Exatas, São José do Rio
Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Pedro Otávio de Souza Mussatto

Princípio de Pontrjagin e Formalismo de Dubovickij-Miljutin

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática junto ao Programa de Pós-Graduação em Matemática do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, *campus* de São José do Rio Preto

Financiador: CNPq — Processo nº 165080/2021-3

Comissão Examinadora

Prof. Dr. Geraldo Nunes Silva

Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, *campus* de São José do Rio Preto

Orientador

Prof. Dr. Valeriano Antunes de Oliveira

Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, *campus* de São José do Rio Preto

Prof. Dr. Marko Antonio Rojas Medar

Facultad de Ciencias da Universidad de Tarapacá

São José do Rio Preto

13 de março de 2024

Agradecimentos

A nosso Senhor por me dar as graças e bênçãos necessárias para prosseguir com meus estudos, sem elas, é provável que a primeira dificuldade me fizesse desistir.

À minha família pelo apoio emocional dado durante esses estudos, especialmente, à minha mãe por me acompanhar durante toda a viagem à Itália e por ser meu abrigo durante o período final de elaboração da dissertação, que foi muito estressante.

A meu orientador, Prof. Dr. Geraldo Nunes Silva, pelo comprometimento em guiar meus estudos, me avisar da oportunidade de estudar na Itália, pela influência direta na segunda parte desta dissertação e pelo acompanhamento feito durante todo o período.

Ao Centro Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por financiar os estudos por meio do processo 165080/2021-3 e pela ajuda com as questões burocráticas ligadas à remuneração do estágio na Itália.

Ao Conselho de Curso do Programa de Pós-Graduação em Matemática e Seção Técnica de Pós-Graduação do Ibilce/Unesp pela oportunidade de estudar no programa e pela ajuda com as questões burocráticas ligadas à remuneração do estágio na Itália e à conclusão do curso.

À Pró-Reitoria de Pós-Graduação da Unesp pela oportunidade de cursar o curso preparatório para o TOEFL ITP e a Douglas Cunha dos Santos por ministrar esse curso e me auxiliar na preparação para o exame.

Aos professores doutores Marko Antonio Rojas Medar e Valeriano Antunes de Oliveira por influenciarem, diretamente, na segunda parte da dissertação indicando o tema e contribuindo com material e resultados necessários para a construção dela.

Ao Departamento de Matemática da Universidade de Pádua por me permitir realizar o estudo que constitui a primeira parte da dissertação por meio de um estágio de três meses pelo programa *Marco Garutti*, em particular, à Sr^a. Laura Provasi, que me acompanhou durante todo o processo seletivo e o estágio me ajudando com questões burocráticas, e à Prof^a. Dr^a. Monica Motta, minha orientadora durante esse estágio e que me ajudou em minha fixação na cidade.

Resumo

A primeira parte desta dissertação de Mestrado apresenta uma versão de mais alta ordem do *Princípio de Pontrjagin*, resultado de extrema importância na Teoria de Controle Ótimo e que marcou sua emergência como área de pesquisa distinta na Matemática, para problemas de controle espaço-temporais cujos lagrangianos crescem, sub linearmente, em relação às variáveis de controle. A principal consequência de tal crescimento é o possível aparecimento de impulsos no controle ótimo. A demonstração combina as clássicas *variações de agulha*, usadas na demonstração da versão de primeira ordem, com as recém-introduzidas *variações de tipo-colchete* construídas a partir de colchetes de Lie arbitrários de campos de vetores (em contraste com outros resultados similares, que requerem comutatividade) e com um argumento padrão de *separação de conjuntos*. Já a segunda parte apresenta o *formalismo de Dubovickij-Miljutin* sob uma nova abordagem. Trata-se de um resultado geral de Otimização em espaços vetoriais topológicos introduzido em 1962 e cuja versão clássica se vale de três tipos de cones: de direções de descida, de direções factíveis e de direções tangentes. Sua essência está no fato de que a disjunção entre o conjunto viável e o subnível do funcional objetivo com fronteira no nível ótimo causa uma separação dos cones apresentados. Uma desvantagem dessa abordagem clássica é a de que quase todos os cones, com exceção, possivelmente, dos de direções tangentes, são abertos, o que complica a aplicação do formalismo em problemas reais. Dito isso, a abordagem utilizada aqui se vale do *cone tangente* introduzido por Frank H. Clarke na década de 1970, que é um cone convexo e fechado por natureza e que satisfaz essa mesma separação sob a hipótese mais fraca de isolamento do minimizador.

Palavras-chave. Controle Ótimo Impulsivo. Colchetes de Lie. Princípio de Pontrjagin. Cone tangente. Princípio da Interseção. Formalismo de Dubovickij-Miljutin.

Abstract

The first part of this Master's thesis outlays a higher-order version of *Pontrjagin's Principle*, a landmark result in Optimal Control Theory that marked its rise as a one-of-a-kind research area in Mathematics, for space-time control problems whose Lagrangians grow sub-linearly with respect to the control variables. The main outcome of this growth is the possible appearance of *impulses* in the optimal control. The proof melds the classical *needle variations*, wielded in proving the first-order version, with the newly brought *bracket-like variations* crafted with arbitrary Lie brackets of vector fields (unlike other similar results, which need commutativity) and with a standard *set separation* argument. The second part brings a version of the *Dubovickij-Miljutin formalism* under a new approach. It's a broad result of Optimization in topological vector spaces introduced in 1962 and whose classical version uses three kinds of cones: descent directions, feasible directions and tangent directions. Its proof's core lies in the fact that the disjunction between the feasible set and the objective functional's sublevel with boundary the optimal level causes a separation of the presented cones. A disadvantage of this classical approach is that almost all cones, with the exception perhaps of the tangent directions' one, are open, which toughens the formalism's application to real problems. That said, the approach taken here wields the *tangent cone* introduced by Frank H. Clarke in the 1970s, which is a convex and closed cone fulfilling the same separation under the weaker assumption of the minimizer's isolation.

Keywords. Impulsive Optimal Control. Lie brackets. Pontrjagin's Principle. Tangent cone. Intersection Principle. Dubovickij-Miljutin formalism.

Lista de notações

Notação	Significado
e_j	j° vetor canônico de um espaço euclidiano
$\text{ess inf } f$	ínfimo essencial de um funcional f
$\text{int } C$	interior de um conjunto C
$\bar{C}$	fecho de um conjunto C
$\text{diam } C$	diâmetro de um conjunto C
$\text{card } C$	cardinalidade de um conjunto C
$\mathfrak{D}(\mathbf{x})$	dicionário formal de uma família $\mathbf{x}$ de caracteres
$\mathfrak{ib}(\mathbf{x})$	conjunto de colchetes iterados formais de uma família $\mathbf{x}$ de caracteres
$\mathcal{C}^1(X; \mathbb{R}^n)$	conjunto das aplicações $f: X \rightarrow \mathbb{R}^n$ que admitem diferenciais que são contínuas em todo ponto de X
$\mathcal{C}^{B,k}$	conjunto das úplas de campos de vetores de classe k de diferenciabilidade em relação a um colchete B
$\mathfrak{L}^k$	conjunto de colchetes de Lie de comprimento ≥ 2 e de classe k de diferenciabilidade
X'	dual algébrico do espaço vetorial X
$\langle x, \varphi \rangle$	valor de um funcional linear $\varphi \in X'$ em um ponto $x \in X$
$\langle Y, \varphi \rangle$	conjunto dos valores de um funcional linear $\varphi \in X'$ em pontos do conjunto $Y \subseteq X$
$\langle x, \Psi \rangle$	conjunto dos valores de funcionais lineares de um conjunto $\Psi \subseteq X'$ em um ponto $x \in X$
$\varsigma(X, \Phi)$	topologia gerada por conjuntos $V(x_0, \varphi, \epsilon) := \{x \in X \mid \langle x - x_0, \varphi \rangle < \epsilon\}$ com $x_0 \in X$, $\varphi \in \Phi$ e $\epsilon > 0$
$x_i \xrightarrow{\Phi} x$	indica que $\{x_i\} \subseteq X$ converge para x na topologia $\varsigma(X, \Phi)$
$B_X(x_0, r)$	bola aberta, no espaço X , de centro x_0 e raio $r > 0$
$B_X[x_0, r]$	bola fechada, no espaço X , de centro x_0 e raio $r > 0$
$\ \cdot\ _p$	norma do espaço L^p para $p \in [1, \infty]$
$B_p(x_0, r)$	bola aberta, no espaço L^p , de centro x_0 e raio $r > 0$
$B_p[x_0, r]$	bola fechada, no espaço L^p , de centro x_0 e raio $r > 0$
$\wp(Y)$	conjunto das partes de Y
$\Gamma: X \rightsquigarrow Y$	multifunção Γ de X para Y
$\text{gr } \Gamma$	gráfico da multifunção Γ
$\text{def } \Gamma$	domínio efetivo da multifunção Γ
$\Gamma^{-1}: Y \rightsquigarrow X$	inversa da multifunção Γ

$\text{co } K$	envoltória convexa do conjunto K
K^Δ	cone polar do conjunto K
$\text{aff } A$	envoltória afim do conjunto A
$\text{ri } C$	interior relativo do conjunto C
$f^\circ(x_0; v)$	derivada direcional geral de um funcional f em um ponto x_0 e em uma direção v
$\partial f(x_0)$	diferencial geral de um funcional f em um ponto x_0
$F'_s(x_0)$	derivada estrita de uma aplicação F entre dois espaços de Banach reais em um ponto x_0
$\mathfrak{C}(Q; x_0)$	cone contingente a um conjunto $Q \subseteq X$ em um ponto x_0
$\mathfrak{T}(Q; x_0)$	cone tangente a um conjunto $Q \subseteq X$ em um ponto x_0
$\text{epi } f$	epígrafo de um funcional f
$\mathfrak{H}(Q; x_0)$	cone hipertangente a um conjunto $Q \subseteq X$ em um ponto x_0
$\mathfrak{N}(Q; x_0)$	cone normal a um conjunto $Q \subseteq X$ em um ponto x_0
$\widehat{\mathfrak{N}}(Q; x_0)$	cone normal estrito a um conjunto $Q \subseteq X$ em um ponto x_0

Sumário

1	Introdução	17
1.1	Cálculo das Variações	17
1.2	Controle Ótimo	18
1.2.1	Problemas impulsivos	20
1.3	Unificação da otimização	21
1.4	Estrutura da dissertação	22
2	Conceitos Preliminares	25
2.1	Colchetes de Lie	25
2.2	Análise não-suave em espaços de Banach	27
2.2.1	Diferenciais gerais	27
2.2.2	Regras de derivação	29
2.3	Cones e polaridade	33
2.3.1	Cones aproximantes de Boltjanskij	39
2.4	Princípios Variacionais	42
2.4.1	Penalização exata	42
2.4.2	Princípio Variacional de Ekeland	43
3	Princípio de Pontrjagin	45
3.1	Problemas de controle ótimo	45
3.1.1	Problemas sentido-estrito	45
3.1.2	Problemas espaço-temporais	46
3.2	Versão de primeira ordem	50
3.3	Versão de alta ordem	53
3.3.1	Sistemas lineares	54
3.3.2	Processos, totalmente, impulsivos	55
3.3.3	Processos não impulsivos	57
3.4	Exemplo de aplicação	59
3.5	Demonstrando a versão de alta ordem	63
3.5.1	Aproximações das variáveis de controle	65
3.5.2	Aproximações compostas	69
3.5.3	Separação de conjuntos	72
3.5.4	Extensão por Interseção Finita	75
4	Formalismo de Dubovickij-Miljutin com cone tangente	79
4.1	Introdução	79

4.2	Cones especiais	79
4.2.1	Cone contingente	79
4.2.2	Cone tangente	80
4.2.3	Cone hipertangente	82
4.2.4	Cone normal estrito	83
4.2.5	Cone normal	84
4.2.6	Relações entre os cones	84
4.3	Formulando o formalismo	87
4.3.1	Diferenciais gerais de funcionais com valores na reta estendida	87
4.3.2	Resultado principal	91
5	Considerações finais e próximos passos	93
	Referências	95
	Apêndice A Conceitos de Análise Funcional	99
A.1	Teorema de Hahn-Banach	99
A.2	Topologias fracas	101
A.3	Espaços normados	108
A.3.1	Topologias fracas em espaços normados	112
	Apêndice B Multifunções e limites	115
	Apêndice C Envoltória afim e interior relativo	119

1 Introdução

Problemas de otimização podem ser estudados em dimensão finita ou infinita. No primeiro caso, são abordados modelos matemáticos que se enquadram em minimizar uma função real definida sobre um espaço euclidiano. Já no segundo caso, são abordados modelos nos quais se minimiza um funcional real definido em um espaço de Banach real, *id est*, um espaço vetorial real que admite uma norma e, em relação à métrica induzida por tal norma, é completo. Em qualquer um desses modelos, é possível que o conjunto de busca das soluções não seja o espaço todo, e isso, na verdade, melhora a exatidão destes em trabalhar com problemas reais — tais problemas modelados são ditos *com restrições*.

Seu surgimento data de meados do século XVII com problemas formulados em espaços funcionais com restrições, somente, em relação aos valores a serem tomados na fronteira, na área de pesquisa que seria, cerca de um século depois, nomeada *Cálculo das Variações*. No século XX, surgem duas novas áreas da otimização distintas do Cálculo das Variações não são aplicáveis: a Programação Matemática, que surge para lidar com problemas de planejamento de produção e de uso de recursos; e a Teoria de Controle Ótimo, que surge para lidar com novas tecnologias aeroespaciais. Neste trabalho, nosso foco será em problemas de otimização em dimensão infinita, começando nas próximas seções, nas quais discutiremos, brevemente, sobre algumas das contribuições no estudo desses problemas. Para uma boa leitura, é necessário ter conhecimentos, principalmente, de Topologia Geral, Análise Real e Análise Convexa.

1.1 Cálculo das Variações

De acordo com Herman H. Goldstine [22], os primeiros trabalhos de impacto em problemas de otimização podem ser vistos com as duas últimas cartas de Pierre de Fermat a Marin C. de la Chambre, enviadas em 1662 e intituladas *Análise de refrações* e *Síntese de refrações*, nas quais Fermat enuncia o princípio de que “a natureza opera de maneiras e métodos que são *mais fáceis e mais rápidos*” clareando que não se trata de atuar no *caminho mais curto*, mas, sim, no de *menor tempo* e citando o trabalho de Galileu Galilei em relação à gravidade. Esse trabalho foi a inspiração de Johann Bernoulli para formulação do *problema da braquistócrona* em 1696²², que desempenhou papel importante no desenvolvimento do Cálculo Variacional²⁰. Ainda no século XVII, destaca-se, também, o *problema da resistência mínima* formulado e resolvido, pela primeira vez, por Isaac Newton [36], mais precisamente, sua Proposição 34 afirma que, “se, em um meio consistindo de partículas iguais dispostas em distâncias iguais de cada uma” (ou seja, um fluido homogêneo), “um globo e um cilindro descritos no mesmo diâmetro se movem com velocidades iguais na direção do eixo do cilindro, a resistência do globo será, porém, metade daquela experimentada pelo cilindro.”

Porém, o grande marco nessa área são os trabalhos de Leonhard Euler [18] e Joseph-Louis

Lagrange [26], que, a partir de 1755, iniciaram um estudo uma condição necessária de otimalidade para problemas do tipo

$$\begin{cases} \text{minimizar} & I(\mathbf{x}) := \int_a^b L(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \\ \text{sujeito a} & \mathbf{x} \in \mathcal{C}^1([a, b]; \mathbb{R}^n), \\ & \mathbf{x}(a) = \mathbf{x}_a, \quad \mathbf{x}(b) = \mathbf{x}_b, \end{cases} \quad (1.1)$$

em que $L \in \mathcal{C}^1([a, b] \times \mathbb{R}^{2n}; \mathbb{R})$ e $\mathbf{x}_a, \mathbf{x}_b \in \mathbb{R}^n$ são dados. A principal motivação para essa equação foi o *problema da tautócrona*, originalmente resolvido por Christiaan Huygens [24], porém, usando uma abordagem geométrica, enquanto que a abordagem de Euler e Lagrange foi analítica. É, também, com Euler que a área de pesquisa recebe o nome de *Cálculo das Variações*, no trabalho [19].

Já no século XVIII, destacam-se, também, as condições de Weierstrass-Erdmann, que estendem o trabalho de Euler e Lagrange para problemas como (1.1) nos quais as curvas $\mathbf{x}$, ao invés de suaves, são suaves por partes. Por fim, mencionamos, também, os trabalhos de David Hilbert e Konstantínos Karatheodorí na virada para o século XX, sendo este o primeiro a diferenciar *variáveis de estado* de *variáveis de controle* sem, no entanto, adotar essa nomenclatura.

1.2 Controle Ótimo

Antes de discutir sobre Controle Ótimo, destacamos que, visando melhorias na organização do planejamento e da produção, ou seja, melhorias no processo de distribuição do trabalho entre as máquinas da empresa, distribuição de matéria prima, combustível, entre outros fatores¹⁴, quase que simultaneamente ao início da Segunda Guerra Mundial, surge a Otimização Linear a partir dos trabalhos de Leonid V. Kantorovič [25] e George B. Dantzig e que, segundo Igor' V. Girsanov [21], estuda métodos especiais para determinar extremas de funções suaves em domínios fechados com fronteiras suaves por partes.

Com o advento da Guerra Fria, surge um novo problema cuja solução parecia estar ao alcance dos matemáticos da época, mas que escondia elementos não-reveláveis pelo Cálculo das Variações: encontrar a trajetória ótima para guiar um avião de uma posição inicial até uma posição mais favorável para atacar aviões inimigos³⁹, denominado *problema de tempo mínimo de interceptação*. A princípio, tal problema parece ser uma variação do problema da braquistócrona com, somente, a substituição da força da gravidade pela resultante da gravidade impactando na propulsão do motor do avião. No entanto, de acordo com Hans J. Pesch e Michael Plail [39], a alta velocidade do avião fez com que termos não-lineares não mais pudessem ser ignorados e diminuiu a eficácia de uma aproximação linear para resolução. No chamado *bloco ocidental* (isto é, nos Estados Unidos e em seus países aliados), através da corporação RAND, Magnus R. Hestenes [23] conseguiu formular a condição que, mais tarde, seria conhecida como Princípio de Pontrjagin utilizando a condição necessária de otimalidade de Weierstrass-Erdmann, no entanto, sua experiência no Cálculo Va-

riacional (orientando de Gilbert A. Bliss, que, por sua vez, foi orientado por Oskar Bolza, aluno de Karl T. W. Weierstrass e de Felix C. Klein) o impediu de separá-la dessa área e buscar uma generalização, e o memorando citado não foi muito divulgado fora da corporação RAND. Outra contribuição ainda mais importante é o nascimento da Programação Dinâmica por Richard E. Bellman [6], que acertou ao criticar o Cálculo das Variações pela impossibilidade resolver o problema de valor contorno com equações diferenciais não-lineares resultante³⁹. Johann Baumeister e Antonio Leitão [5] afirmam que essa abordagem pode ser aplicada para solucionar problemas de controle bastante gerais (não lineares, não autônomos) e constitui uma alternativa ao Princípio de Pontrjagin (no entanto, seus métodos computacionais são caros comparados a este).

Já no *bloco oriental* (isto é, na União Soviética e em seus países aliados), a situação era muito diferente. Lev S. Pontrjagin, formado topólogo, apenas recentemente mudara seu interesse para o problema de tempo mínimo de interceptação após ser ultrapassado pelos franceses na Topologia e sob pressão do diretor do Departamento de Matemática do Instituto Steklov e da organização do Partido Comunista no instituto. O preparo já começara em 1952 com um seminário conduzido por Pontrjagin e por Mark A. Ajzerman, especialista em controle automático, porém, apenas três anos depois, com o encontro do grupo liderado por Pontrjagin com membros da força aérea, os trabalhos no problema de controle de tempo ótimo se iniciaram, e o grupo contou com engenheiros para fortalecer o caráter de aplicação da teoria, em especial, Alexander Ja. Lerner e Alexander Fel'dbaum focaram a atenção em processos ótimos para controle automático³⁹. Após um trabalho forte e rápido, Pontrjagin e seus colegas Vladimir G. Boltjanskij e Revaz V. Gamkrelidze conseguiram, um ano depois, publicar uma primeira versão do Princípio, que foi imediatamente, comprovada por meio de problemas do tipo Bushaw-Fel'dbaum³⁹, ou seja, problemas em que um ponto no espaço de fase posição $\times$ velocidade deve ser conduzido à origem em tempo mínimo sob o sistema dinâmico $\ddot{x} + x = u$ com controle de norma limitada por um, no qual pode ser encaixado o problema trabalhado pondo x como a distância entre o avião e um míssil. Gamkrelidze, então, mostrou que tal princípio era condição necessária e suficiente para problemas lineares, e Boltjanskij completou o trabalho provando que se tratava de uma condição necessária para problemas gerais, o que culminou na publicação de [40]. Esse resultado, de acordo com Richard Vinter [48], é considerado marco na emergência da Teoria de Controle Ótimo como área de pesquisa distinta e é usado para resolver problemas tanto de maneira analítica quanto por métodos computacionais pela localização de “candidatos” chamados, formalmente, de *extremais*.

No entanto, não é garantia que um extremal encontrado seja solução do problema — não se trata de uma complicação exclusiva de problemas de Controle Ótimo, mas abrange problemas de otimização em geral; por exemplo, a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $(x, y) \mapsto xy$ possui a origem como extremal, porém, tal extremal não é minimizador (na verdade, tal função não possui minimizador, porém, não é o foco estudá-la). Diante disso, torna-se de grande valor encontrar condições necessárias de mais alta ordem para melhor filtrar os extremais obtidos.

Os primeiros problemas para os quais o princípio pode ser formulado são os denominados *sensitivo-estrito*: tratam-se de problemas da forma

$$\begin{aligned}
&\text{minimizar} && \Psi(T, x(T)) + \int_0^T \ell(x, u, a)(t) dt \\
&\text{sujeito a} && T > 0, \quad (u, a) \in L^1([0, T]; C \times A), \quad (x, v) \in \mathcal{AC}([0, T]; \mathbb{R}^n \times [0, \infty[), \\
&&& \begin{cases} \dot{x} = f(x, a) + \sum_{i=1}^m g_i(x) u_i, \\ \dot{v} = |u| \end{cases} \quad \text{quase sempre em } [0, T], \\
&&& (x, v)(0) = (\check{x}, 0), \quad (T, x(T)) \in \mathcal{T}, \quad v(T) \leq K,
\end{aligned} \tag{1.2}$$

em que:

- $K \in [0, \infty]$ e $\check{x} \in \mathbb{R}^n$;
- O alvo $\mathcal{T}$ é um subconjunto fechado de $[0, \infty[\times \mathbb{R}^n$;
- A é um subconjunto compacto de $\mathbb{R}^q$;
- C é um cone fechado em $\mathbb{R}^m$;
- $\mathcal{AC}([0, T]; \mathbb{R}^n \times [0, \infty[)$ é o conjunto dos pares ordenados de aplicações, absolutamente, contínuas com domínio em $[0, T]$ cuja primeira coordenada tem contradomínio em $\mathbb{R}^n$ e segunda coordenada tem domínio em $[0, \infty[$;
- $f: \mathbb{R}^n \times A \rightarrow \mathbb{R}$, $g_1, \dots, g_m: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\ell: \mathbb{R}^n \times C \times A \rightarrow \mathbb{R}$ e $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}$ satisfazem hipóteses padrões de regularidade¹.

1.2.1 Problemas impulsivos

Aqui, serão, também, feitas as hipóteses de que os controles u são ilimitados e que o lagrangiano ℓ cresce, sublinearmente, em relação a estes. De acordo com Maria S. Aronna, Monica Motta e Franco Rampazzo [1], essas duas hipóteses são motivadas por diversas aplicações, por exemplo, dinâmicas afins em variáveis de controle ilimitadas descrevem o movimento de corpos rígidos, mutualmente, ligados. Por outro lado, o crescimento demasiado lento do lagrangiano pode causar o aparecimento de *impulsos*, por exemplo, o problema de tempo mínimo no qual o alvo intercepta a órbita de g_1 partindo de $\check{x}$ e a trajetória ótima percorre-a com velocidade infinita¹.

Outros problemas em que esse fenômeno acontece, chamados *problemas de controle impulsivos*, ocorrem, classicamente, no cálculo de trajetórias ótimas de voo (veja, por exemplo, [28, 30]). Dentre as contribuições para o estudo de tais problemas, destacam-se os trabalhos de Raymond W. Rishel [41] e Jack Warga [49], que apresentaram a ideia de “compactificar” o problema (1.2) ao imergi-lo no problema

$$\begin{aligned}
&\text{minimizar} && \Phi(y_0, y)(S) + \int_0^S \ell^e(y, w_0, w, a)(s) \, ds \\
&\text{sujeito a} && S > 0, \\
&&& (w_0, w, \alpha) \in L^\infty([0, S]; [0, \infty[\times C \times A), \\
&&& (y_0, y, \beta) \in \mathcal{AC}([0, S]; [0, \infty[\times \mathbb{R}^n \times [0, \infty[), \\
&&& \text{ess inf}(w_0 + |w|) > 0, \\
&&& \begin{cases} \dot{y}_0 = w_0, \\ \dot{y} = f(y, \alpha)w_0 + \sum_{i=1}^m g_i(y)w_i, \\ \dot{\beta} = |w| \end{cases} \quad \text{quase sempre em } [0, S], \\
&&& (y_0, y, \beta)(0) = (0, \check{x}, 0), \quad (y_0, y)(S) \in \mathcal{T}, \quad \beta(S) \in [0, K].
\end{aligned} \tag{1.3}$$

Problemas dessa classe, denominados *espaço-temporais*, podem ser reduzidos a problemas como (1.2) quando $w_0 > 0$ quase sempre por meio de uma simples reparametrização da variável temporal, enquanto que, quando w_0 se anula em um conjunto de medida positiva, a variável temporal permanece “estacionada” enquanto a variável espacial continua evoluindo, o que caracteriza um *impulso*.

A abordagem em questão é o padrão para estudar problemas impulsivos atualmente e, por conta disso, vários trabalhos a seguem, como, por exemplo, [8, 38, 31, 32, 44], no entanto, a versão apresentada em [1] e que será o foco do estudo aqui possui ordem mais alta e pode ser utilizada para filtrar processos extremais obtidos com a correspondente versão de primeira ordem, como mostra o exemplo da seção 1.6. Seu caráter é geométrico por fazer uso dos colchetes de Lie, ferramenta clássica utilizada para demonstrações de resultados desse caráter (veja, por exemplo, [43, 46]), porém, de acordo com Aronna, Motta e Rampazzo [1], os resultados de segunda ordem obtidos com essa ferramenta lidavam, apenas, com o *caso comutativo*, por exemplo, nos trabalhos [2, 3, 16], de modo que o uso de colchetes arbitrários constitui a principal novidade.

1.3 Unificação da otimização

Apesar de suas inegáveis contribuições para a Teoria de Controle Ótimo, o Princípio de Pontrjagin destacou, ainda mais, a falta de um resultado unificado para tratar de problemas extremais.

Essa lacuna, porém, foi preenchida com a publicação de [15], que estabeleceu uma condição necessária de otimalidade para problemas de otimização em espaços vetoriais topológicos que tomou o formato de uma equação posta na linguagem da Análise Funcional²¹ e que conseguiu recuperar a unidade teórica do Cálculo Variacional até, então, perdida.

Sua demonstração se dá através da aproximação do conjunto viável por cones de direções viáveis e tangentes e a construção de um cone de direções de descida e utilizando uma propriedade de tais

cones: o fato de que o extremal analisado ser minimizador implicar que tais cones são separáveis. A partir daí, é possível encontrar funcionais não-nulos em seus respectivos cones polares cuja somatória é nula.

No entanto, G. G. Watkins [50] menciona que a principal desvantagem da abordagem usada é a de quase todos os cones (especificamente, os de direções de descida e viáveis) serem abertos e que essa desvantagem foi eliminada com a publicação de [7], estabelecendo o chamado *método das tendas*. Esse método, porém, só funciona quando os dados (aqui, entende-se como *dados* as aplicações e/ou funcionais que caracterizam o objetivo e/ou as restrições) são diferenciáveis. Por exemplo, o Teorema de Ljusternik requer que a aplicação que caracteriza a restrição de igualdade seja diferenciável em uma vizinhança do minimizador e que sua derivada seja sobrejetiva nele.

A partir disso, é de grande importância estudar versões desse resultado para problemas não-suaves. Embora Watkins [50] já demonstrara que o *cone tangente* introduzido em [12] satisfaz o *Princípio da Interseção* comentado no Capítulo 1, a demonstração apresentada utiliza diversos conceitos, elementos e resultados válidos para espaços de dimensão finita. O foco deste capítulo é apresentar uma versão para espaços de dimensão infinita utilizando, também, o cone tangente de Clarke e sob a hipótese de interior não-vazio de alguns dos cones, que representa uma melhora em relação à necessidade de abertura.

1.4 Estrutura da dissertação

O capítulo 2 aborda os conceitos preliminares mais importantes para entender os resultados principais da dissertação. Mais especificamente, suas primeiras duas seções apresentam as definições e algumas propriedades do *cone aproximante de Boltjanskij*, que é o aproximante tangencial utilizado na formulação de ambas as versões do Princípio de Pontrjagin e dos *colchetes iterados formais*, incluindo *colchetes de Lie*, objeto utilizado para formular a versão de ordem mais alta. As demais seções são destinadas a conceitos necessários para o entendimento do formalismo de Dubovickij-Miljutin: a primeira delas (terceira do capítulo) apresenta as definições e diversas propriedades das *derivadas direcionais gerais e diferenciais gerais*, utilizados para definir os cones contingente, tangente, hipertangente, normal e normal estrito (nesta dissertação, o termo *cones especiais* será usado para se referir aos cinco cones mencionados quando não houver necessidade de especificar, ao menos, um deles) e mostrar diversas propriedades, incluindo relações entre tais cones; já a segunda apresenta definições e propriedades gerais de *cones* e *cones polares* necessárias para compreender propriedades dos cones especiais, incluindo um teorema de equivalência de separabilidade que será envolvido na formulação do formalismo; por fim, a terceira apresenta o Princípio Variacional de Ekeland utilizado para demonstrar relações entre os cones tangente e contingente.

O capítulo 3 aborda os problemas de Controle Ótimo para os quais o Princípio de Pontrjagin será formulado. Sua primeira seção apresenta as classes de problemas *sentido-estrito*, que podem ser, também, denominados *não-impulsivos* e que, como comentado acima, foram os primeiros proble-

mas para os quais o resultado foi feito, e a classe mais geral dos problemas *espaço-temporais*, que podem apresentar comportamento impulsivo, e o método para imergir os problemas sentido-estrito nela. A segunda seção apresenta a versão de primeira ordem do Princípio de Pontrjagin para problemas espaço-temporais sem, no entanto, realizar sua demonstração, que nada mais é do que uma adaptação da correspondente demonstração de alta ordem com a argumentação envolvendo colchetes retirada, e algumas de suas consequências. A terceira seção apresenta a versão de ordem mais alta do Princípio de Pontrjagin para problemas espaço-temporais e apresenta e demonstra algumas de suas consequências, incluindo o correspondente resultado para problemas sentido-estrito e condições para que um processo ótimo seja, totalmente, impulsivo. A quarta seção é dedicada a apresentar um exemplo no qual essa versão de ordem mais alta é utilizada para filtrar um extremal não-ótimo. Por fim, a quinta seção demonstra essa versão de ordem mais alta.

O capítulo 4 aborda o formalismo de Dubovickij-Miljutin utilizando o cone tangente de Clarke. Sua primeira seção introduz os cones especiais e demonstra diversas propriedades deles, incluindo convexidade, fechamento ou abertura, relações entre estes e relações entre os cones tangente e normal com o gradiente generalizado e com epígrafos. Já sua segunda seção inicia-se com a apresentação de extensões de resultados da primeira para funcionais que não sejam, localmente, lipschitzianos, que são necessárias para extrair a propriedade mais importante dos cones tangente e normal, que diz respeito a interseção de conjuntos e será usada para formular o formalismo, e se encerra com a apresentação do formalismo em si.

O capítulo 5 retoma, brevemente, os conceitos apresentados comentando suas diferenças para outros trabalhos e apresenta uma leve previsão dos trabalhos a serem desenvolvidos a partir do que foi apresentado aqui. No entanto, não há garantia de que esses trabalhos serão, de fato, desenvolvidos, visto que a situação pode mudar ao longo dos anos.

Finalmente, abordamos conceitos e propriedades relacionados a separação de conjuntos convexos, topologias de espaços normados, multifunções e seus limites e envoltória afim e interior relativo. Embora sejam de extrema importância para compreensão completa dos capítulos 2 a 4, tais assuntos foram colocados em apêndices devido à estrutura da dissertação: os conceitos relativos ao Princípio de Pontrjagin são apresentados na mesma ordem em que aparecem no artigo [1], e os relativos ao formalismo de Dubovickij-Miljutin seguem a ordem apresentada em [21].

2 Conceitos Preliminares

2.1 Colchetes de Lie

Em primeiro lugar, veremos algumas definições e propriedades de colchetes formais apresentadas por Aronna, Motta e Rampazzo [1].

Seja $\mathbf{x} := \{x_i \mid i \in \mathbb{Z}_+\}$ uma família de *caracteres*.

Definição 1. As combinações finitas de caracteres de $\mathbf{x}$ (sem repetição), colchetes e vírgulas são denominadas **palavras formais de $\mathbf{x}$** , e o conjunto delas é denominado **dicionário formal de $\mathbf{x}$** e denotado por $\mathfrak{D}(\mathbf{x})$.

Definição 2. Um elemento de $\mathfrak{D}(\mathbf{x})$ é denominado **colchete iterado formal de $\mathbf{x}$** se é escrito como uma combinação de caracteres de $\mathbf{x}$ com elementos da forma $[\cdot, \cdot]$. O conjunto de colchetes iterados formais de $\mathbf{x}$ será denotado por $\mathfrak{ib}(\mathbf{x})$. O número de caracteres de um colchete iterado formal é denominado **comprimento de tal colchete**.

Exemplo 3. Sendo $\mathbf{x} = \{x_1, x_2, x_3\}$, o conjunto de colchetes iterados de $\mathbf{x}$ é

$$\begin{aligned} \mathfrak{ib}(\mathbf{x}) = \{ & x_1, x_2, x_3, [x_1, x_2], [x_2, x_3], [x_3, x_1], [x_2, x_1], [x_3, x_2], [x_1, x_3], [x_1, [x_2, x_3]], [x_3, [x_1, x_2]], \\ & [x_2, [x_3, x_1]], [x_3, [x_2, x_1]], [x_1, [x_3, x_2]], [x_2, [x_1, x_3]], [[x_2, x_3], x_1], [[x_1, x_2], x_3], [[x_3, x_1], \\ & x_2], [[x_2, x_1], x_3], [[x_3, x_2], x_1], [[x_1, x_3], x_2] \} \end{aligned}$$

Propriedade 4. Dados $B_1, B_2 \in \mathfrak{ib}(\mathbf{x})$, o comprimento de $[B_1, B_2]$ é a soma dos comprimentos de B_1 e de B_2 . Ainda, qualquer $B \in \mathfrak{ib}(\mathbf{x})$ admite uma única escrita da forma $[B_1, B_2]$ com $B_1, B_2 \in \mathfrak{ib}(\mathbf{x})$.

Definição 5. O colchete B_1 da **Propriedade 4** é denominado **fator esquerdo de B** , o colchete B_2 desta é denominado **fator direito de B** , e o par (B_1, B_2) é denominado **fatoração de B** .

Definição 6. Sendo $B \in \mathfrak{ib}(\mathbf{x})$, a **contagem de diferenciação de um subcolchete de B em relação a B** é o número $\mathfrak{d}(\cdot; B)$ definindo por: $\mathfrak{d}(B; B) = 0$; se B_1, B_2 são sub colchetes próprios de B , então, $\mathfrak{d}(B_1; B) = \mathfrak{d}(B_2; B) = 1 + \mathfrak{d}([B_1, B_2]; B)$.

Exemplo 7. Sendo $\mathbf{x} = \{x_1, x_2, x_3, x_4, x_5\}$ e $B = \left[x_1, \left[[x_2, x_3], x_4 \right], x_5 \right]$, temos que $\mathfrak{d}(x_1; B) = 1$, $\mathfrak{d}([x_2, x_3]; B) = \mathfrak{d}(x_4; B) = 3$ e $\mathfrak{d}(x_5; B) = 2$.

Definição 8. Sendo $B \in \mathfrak{ib}(\mathbf{x})$ de comprimento $m \in \mathbb{N}$, suponha que os caracteres de B são $x_{\pi(1)}, \dots, x_{\pi(m)}$ (em que $\pi: \mathbb{N} \rightarrow \mathbb{N}$ é uma permutação fixada). Dado $\mu \geq \max \pi(\{1 : m\})$ e $k \in \mathbb{N}$, uma μ -úpla $\mathbf{f} = (f_1, \dots, f_\mu)$ de campos de vetores definidos sobre uma mesma variedade

topológica¹ é dita **de classe k de diferenciabilidade em relação a B** se cada f_j possui classe de diferenciabilidade $\mathfrak{d}(x_{\pi(j)}; B) + k$.

Lema 9. Sendo B , k e f como na **Definição 8** e (B_1, B_2) a fatoração de B , $f \in \mathcal{C}^{B,k}$ se, e somente se, $f \in \mathcal{C}^{B_1, k+1}$ e $f \in \mathcal{C}^{B_2, k+1}$. Aqui, $\mathcal{C}^{B,k}$ denota classe k de diferenciabilidade em relação a B .

Demonstração. Claramente, dizer que $f \in \mathcal{C}^{B,k}$ é equivalente a dizer que cada $f_{\pi(j)}$ possui classe de diferenciabilidade $\mathfrak{d}(x_{\pi(j)}; B) + k$ para $j \in \{1 : m\}$. Sabendo que

$$\begin{aligned} \mathfrak{d}(x_{\pi(j)}; B_2) + 1 &= \mathfrak{d}(x_{\pi(j)}; B_2) + \mathfrak{d}(B_2; B) = \mathfrak{d}(x_{\pi(j)}; B) = \mathfrak{d}(x_{\pi(j)}; B_1) + \mathfrak{d}(B_1; B) \\ &= \mathfrak{d}(x_{\pi(j)}; B_1) + 1, \end{aligned}$$

a equivalência fica provada. □

Definição 10. Dado $B \in \mathfrak{ib}(\mathbf{x})$, o **número de troca de B** é: 1 se o comprimento de B for um; senão, é o dobro a soma dos números de troca de seus fatores esquerdo e direito. Esse número será denotado por $r(B)$.

Exemplo 11. Sendo $\mathbf{x} = \{x_1, x_2, x_3, x_4\}$, temos que

$$\begin{aligned} r(x_1) &= 1, \\ r([x_1, x_2]) &= 2[r(x_1) + r(x_2)] = 4, \\ r([[x_1, x_2], x_3]) &= 2\{r([x_1, x_2]) + r(x_3)\} = 2\{4 + 1\} = 10, \\ r([[x_1, x_2], [x_3, x_4]]) &= 2\{r([x_1, x_2]) + r([x_3, x_4])\} = 4 \cdot 4 = 16, \\ r([[[x_1, x_2], x_3], x_4]) &= 2\{r([[x_1, x_2], x_3]) + r(x_4)\} = 2 \cdot \{10 + 1\} = 22. \end{aligned}$$

Definição 12. Na hipótese geral da **Definição 8**, sejam $\mu \geq \max \pi(\{1 : m\})$ e $f = (f_1, \dots, f_\mu) \in \mathcal{C}^{B,0}$. Sendo s um sub colchete de B , defina o campo de vetores $\mathfrak{b}(s, f)$ como segue: se $s = x_{\pi(j)}$ para algum $j \in \{1 : m\}$, ponha $\mathfrak{b}(s, f) = f_{\pi(j)}$; senão, sendo (s_1, s_2) a fatoração de s , ponha $\mathfrak{b}(s, f) = [\mathfrak{b}(s_1, f); \mathfrak{b}(s_2, f)]$. O campo de vetores $\mathfrak{b}(s, f)$ é denominado **colchete de Lie de f associado a s** .

Por essa definição, é possível tratar os colchetes de Lie como casos particulares de colchetes iterados formais nos quais os caracteres são os devidos campos de vetores. Sendo assim, faz sentido dizer “sendo B um colchete de Lie”, e os conceitos de subcolchetes, fatores, número de troca e contagem de diferenciação estão bem definidos.

Definição 13. Dados $k \in \mathbb{N}$, B um colchete formal iterado e f uma úpla de campos de vetores definidos sobre uma mesma variedade topológica, o colchete de Lie $\mathfrak{b}(B, f)$ é dito **de classe k de diferenciabilidade** se $f \in \mathcal{C}^{B,k}$. O conjunto dos colchetes de Lie $\mathcal{C}^k$ -admissíveis de comprimento maior que um será denotado por $\mathfrak{B}^k$.

¹Os capítulos 4 e 5 de [29] definem o que é uma variedade topológica e o que é uma aplicação diferenciável entre variedades topológicas. Usando tais definições e as de campos de vetores em uma superfície, definidos nos capítulos 2 e 3 de [29], pode-se definir o que é um campo de vetores em uma variedade.

2.2 Análise não-suave em espaços de Banach

2.2.1 Diferenciais gerais

Quando nada contrário for especificado, cada resultado desta subseção será feito supondo que X é um espaço de Banach real e $f: X \rightarrow \mathbb{R}$ é lipschitziano de posto $K > 0$ em uma vizinhança de um ponto $x_0 \in X$.

Definição 14 (página 194 de [11]). Dado $v \in X$, a **derivada direcional geral de f em x_0 na direção v** é o número

$$f^\circ(x_0; v) := \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + tv) - f(x)}{t}.$$

O primeiro resultado fornece algumas propriedades importantes dessa derivada.

Propriedade 15 (página 195 de [11]). *A derivada direcional geral satisfaz as seguintes propriedades:*

- O funcional $f^\circ(x_0; \cdot)$ é lipschitziano de posto K , subaditivo e positivamente, homogêneo;
- A aplicação f° é semicontínua superior;
- Dado $v \in X$, vale $f^\circ(x_0; -v) = (-f)^\circ(x; v)$.

Demonstração. • Dados $v_1, v_2 \in X$, vale

$$\begin{aligned} |f^\circ(x_0; v_1) - f^\circ(x_0; v_2)| &= \left| \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + tv_1) - f(x + tv_2)}{t} \right| \\ &= \limsup_{x \rightarrow x_0, t \downarrow 0} \left| \frac{f(x + tv_1) - f(x + tv_2)}{t} \right| \\ &\leq K \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{t\|v_1 - v_2\|_X}{t} = K\|v_1 - v_2\|_X. \end{aligned}$$

Ainda, dados $v_1, v_2 \in X$, da **Definição 14**,

$$\begin{aligned} f^\circ(x_0, v_1 + v_2) &= \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + t(v_1 + v_2)) - f(x)}{t} \\ &= \lim_{x \rightarrow x_0, t \downarrow 0} \left[\frac{f(y + tv_1) - f(y)}{t} + \frac{f(x + tv_2) - f(x)}{t} \right] \end{aligned}$$

com $y := x + tv_2$. Como $y \rightarrow x_0$, é possível concluir que

$$\begin{aligned} f^\circ(x_0, v_1 + v_2) &\leq \limsup_{y \rightarrow x_0, t \downarrow 0} \frac{f(y + tv_1) - f(y)}{t} + \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + tv_2) - f(x)}{t} \\ &= f^\circ(x_0, v_1) + f^\circ(x_0, v_2). \end{aligned}$$

Por fim, dados $\alpha \geq 0$ e $v \in X$, temos que

$$\begin{aligned} f^\circ(x_0; \alpha v) &= \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + t\alpha v) - f(x)}{t} = \alpha \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + \alpha t v) - f(x)}{\alpha t} \\ &= \alpha \limsup_{x \rightarrow x_0, s \downarrow 0} \frac{f(x + sv) - f(x)}{s} = \alpha f^\circ(x_0, v); \end{aligned}$$

- Dadas sequências $\{x_i\}, \{v_i\} \subseteq X$ tendendo x_0 e v respectivamente, existem $\{y_i\} \subseteq X$ e $\{t_i\} \subseteq]0, \infty[$ de modo que $\|y_i - x_i\|_X + t_i < \frac{1}{i}$ e

$$\begin{aligned} f^\circ(x_i; v_i) - \frac{1}{i} &\leq \frac{f(y_i + t_i v_i) - f(y_i)}{t_i} = \frac{f(y_i + t_i v) - f(y_i)}{t_i} + \frac{f(y_i + t_i v_i) - f(y_i + t_i v)}{t_i} \\ &\leq \frac{f(y_i + t_i v) - f(y_i)}{t_i} + K\|v_i - v\|_X, \end{aligned}$$

assim, ao tomar o $\limsup$ para $i \uparrow \infty$, vem $\limsup_{i \uparrow \infty} f^\circ(x_i; v_i) \leq f^\circ(x_0, v)$;

- Dado $v \in X$ e chamando $y = x - tv$, temos que

$$\begin{aligned} f^\circ(x_0; -v) &= \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x - tv) - f(x)}{t} = \limsup_{y \rightarrow x_0, t \downarrow 0} \frac{(-f)(y + tv) - (-f)(y)}{t} \\ &= (-f)^\circ(x_0; v). \end{aligned}$$

□

Propriedade 16 (página 200 de [11]). *Se f e g são lipschitzianos em uma vizinhança de x_0 , então, qualquer $v \in X$ satisfaz $(f + g)^\circ(x_0; v) \leq f^\circ(x_0; v) + g^\circ(x_0; v)$.*

Demonstração. Dado $v \in X$, segue que

$$\begin{aligned} (f + g)^\circ(x_0; v) &= \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{(f + g)(x + tv) - (f + g)(x)}{t} \\ &\leq \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + tv) - f(x)}{t} + \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{g(x + tv) - g(x)}{t} = f^\circ(x_0; v) \\ &\quad + g^\circ(x_0; v). \end{aligned}$$

□

Definição 17 (página 196 de [11]). A **diferencial geral de f em x_0** , denotado por $\partial f(x_0)$, é o conjunto não-vazio, convexo e compacto fraco^{*} de X^* cujo funcional de suporte é $f^\circ(x_0; \cdot)$.

Essa definição indica que, dado $v \in X$, $f^\circ(x_0, v) = \sup\{\langle v, \zeta \rangle \mid \zeta \in \partial f(x_0)\}$.

Proposição 18 (página 196 de [11]). $\partial f(x_0) \subseteq B_{X^*}[0, K]$.

Demonstração. Dado $\zeta \in \partial f(x_0)$, da **Propriedade 15**, vale $\langle v, \zeta \rangle \leq f^\circ(x_0, v) \leq K\|v\|_X$ para qualquer $v \in X$. $\square$

Proposição 19 (página 197 de [11]). *Se f admite máximo ou mínimo local em x_0 , então, $0 \in \partial f(x_0)$.*

Demonstração. Faremos a demonstração para o caso em que f admite mínimo local, já que, no caso do máximo local, $-f$ admite mínimo local em x_0 e seu gradiente generalizado é formado pelos opostos dos elementos de $\partial f(x_0)$ (**Propriedade 21**). Dito isso, supondo que exista uma direção $v \in X$ na qual $f^\circ(x_0; v) < 0$, em particular, temos que

$$\lim_{i \uparrow \infty} i \left[f\left(x_0 + \frac{v}{i}\right) - f(x_0) \right] < 0,$$

logo, deve existir um $i_0 \in \mathbb{N}$ para o qual $i > i_0 \Rightarrow f(x_0 + \frac{v}{i}) < f(x_0)$, porém, essa conclusão contradiz a minimalidade local de f em x_0 . A partir disso, é possível afirmar que $f^\circ(x_0; v) \geq 0 = \langle 0, v \rangle$ para todo $v \in X$, o que mostra que $0 \in \partial f(x_0)$. $\square$

Proposição 20 (página 199 de [11]). *∂f é uma multifunção de gráfico fechado.*

Demonstração. Dados $x_0 \in X$, $\{x_i\} \subseteq X$ tendendo a x_0 e $\{\zeta_i\} \subseteq X^*$ de modo que $\zeta_i \in \partial f(x_i)$ para todo $i \in \mathbb{N}$, seja ζ um ponto de acumulação de $\{\zeta_i\}$ (que existe pelos fatos de que, sendo K o posto de f como aplicação lipschitziana, $\{\zeta_i\} \subseteq B_{X^*}[0, K]$ e de que, pelo **Corolário 178**, este é compacto fraco^{*}). Sob essa suposição, fixado $v \in X$, a sequência $\{\langle v, \zeta_i \rangle\}$ é limitada em $\mathbb{R}$ e contém termos, arbitrariamente, próximos a $\langle v, \zeta \rangle$, o que permite extrair uma subsequência $\{\zeta_j\}$ de $\{\zeta_i\}$ de modo que $\langle v, \zeta_j \rangle \rightarrow \langle v, \zeta \rangle$. A partir disso, sabendo, da **Propriedade 15**, que f° é semicontínua superior, vem

$$f^\circ(x_0; v) = \lim_{j \rightarrow \infty} f^\circ(x_j; v) \geq \lim_{j \rightarrow \infty} \langle v, \zeta_j \rangle = \langle v, \zeta \rangle.$$

Da arbitrariedade de v , concluímos que $\zeta \in \partial f(x_0)$. $\square$

2.2.2 Regras de derivação

Propriedade 21 (página 199 de [11]). *Para qualquer $\lambda \in \mathbb{R}$, vale $\partial(\lambda f)(x_0) = \lambda \partial f(x_0)$.*

Demonstração. Note que λf é um funcional lipschitziano de posto λK em uma vizinhança de x_0 . No caso em que $\lambda > 0$, da **Definição 14**, $(\lambda f)^\circ(x_0; \cdot) = \lambda f^\circ(x_0; \cdot)$, assim, como o gradiente generalizado é convexo e, fracamente^{*}, compacto, o resultado segue. Já no caso em que $\lambda = -1$, fixado $\zeta \in X^*$, é válido que

$$\begin{aligned} \zeta \in \partial(-f)(x_0) &\Leftrightarrow (-f)^\circ(x_0; v) \geq \langle v, \zeta \rangle \quad \forall v \in X \\ &\stackrel{\text{Propriedade 15}}{\Longleftrightarrow} f^\circ(x_0; -v) \geq \langle v, \zeta \rangle = \langle -v, -\zeta \rangle \quad \forall v \in X \\ &\Leftrightarrow -\zeta \in \partial f(x). \end{aligned}$$

$\square$

Definição 22 (página 199 de [11]). O funcional f é dito **regular em x_0** se admite derivadas direcionais usuais $f'(x_0; v)$ em relação a qualquer direção $v \in X$ e $f^\circ(x_0; \cdot) = f'(x_0; \cdot)$.

Teorema 23 (página 200 de [11]). Se f e g são lipschitzianos em uma vizinhança de x_0 , então,

$$\partial(f+g)(x_0) \subseteq \partial f(x_0) + \partial g(x_0).$$

Ainda, se f e g são regulares em x_0 , então, vale a igualdade.

Demonstração. O funcional de suporte do lado esquerdo da desigualdade é $(f+g)^\circ(x_0; \cdot)$, e o do lado direito é $f^\circ(x_0; \cdot) + g^\circ(x_0; \cdot)$, logo, a **Propriedade 16** garante a inclusão desejada. Ainda, se f e g são regulares em x_0 , então, qualquer $v \in X$ satisfaz

$$f^\circ(x_0; v) + g^\circ(x_0; v) = f'(x_0; v) + g'(x_0; v) = (f+g)'(x_0; v) \leq (f+g)^\circ(x_0; v).$$

Isto, em conjunto com a **Propriedade 16**, garante que $(f+g)^\circ(x_0; \cdot) = f^\circ(x_0; \cdot) + g^\circ(x_0; \cdot)$. $\square$

Propriedade 24 (página 201 de [11]). A combinação linear positiva de funcionais regulares em x_0 resulta em um funcional regular em x_0 .

Demonstração. Supondo que $f_1, \dots, f_n$ sejam funcionais regulares em x_0 e sendo $\lambda_1, \dots, \lambda_n \geq 0$, todo $v \in X$ satisfaz

$$\left(\sum_{i=1}^n \lambda_i f_i \right)'(x_0; v) = \sum_{i=1}^n \lambda_i f'_i(x_0; v) = \sum_{i=1}^n \lambda_i f^\circ_i(x_0; v) = \left(\sum_{i=1}^n \lambda_i f_i \right)^\circ(x_0; v).$$

$\square$

Proposição 25 (página 201 de [11]). Se $f_1, \dots, f_n$ são lipschitzianos em uma vizinhança de x_0 e $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, então, $f := \sum_{i=1}^n \lambda_i f_i$ é, também, lipschitziano em uma vizinhança de x_0 e

$$\partial f(x_0) \subseteq \sum_{i=1}^n \lambda_i \partial f_i(x_0).$$

Ainda, a igualdade vale se cada f_i é regular e cada λ_i é não-negativo.

Lema 26 (página 201 de [11]). Dados $x, y \in X$, seja $f: X \rightarrow \mathbb{R}$ um funcional lipschitziano de posto K em uma vizinhança (tubular) do segmento $[x, y]$ — parametrizado por $\gamma(t) = (1-t)x + ty$ com $t \in [0, 1]$. Nessas condições, a função $g := f \circ \gamma$ é lipschitziana em $]0, 1[$, e fixado $t \in]0, 1[$, qualquer $\eta \in \partial g(t)$ admite um $\zeta \in \partial f(\gamma(t))$ satisfazendo $\eta = \langle y - x, \zeta \rangle$.

Demonstração. A continuidade lipschitziana de g é clara. Dito isso, para mostrar o resultado, é suficiente mostrar que, para $v \in \{-1, 1\}$, vale $\max \partial g(t)v \leq \max \{ \langle y - x, \zeta \rangle v \mid \zeta \in \partial f(\gamma(t)) \}$. O lado esquerdo dessa desigualdade coincide com $g^\circ(t; v)$, que satisfaz

$$\begin{aligned}
g^\circ(t; v) &= \limsup_{s \rightarrow t, \lambda \downarrow 0} \frac{g(s + \lambda v) - g(s)}{\lambda} \\
&= \limsup_{s \rightarrow t, \lambda \downarrow 0} \frac{f((1 - s - \lambda v)x + (s + \lambda v)y) - f((1 - s)x + sy)}{\lambda} \\
&\leq \limsup_{s \rightarrow t, \lambda \downarrow 0} \frac{f(\gamma(s) + \lambda v(y - x)) - f(\gamma(s))}{\lambda} \leq f^\circ(\gamma(t); v(y - x)) \\
&= \max\{\langle v(y - x), \zeta \rangle \mid \zeta \in \partial f(\gamma(t))\}.
\end{aligned}$$

□

Teorema 27 (página 201 de [11]). *Nas condições do **Lema 26**, existem $t_0 \in]0, 1[$ e $\zeta \in \partial f(\gamma(t_0))$ de modo que $f(y) - f(x) = \langle y - x, \zeta \rangle$.*

Demonstração. Definindo $\theta(t) = (f \circ \gamma)(t) + t[f(x) - f(y)]$, trata-se de uma função contínua cujo valor em 0 e em 1 é $f(x)$, o que mostra que ela admite um minimizador (ou maximizador) $t_0 \in]0, 1[$, assim, da **Proposição 19**, $0 \in \partial \theta(t_0)$. Usando a **Propriedade 21**, a **Proposição 25** e o **Lema 26**, o resultado segue. □

Corolário 28. *Nas condições do **Lema 26**, existe $t_0 \in]0, 1[$ tal que $f(y) - f(x) \leq f^\circ(\gamma(t_0); y - x)$.*

Definição 29 (página 30 de [13]). Sendo Y um espaço de Banach real e $x_0 \in X$, uma aplicação $F : X \rightarrow Y$ é dita ser, **estritamente, diferenciável em x_0** se existe uma transformação linear $T : X \rightarrow Y$ satisfazendo

$$\lim_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \frac{F(x + tv) - F(x)}{t} = T(v) \quad \forall v \in X.$$

Nesse caso, T é denominada **diferencial estrita de F em x_0** e denotada por $F'_s(x_0)$.

Teorema 30 (página 202 de [11]). *Suponha que Y seja um espaço de Banach real, que $F : X \rightarrow Y$ seja, estritamente, diferenciável em x_0 e que $g : Y \rightarrow \mathbb{R}$ seja lipschitziano em uma vizinhança de $h(x_0)$. Nessas condições, $f := g \circ F$ é lipschitziano em uma vizinhança de x_0 , e vale*

$$\partial f(x_0) \subseteq [F'_s(x_0)]^* (\partial g(F(x_0))),$$

em que T^* denota a transformação adjunta de T . Se $F'_s(x_0)$ for sobrejetiva, valerá a igualdade.

Demonstração. Fixado $v \in X$, tome sequências $x_i \rightarrow x_0$ e $t_i \downarrow 0$ satisfazendo

$$\lim_{i \uparrow \infty} \frac{f(x_i + t_i v) - f(x_i)}{t_i} = f^\circ(x_0; v).$$

Do **Teorema 27**, para cada $i \in \mathbb{N}$ alto o suficiente, existem $z_i \in [F(x_i), F(x_i + t_i v)]$ e $\zeta_i \in \partial g(z_i)$ de modo que

$$\frac{f(x_i + t_i v) - f(x_i)}{t_i} = \left\langle \frac{F(x_i + t_i v) - F(x_i)}{t_i}, \zeta_i \right\rangle.$$

Sabendo que, da **Proposição 18**, a sequência ζ_i é limitada, do **Corolário 178**, ela admite uma subsequência ζ_j que converge, fracamente^{*}, para um certo $\zeta \in \partial g(F(x_0))$ pela **Proposição 20**, logo,

$$f^\circ(x_0; v) = \lim_{j \rightarrow \infty} \frac{f(x_j + t_j v) - f(x_j)}{t_j} = \langle [F'_s(x_0)](v), \zeta \rangle = \langle v, [F'_s(x_0)]^*(\zeta) \rangle.$$

Agora, supondo que $F'_s(x_0)$ seja sobrejetora, F associa cada vizinhança de x_0 a uma vizinhança de $F(x_0)$, assim, como

$$\lim_{x \rightarrow x_0, t \downarrow 0} \frac{F(x + tv) - F(x) - t[F'_s(x_0)](v)}{t} = 0,$$

segue que

$$\begin{aligned} g^\circ(F(x_0); [F'_s(x_0)](v)) &= \limsup_{y \rightarrow F(x_0), t \downarrow 0} \frac{g(y + t[F'_s(x_0)](v)) - g(y)}{t} \\ &= \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{g(F(x) + t[F'_s(x_0)](v)) - g(F(x))}{t} \\ &= \limsup_{x \rightarrow x_0, t \downarrow 0} \frac{f(x + tv) - f(x)}{t} = f^\circ(x_0; v). \end{aligned}$$

Como v é arbitrário, a igualdade entre os dois conjuntos fica mostrada. $\square$

Teorema 31 (página 203 de [11]). *Suponha que $F : X \rightarrow \mathbb{R}^n$ seja lipschitziana de posto K_F em uma vizinhança de x_0 e que $g : \mathbb{R}^n \rightarrow \mathbb{R}$ seja, estritamente, diferenciável em $F(x_0)$. Nessas condições, $f := g \circ F$ é lipschitziana em uma vizinhança de x_0 e vale que*

$$\partial f(x_0) = \partial \langle F, g'_s(F(x_0)) \rangle(x_0).$$

Demonstração. Chamando $h_0 = g'_s(F(x_0))$ e $\eta = \langle F, h_0 \rangle$, provaremos que, para $\epsilon > 0$, existe uma vizinhança V_ϵ de x_0 na qual $f - \eta$ é lipschitziana de posto ϵ . De fato, fixados y, z na vizinhança de x_0 na qual f é lipschitziana, como, pelo Teorema do Valor Médio, existe $u \in [F(y), F(z)]$ tal que $f(y) - f(z) = \langle F(y) - F(z), g'_s(u) \rangle$, temos que

$$\begin{aligned} (f - \eta)(y) - (f - \eta)(z) &= \langle F(y) - F(z), g'_s(u) \rangle - \langle F(y) - F(z), h_0 \rangle \\ &= \langle F(y) - F(z), g'_s(u) - h_0 \rangle \leq \|g'_s(u) - h_0\|_{(\mathbb{R}^n)^*} \|F(y) - F(z)\|_{\mathbb{R}^n} \\ &\leq K_F \|g'_s(u) - h_0\|_{(\mathbb{R}^n)^*} \|y - z\|_X, \end{aligned}$$

logo, restringido y e z a uma bola fechada, suficientemente, pequena e centrada em x_0 , é possível fazer com que $K_F \|g'_s(u) - h_0\|_{(\mathbb{R}^n)^*} \leq \epsilon$.

A partir disso, da **Proposição 18** e do **Teorema 23**, segue que $\partial f(x_0) \subseteq \partial \eta(x_0) + B_{X^*}[0, \epsilon]$ e que $\partial \eta(x_0) \subseteq \partial f(x_0) + B_{X^*}[0, \epsilon]$, provando a igualdade. $\square$

Corolário 32 (página 204 de [11]). *Se f e g são lipschitzianas em uma vizinhança de x_0 , então, o produto fg , também é lipschitziano em vizinhança de x_0 e vale que*

$$\partial(fg)(x_0) \subseteq f(x_0)\partial g(x_0) + g(x_0)\partial f(x_0).$$

Ainda, a igualdade vale se f e g forem regulares em x_0 e o produto delas for não-negativo em tal ponto.

Demonstração. Chamando G a função produto das coordenadas e $F = (f, g)$, é claro que $fg = G \circ F$, que a primeira é, estritamente, diferenciável e que a segunda é lipschitziana em uma vizinhança de x_0 — o posto dela coincide com a raiz quadrada da somatória dos quadrados dos postos de f e g . Sendo assim, do **Teorema 31** e da **Propriedade 16**, vem

$$\begin{aligned} \partial(fg)(x_0) &= \partial\langle(f, g); (g(x_0), f(x_0))\rangle(x_0) \\ &= \partial[g(x_0)f + f(x_0)g](x_0) \subseteq f(x_0)\partial g(x_0) + g(x_0)\partial f(x_0). \end{aligned}$$

Se f e g forem regulares em x_0 e cada uma delas for não-negativa em tal ponto, a inclusão torna-se igualdade pela **Propriedade 16**, pois $g(x_0)f$ e $f(x_0)g$ são, nesse caso, regulares em x_0 . Já se ambas forem não-positivas em x_0 , suas opostas são regulares e não-negativas em x_0 , logo, basta aplicar as Propriedades 16 e 21 para obter a igualdade desejada. $\square$

2.3 Cones e polaridade

Os resultados dessa seção serão feitos sob a suposição de que X é um espaço de Banach real.

Definição 33 (página 30 de [21]). • Um conjunto $K \subseteq X$ é denominado **cone com vértice na origem** se $\lambda K = K$ para todo $\lambda > 0$;

- Sendo K um cone com vértice na origem, o conjunto $x_0 + K$ é denominado **cone com vértice em x_0** ;
- Um cone é denominado **afiado** se contém seu vértice e **cego** caso contrário, e é denominado **saliente** se não contém subespaços de X não-triviais.

Lema 34 (página 30 de [21]). Supondo que $K \subseteq X$ seja um cone com vértice em x_0 e que $f \in X^*$ seja tal que $\langle K, f \rangle$ admite uma cota superior $\alpha \in \mathbb{R}$, o valor $\langle x_0, f \rangle$ é, também, uma cota superior de $\langle K, f \rangle$.

Demonstração. Sendo $y \in K$, por definição, a semirreta $\{(1-t)x_0 + ty \mid t \geq 0\} \subseteq K$, assim, como f é linear e α é cota superior de $\langle K, f \rangle$, vale $(1-t)\langle x_0, f \rangle + t\langle y, f \rangle \leq \alpha$ para todo $t \geq 0$. Em particular, para todo $n \in \mathbb{N}$, sabendo que

$$(1-n)\langle x_0, f \rangle + n\langle y, f \rangle \leq \alpha \Leftrightarrow n\langle y, f \rangle \leq \alpha + (n-1)\langle x_0, f \rangle \Leftrightarrow \langle y, f \rangle \leq \frac{\alpha}{n} + \left(1 - \frac{1}{n}\right)\langle x_0, f \rangle$$

e fazendo $n \uparrow \infty$, obtemos $\langle y, f \rangle \leq \langle x_0, f \rangle$. Da arbitrariedade de y , o resultado segue. $\square$

Definição 35 (adaptado de [47]). Dois cones convexos $C_1, C_2 \subseteq X$ são ditos **transversais** se $C_1 - C_2 = X$ (diferença de Kuratowski) e, **fortemente, transversais** se, além de transversais, possuem interseção não-trivial.

Proposição 36 (página 239 de [43]). *Se dois cones convexos não-vazios $C_1, C_2 \subseteq X$ são inseparáveis, então, eles são transversais. Também, se nenhum deles é um subespaço vetorial, então, eles são, fortemente, transversais.*

Demonstração. Sabendo que $C_1 - C_2$ é um cone convexo, se tal cone não coincide com X , então, existe um $\lambda \in X^* \setminus \{0\}$ para o qual $\langle C_1 - C_2, \lambda \rangle \subseteq [0, \infty[$. Em particular, para $x \in C_1$, $r > 0$ e $y \in C_2$, vale

$$\begin{aligned} \langle x, \lambda \rangle &\geq r \langle y, \lambda \rangle & \forall x \in C_1, \quad y \in C_2, \quad r > 0, \\ r \langle x, \lambda \rangle &\geq \langle y, \lambda \rangle & \forall x \in C_1, \quad y \in C_2, \quad r > 0. \end{aligned}$$

Fazendo $r \downarrow 0$, é obtido $\langle C_1, \lambda \rangle \subseteq [0, \infty[$ e $\langle C_2, \lambda \rangle \subseteq]-\infty, 0]$, o que prova que C_1 e C_2 são separáveis. Isso contradiz a hipótese de inseparabilidade, portanto, C_1 e C_2 são transversais.

Agora, se $C_1 \cap C_2 = \{0\}$, dado $z \in C_2$, existem $x \in C_1$ e $y \in C_2$ satisfazendo $z = x - y$, então, $z + y = x \in C_1 \cap C_2 = \{0\}$, o que mostra que $z = -y$ e, conseqüentemente, $-z \in C_2$. Sabendo que isso vale para qualquer $z \in C_2$, a conclusão é que se trata de um subespaço, mas isso contradiz a hipótese de que nenhum deles é subespaço. Portanto, eles são, fortemente, transversais. $\square$

Corolário 37 (página 31 de [21]). *Nas condições do Lema 34, é válido que:*

- Se um cone K de vértice x_0 é separável de um conjunto A por um hiperplano fechado, então, ele é separável, também, pelo hiperplano paralelo passando por seu vértice;
- Um cone convexo fechado é separável de qualquer semirreta emanando de seu vértice não contida nele;
- Se dois cones K_1 e K_2 com vértice na origem são separados por um hiperplano, então, este deve passar pela origem. Em particular, se K_2 é um subespaço, então, ele deve ser subconjunto do referido hiperplano.

Demonstração. • Supondo que o hiperplano de equação $f = \alpha$ separe A e K , α é uma cota inferior de $\langle A, f \rangle$ e, do Lema 34, $\langle x_0, f \rangle$ é cota superior de $\langle K, f \rangle$, o que mostra que este é, também, uma cota inferior de $\langle A, f \rangle$, portanto, $f = \langle x_0, f \rangle$ separa A de K ;

- Supondo que K seja um cone de vértice x_0 e x_1 seja um ponto de uma semirreta emanando do cone e não contida em K , segue que $x_1 \notin K$, assim, do Teorema 159, $\{x_1\}$ podem ser separados por um hiperplano de equação $f = \alpha$, o que acarreta que o hiperplano de equação $f = \langle x_0, f \rangle$ separa, também, $\{x_1\}$ e K . A partir disso, segue que $\langle (1-t)x_0 + tx_1, f \rangle = \langle x_0, f \rangle + t(\langle x_1, f \rangle - \langle x_0, f \rangle) \geq \langle x_0, f \rangle$ para qualquer $t \geq 0$, o que mostra que $f = \langle x_0, f \rangle$ separa o cone e a semirreta;

- Supondo que a equação de tal hiperplano seja $f = \alpha$ e tomando $x_1 \in K_1$ e $x_2 \in K_2$, para cada $n \in \mathbb{N}$, vale $\langle \frac{x_1}{n}, f \rangle \geq \alpha \geq \langle \frac{x_2}{n}, f \rangle$, assim, fazendo $n \rightarrow \infty$, a continuidade de f garante que $\langle 0, f \rangle \geq \alpha \geq \langle 0, f \rangle$, o que mostra que $\alpha = \langle 0, f \rangle = 0$. Portanto, o hiperplano deve passar pela origem.

Agora, se K_2 é subespaço, então, dado $x \in K_2$, vale $\langle x, f \rangle, \langle -x, f \rangle \leq 0$, o que mostra que $\langle x, f \rangle = 0$. Portanto, K_2 está contido no hiperplano de equação $f = 0$.

□

Teorema 38 (página 32 de [21]). *Sendo K um cone convexo com vértice na origem e interior não-vazio e L um subespaço que intercepte $\text{int } K$, se $g \in X^*$ é não-positivo em $K \cap L$, então, existe $f \in X^*$ contínuo que coincide com g em L e é não-positivo em K .*

Demonstração. No caso em que $g \equiv 0$ em L , o funcional $f \equiv 0$ satisfaz as condições requeridas. Supondo, assim, $g \not\equiv 0$ em L e chamando $Q_1 = \ker g$, trata-se de um cone convexo e não-vazio (contém a origem) que é disjunto de $\text{int } K$: de fato, se existisse $x_0 \in Q_1 \cap \text{int } K$, existiria, pela continuidade e linearidade de g , um ponto $x_1 \in L \cap K$ satisfazendo $g(x_1) > 0$, o que contradiz a hipótese. Assim, do Teorema 159, existe um hiperplano separando Q_1 e $\text{int } K$, o que implica, pelo item 3 do Corolário 37, na existência de um funcional linear $h: X \rightarrow \mathbb{R}$ contínuo negativo em $\text{int } K$ e nulo em Q_1 .

Agora, chamando $Q_2 = (\ker h) \cap L$, como h se anula em Q_1 , vale $Q_1 \subseteq Q_2$, assim, do Lema 172, $h|_L \equiv 0$ ou $h|_L = \lambda g$ para algum $\lambda \neq 0$. O primeiro caso é impossível pelo fato de que h é negativo em $\text{int } K \cap L$, que, por hipótese, é não-vazio, assim, o segundo caso é o que ocorre. Mais ainda, como g e h possuem o mesmo sinal em $K \cap L$, o número λ deve ser positivo, portanto, pondo $f = \lambda g$, o resultado segue. □

Definição 39 (página 72 de [11]). *Sendo K um cone com vértice na origem, o cone K^Δ definido a seguir é denominado **cone polar de K** :*

- Se $K \subseteq X$, K^Δ é formado pelos funcionais $f \in X^*$ para os quais 0 é cota superior de $\langle K, f \rangle$;
- Se $K \subseteq X^*$, K^Δ é formado pela interseção das imagens inversas de $] -\infty, 0]$ por todos os funcionais de K .

Da definição, é possível afirmar que K^Δ é um cone convexo afiado com vértice na origem.

Propriedade 40 (página 33 de [21]). *O cone K^Δ é fechado na topologia mais fraca de seu espaço.*

Demonstração. • $K \subseteq X$: tome $f_i \xrightarrow{*} f_0$. Dado $x \in K$, vale $\langle x, f_i \rangle \leq 0$ para todo $i \in \mathbb{N}$, assim, $\langle x, f_0 \rangle = \lim \langle x, f_i \rangle \leq 0$. Isso mostra que $f_0 \in K^\Delta$;

- $K \subseteq X^*$: tome $x_i \rightarrow x_0$. Dado $f \in K$, vale $\langle x_i, f \rangle \leq 0$ para todo $i \in \mathbb{N}$, assim, $\langle x_0, f \rangle = \lim \langle x_i, f \rangle \leq 0$. Isso mostra que $x_0 \in K^\Delta$.

□

Propriedade 41 (página 34 de [21]). Sendo $K_1, K_2 \subseteq X$ cones com vértice na origem, se $K_1 \subseteq K_2 \subseteq X$, então, $K_1^\Delta \supseteq K_2^\Delta$.

Propriedade 42 (página 33 de [21]). O cone polar de K coincide com o de $\bar{K}$ e o de $\text{co } K$.

Demonstração. • $K \subseteq X$: tomando f um funcional linear contínuo e não-positivo em K , seja $x \in \bar{K}$. Se $x \in K$, é imediato que $\langle x, f \rangle \leq 0$. Senão, x é ponto de acumulação de K , desse modo, caso $\langle x, f \rangle > 0$, $f^{-1}(]0, \infty[)$ seria uma vizinhança de x e, conseqüentemente, conteria pontos de K , o que implica que $\langle x, f \rangle \leq 0$. Portanto, $K^\Delta \subseteq \bar{K}^\Delta$. Já a outra inclusão segue da **Propriedade 41**. Portanto, os cones polares de K e $\bar{K}$ coincidem.

Agora, tomando f um funcional linear contínuo e não-positivo em K , como cada ponto de $\text{co } K$ é combinação convexa de pontos de K , a linearidade de f garante que f é não-positivo em $\text{co } K$. Já a outra inclusão segue da **Propriedade 41**. Portanto, os cones polares de K e $\text{co } K$ coincidem;

- $K \subseteq X^*$: tomando x um ponto no qual todo elemento de K assume valor não-positivo, seja $f \in \bar{K}$. Se $f \in K$, é imediato que $\langle x, f \rangle \leq 0$. Senão, f é ponto de acumulação de K , desse modo, caso $\langle x, f \rangle > 0$, a imagem inversa de $]0, \infty[$ por $\langle x, \cdot \rangle$ seria uma vizinhança de f e, conseqüentemente, conteria pontos de K , o que implica que $\langle x, f \rangle \leq 0$. Portanto, $K^\Delta \subseteq \bar{K}^\Delta$. Já a outra inclusão segue da **Propriedade 41**. Portanto, os cones polares de K e $\bar{K}$ coincidem.

Agora, tomando x um ponto no qual todo elemento de K assume valor não-positivo, como cada elemento de $\text{co } K$ é combinação convexa de elementos de K , cada elemento daquele assume valores não-positivos em x . Já a outra inclusão segue da **Propriedade 41**. Portanto, os cones polares de K e $\text{co } K$ coincidem;

□

Propriedade 43. Sendo K_1, K_2 cones com vértice na origem, $(K_1 \times K_2)^\Delta = K_1^\Delta \times K_2^\Delta$.

Demonstração. • $K_1, K_2 \subseteq X$: supondo que $f = (f_1, f_2) \in (K_1 \times K_2)^\Delta$, para quaisquer $x_1 \in K_1$ e $x_2 \in K_2$, vale $\langle x_1, f_1 \rangle = \langle (x_1, 0); f \rangle \leq 0$ e $\langle x_2, f_2 \rangle = \langle (0, x_2); f \rangle \leq 0$, assim, $f_1 \in K_1^\Delta$ e $f_2 \in K_2^\Delta$, ou seja, $f \in K_1^\Delta \times K_2^\Delta$. Reciprocamente, supondo que $f = (f_1, f_2) \in K_1^\Delta \times K_2^\Delta$, para qualquer $x = (x_1, x_2) \in K_1 \times K_2$, vale que $\langle x, f \rangle = \langle x_1, f_1 \rangle + \langle x_2, f_2 \rangle \leq 0$, o que mostra que $f \in (K_1 \times K_2)^\Delta$;

- $K_1, K_2 \subseteq X^*$: supondo que $x = (x_1, x_2) \in (K_1 \times K_2)^\Delta$, para quaisquer $f_1 \in K_1$ e $f_2 \in K_2$, vale $\langle x_1, f_1 \rangle = \langle x; (f_1, 0) \rangle \leq 0$ e $\langle x_2, f_2 \rangle = \langle x; (0, f_2) \rangle \leq 0$, assim, $x_1 \in K_1^\Delta$ e $x_2 \in K_2^\Delta$, ou seja, $x \in K_1^\Delta \times K_2^\Delta$. Reciprocamente, supondo que $x = (x_1, x_2) \in K_1^\Delta \times K_2^\Delta$, para qualquer $f = (f_1, f_2) \in K_1 \times K_2$, vale que $\langle x, f \rangle = \langle x_1, f_1 \rangle + \langle x_2, f_2 \rangle \leq 0$, o que mostra que $f \in (K_1 \times K_2)^\Delta$.

□

Propriedade 44 (página 34 de [21]). Sendo $\{K_\alpha \mid \alpha \in A\}$ uma coleção de cones com vértice na origem,

$$\left(\bigcup_{\alpha \in A} K_\alpha \right)^\Delta = \bigcap_{\alpha \in A} K_\alpha^\Delta.$$

Demonstração. • $\{K_\alpha \mid \alpha \in A\} \subseteq X$: sendo $f \in \left(\bigcup_{\alpha \in A} K_\alpha \right)^\Delta$, em particular, dados $\alpha \in A$ e $x \in K_\alpha$, vale $\langle x, f \rangle \leq 0$, assim, $f \in \bigcap_{\alpha \in A} K_\alpha^\Delta$. Reciprocamente, se $f \in \bigcap_{\alpha \in A} K_\alpha^\Delta$, então, dados $\alpha \in A$ e $x \in K_\alpha$, vale $\langle x, f \rangle \leq 0$, o que mostra que $f \in \left(\bigcup_{\alpha \in A} K_\alpha \right)^\Delta$;

• $\{K_\alpha \mid \alpha \in A\} \subseteq X^*$: sendo $x \in \left(\bigcup_{\alpha \in A} K_\alpha \right)^\Delta$, em particular, dados $\alpha \in A$ e $f \in K_\alpha$, vale $\langle x, f \rangle \leq 0$, assim, $x \in \bigcap_{\alpha \in A} K_\alpha^\Delta$. Reciprocamente, se $x \in \bigcap_{\alpha \in A} K_\alpha^\Delta$, então, dados $\alpha \in A$ e $f \in K_\alpha$, vale $\langle x, f \rangle \leq 0$, o que mostra que $x \in \left(\bigcup_{\alpha \in A} K_\alpha \right)^\Delta$.

□

Propriedade 45 (página 34 de [11]). *O cone polar de K^Δ coincide com a envoltória convexa fechada de K .*

Demonstração. Da Propriedade 42, o cone polar de $\overline{\text{co}}K$ coincide com o de K , assim:

- Dado $x \in \overline{\text{co}}K$, qualquer $f \in K^\Delta$ satisfaz $\langle x, f \rangle \leq 0$, o que mostra que $x \in K^{\Delta\Delta}$. Por outro lado, se $x \notin \overline{\text{co}}K$, do **Teorema 159**, existe $f \in X^*$ que é positivo em x e negativo em $\overline{\text{co}}K$, o que, da **Propriedade 42**, acarreta $f \in K^\Delta$, assim, $x \notin K^{\Delta\Delta}$;
- Dado $f \in \overline{\text{co}}K$, qualquer $x \in K^\Delta$ satisfaz $\langle x, f \rangle \leq 0$, o que mostra que $f \in K^{\Delta\Delta}$. Por outro lado, se $f \notin \overline{\text{co}}K$, do **Teorema 159**, existe $x \in X$ no qual f assume valor positivo e no qual qualquer $\overline{\text{co}}K$ assume valor negativo, o que, da **Propriedade 42**, acarreta $x \in K^\Delta$, assim, $f \notin K^{\Delta\Delta}$.

□

Propriedade 46 (página 34 de [21]). *Se $\{K_\alpha \mid \alpha \in A\}$ é uma coleção de cones convexos e fechados na topologia mais fraca de seu espaço, então,*

$$\left(\bigcap_{\alpha \in A} K_\alpha \right)^\Delta = \overline{\sum_{\alpha \in A} K_\alpha^\Delta}.$$

Demonstração. Chamando de Q o fecho da referida somatória, como

$$\text{co} \left(\bigcup_{\alpha \in A} K_\alpha^\Delta \right) = \sum_{\alpha \in A} K_\alpha^\Delta,$$

das Propriedades 42, 44 e 45, segue que

$$Q^\Delta = \left(\sum_{\alpha \in A} K_\alpha^\Delta \right)^\Delta = \left(\bigcup_{\alpha \in A} K_\alpha^\Delta \right)^\Delta = \bigcap_{\alpha \in A} K_\alpha^{\Delta\Delta} = \bigcap_{\alpha \in A} K_\alpha,$$

o que permite afirmar, usando a **Propriedade 45** novamente, que

$$Q = Q^{\Delta\Delta} = \left(\bigcap_{\alpha \in A} K_\alpha \right)^\Delta.$$

□

Propriedade 47 (página 9 de [7]). *Se $K_1, \dots, K_m \subseteq X$ são cones convexos e inseparáveis, então, seus interiores relativos não são disjuntos.*

Demonstração. Supondo que exista um inteiro $\ell < m$ para o qual

$$\bigcap_{i=1}^{\ell} \text{ri } K_i \neq \emptyset, \quad \bigcap_{i=1}^{\ell+1} \text{ri } K_i = \emptyset,$$

do **Teorema 195**, o conjunto $\bigcap_{i=1}^{\ell} K_i$ possui interior relativo não-vazio e os interiores deste e de $K_{\ell+1}$ são disjuntos, assim, o **Teorema 159** garante que $K_{\ell+1}$ é separável de $\bigcap_{i=1}^{\ell} K_i$ e, consequentemente, de $\bigcap_{i \in \{1 : m\} \setminus \{\ell+1\}} K_i$, o que contradiz a hipótese de inseparabilidade. □

Teorema 48 (página 7 de [7]). *Dados $K_1, \dots, K_m \subseteq X^*$ cones, fracamente*, fechados, se $K := \sum_{i=1}^m K_i$ não é, fracamente*, fechado, então, existem $y_1, \dots, y_m$ não todos nulos cuja somatória é nula e de modo que cada $y_i \in K_i$.*

Demonstração. Tomando $x \in \bar{K} \setminus K$, existe uma sequência $\{x_j\} \subseteq K$ tendendo a x , e podemos escrever

$$x_j = \sum_{i=1}^m x_{i,j}, \quad x_{i,j} \in K_i \quad \forall i \in \{1 : m\}.$$

Pondo, agora, $a_j = \max\{|x_{1,j}|, \dots, |x_{m,j}|\}$ para todo $j \in \mathbb{N}$, se $a_j \nrightarrow \infty$, ela seria uma sequência limitada, o que permitiria encontrar, para cada $i \in \{1 : m\}$ uma subsequência de $\{x_{i,j}\}$ convergente (**Corolário 178**) e, consequentemente, concluir que $x \in K$. Agora, pondo $y_{i,j} = \frac{x_{i,j}}{a_j}$ para todos $i \in \{1 : m\}$ e $j \in \mathbb{N}$, vemos que $\max\{|y_{1,j}|, \dots, |y_{m,j}|\} = 1$ para todo $j \in \mathbb{N}$, o que permite assumir, sem perda de generalidade, que $y_{i,j} \rightarrow y_i$ para todo $i \in \{1 : m\}$ (novamente, **Corolário 178**), e, consequentemente, que, ao menos, um dos y é não-nulo e que cada $y_i \in K_i$. Finalmente,

$$\sum_{i=1}^m y_i = \lim_{j \uparrow \infty} \sum_{i=1}^m y_{i,j} = \lim_{j \uparrow \infty} \frac{1}{a_j} \sum_{i=1}^m x_{i,j} = \lim_{j \uparrow \infty} \frac{x_j}{a_j} = 0.$$

□

Teorema 49 (página 6 de [7]). *Dados $K_1, \dots, K_m \subseteq X$ cones convexos com vértice na origem, eles são separáveis se, e somente se, existem $y_1, \dots, y_m$ não todos nulos cuja somatória é nula e de modo que cada $y_i \in K_i^\Delta$.*

Demonstração. Seja $\ell \in \{1 : m-1\}$ o menor inteiro para o qual $\{K_1, \dots, K_m\}$ contém uma subfamília separável de $\ell + 1$ cones e assumamos que K_1 é separável de $\bigcap_{i=2}^{\ell+1} K_i$ — não há perda de

generalidade nessa suposição, uma vez que os cones em questão podem ser reordenados se necessário. A partir disso, é possível encontrar $f \in X^* \setminus \{0\}$ de modo que $f \in K_1^\Delta$ e que $-f \in \left(\bigcap_{i=2}^{\ell+1} K_i\right)^\Delta$. Se $\ell = 1$, o teorema está provado. Senão, como $\{K_2, \dots, K_{\ell+1}\}$ não é uma família de cones separáveis por conter menos que $\ell + 1$ cones, a **Propriedade 47** garante que $\bigcap_{i=2}^{\ell+1} \text{ri } K_i \neq \emptyset$, assim, do **Teorema 195**, $\overline{\bigcap_{i=2}^{\ell+1} K_i} = \bigcap_{i=2}^{\ell+1} \overline{K_i}$. Disso e das Propriedades 42 e 46, temos que

$$-f \in \left(\bigcap_{i=2}^{\ell+1} K_i\right)^\Delta = \overline{\sum_{i=2}^{\ell+1} K_i^\Delta}.$$

Se a somatória em questão for, fracamente^{*}, fechada, claramente, existirão $y_2, \dots, y_{\ell+1}$ cuja somatória é $-f$ e de modo que cada $y_i \in K_i^\Delta$, assim, basta por $y_0 = f$ e $y_{\ell+2} = \dots = y_m = 0$ e o resultado seguirá. Senão, do **Teorema 48**, existirão $y_2, \dots, y_{\ell+1}$ não todos nulos cuja somatória é nula e de modo que cada $y_i \in K_i^\Delta$, assim, basta por $y_0 = y_{\ell+2} = \dots = y_m = 0$ e o resultado seguirá.

Reciprocamente, suponha que existam $y_1, \dots, y_m$ não todos nulos, de somatória nula e de modo que cada $y_i \in K_i^\Delta$. Sendo $j \in \{1 : m\}$ o índice do primeiro y não-nulo, temos que $y_j \in K_j^\Delta$ e que

$$\langle x, y_j \rangle = - \sum_{i \in \{1 : m\} \setminus \{j\}} \langle x, y_i \rangle,$$

assim, para $x \in \bigcap_{i \in \{1 : m\} \setminus \{j\}} K_i$, vale $\langle x, y_j \rangle \geq 0$, o que mostra que o hiperplano $y_j^{-1}(0)$ separa K_j de $\bigcap_{i \in \{1 : m\} \setminus \{j\}} K_i$. $\square$

2.3.1 Cones aproximantes de Boltjanskij

Estudaremos, agora, os *cones aproximantes de Boltjanskij*, cones que podem ser assumidos convexos e que satisfazem uma propriedade denominada, por [50], de Princípio da Interseção.

Definição 50 (página 235 de [43]). Sejam $A \subseteq \mathbb{R}^n$ e $q \in \bar{A}$. Um cone $K \subseteq \mathbb{R}^n$ é dito **cone aproximante de Boltjanskij para A em q** se, para quaisquer $\{v_1, \dots, v_k\} \subseteq K$ (com $k \in \mathbb{N}$) e $\epsilon > 0$, existem $\{v'_1, \dots, v'_k\} \subseteq \mathbb{R}^n$ satisfazendo $\|v_j - v'_j\| < \epsilon$ para $j \in \{1 : k\}$, $\delta > 0$ e uma aplicação contínua $\Xi : [0, \delta]^k \rightarrow A \cup \{q\}$ satisfazendo

$$\Xi(z) = q + \sum_{j=1}^k z_j v'_j + o(\|z\|). \quad (2.1)$$

Propriedade 51 (página 237 de [43]). Se K é um cone aproximante de Boltjanskij para A em q , então, o mesmo vale para sua envoltória convexa $\text{co } K$.

Demonstração. Comece tomando $\{v_1, \dots, v_k\} \subseteq \text{co } K$ (com $k \in \mathbb{N}$) e $\epsilon > 0$. Do Teorema de Karatheodori para Envoltórias Convexas, para cada $j \in \{1 : k\}$, existem $\alpha_{1,j}, \dots, \alpha_{n+1,j} \geq 0$ com somatória $\frac{1}{k}$ e $\{v_{1,j}, \dots, v_{n+1,j}\} \subseteq K$ de modo que

$$v_j = k \sum_{i=1}^{n+1} \alpha_{i,j} v_{i,j}.$$

Agora, sabendo que K é cone aproximante de Boltjanskij para A em q , para cada conjunto $\{v_{i,1}, \dots, v_{i,k}\}$, existem $\{v'_{i,1}, \dots, v'_{i,k}\}$ satisfazendo $\|v_{i,j} - v'_{i,j}\| < \epsilon$ para $j \in \{1 : k\}$, $\delta_i > 0$ e uma aplicação contínua $\Xi_i : [0, \delta_i]^k \rightarrow A \cup \{q\}$ satisfazendo

$$\Xi_i(z) = q + \sum_{j=1}^k z_j v'_{i,j} + o(\|z\|). \quad (2.2)$$

Definindo $v'_j = k \sum_{i=1}^{n+1} \alpha_{i,j} v'_{i,j}$, temos que $\|v_j - v'_j\| < \epsilon$ para $j \in \{1 : k\}$. Ainda, pondo $\alpha_i = \sum_{j=1}^k \alpha_{i,j}$ e $\delta = \min\{\delta_1, \dots, \delta_{n+1}\}$, a aplicação $\Xi : [0, \delta]^k \rightarrow A \cup \{q\}$ definida por

$$\Xi(z) = \sum_{i=1}^{n+1} \alpha_i \Xi_i$$

é tal que

$$\Xi(z) = q + \sum_{j=1}^n z_j \left(\sum_{i=1}^{n+1} \alpha_{i,j} v'_{i,j} \right) + o(\|z\|) = q + \sum_{j=1}^n z_j v'_j + o(\|z\|),$$

o que mostra que K é um cone aproximante de Boltjanskij para A em q . □

Teorema 52 (página 237 de [43]). *Se $\mathbb{R}^n$ é um cone aproximante de Boltjanskij para A em q , então, este é um ponto interior de $A \cup \{q\}$.*

Demonstração. A demonstração desse teorema será feita para $q = 0$, porém, quando $q \neq 0$, uma simples translação da aplicação Ξ estende tal resultado para todo o $\mathbb{R}^n$.

Colocando $e_0 = -\sum_{i=1}^n e_i$, o conjunto $\{e_0, \dots, e_n\}$ é, afinamente, independente (isto é, nenhum ponto pode ser escrito como combinação convexa dos demais), sendo assim, denotando sua envoltória convexa por S , todo $s \in S$ pode ser escrito, unicamente, como uma combinação convexa dos e_i com coeficientes z_i . Tais coeficientes serão colocados em um vetor $z \in \mathbb{R}^{n+1}$.

Por hipótese, dado $\epsilon > 0$, existem $v_0, \dots, v_n \in \mathbb{R}^n$ satisfazendo $\|e_i - v_i\| < \epsilon$ para $i \in \{0 : n\}$, um número $\delta > 0$ e uma aplicação contínua $\Xi : [0, \delta]^{n+1} \rightarrow A \cup \{0\}$ satisfazendo

$$\Xi(z) = \sum_{i=0}^n z_i v_i + o(\|z\|).$$

Escolha $\alpha > 0$ de modo que $B_{\mathbb{R}^n}[0, \alpha] \subseteq S$ e, para $r > 0$ pequeno o suficiente, defina a aplicação $\Phi_r : S \rightarrow \mathbb{R}^n$ por $\Phi_r(s) = \frac{1}{r} \Xi(rz)$. Para ϵ e r pequeno o suficiente, qualquer $s \in S$ satisfaz

$$\|\Phi_r(s) - s\| = \left\| \sum_{i=0}^n z_i (v_i - e_i) + \frac{o(r)}{r} \right\| \leq \sum_{i=0}^n z_i \|v_i - e_i\| + \frac{o(r)}{r} \leq \epsilon + \frac{o(r)}{r} \leq \frac{\alpha}{2}.$$

Fixando $q \in B_{\mathbb{R}^n}\left[0, \frac{\alpha}{2}\right]$ e definindo a aplicação contínua $F : S \times B_{\mathbb{R}^n}\left[0, \frac{\alpha}{2}\right] \rightarrow B_{\mathbb{R}^n}[0, \alpha]$ por $F(s, q) = q - \Phi_r(s) + s$, da estimativa acima, $F(\cdot, q)$ pode ser vista como uma aplicação sobre $B_{\mathbb{R}^n}[0, \alpha]$, o que, pelo Teorema do Ponto Fixo de Brouwer², garante que ela possui um ponto fixo $\mu \in B\left[0, \frac{\alpha}{2}\right]$, o que é equivalente a $\Phi_r(\mu) = q$. Isso permite dizer que $\Phi_r^{-1}\left(B_{\mathbb{R}^n}\left[0, \frac{\alpha}{2}\right]\right) \subseteq S$, assim, $B_{\mathbb{R}^n}\left(0, r\frac{\alpha}{2}\right) \subseteq A \cup \{0\}$. □

²Consulte a página 353 de [35]

Teorema 53 (página 240 de [43]). *Sendo $A_1, A_2 \subseteq \mathbb{R}^n$ e $q \in \overline{A_1} \cap \overline{A_2}$, se K_1 e K_2 são os respectivos cones aproximantes de Boltjanskij para A_1 e A_2 para q e são, fortemente, transversais, então, q é ponto limite de $A_1 \cap A_2$.*

Demonstração. Sem perda de generalidade, assumimos, novamente, que $q = 0$ e chamamos $e_0 = -\sum_{i=1}^n e_i$. Da transversalidade entre C_1 e C_2 , é possível encontrar, para cada $i \in \{0 : n\}$, $f_i \in C_1$ e $g_i \in C_2$ satisfazendo $e_i = f_i - g_i$.

Ainda, como existe $z \in C_1 \cap C_2 \setminus \{0\}$, é possível afirmar que, para cada $m > 0$, $e_i = (f_i + mz) - (g_i + mz)$ com $f_i + mz \in C_1$ e $g_i + mz \in C_2$.

A envoltória convexa de $\{f_0 + mz, \dots, f_n + mz\}$, sendo uma translação da de $\{f_0, \dots, f_n\}$, que é um conjunto compacto, por mz , pode ser escolhida de modo a não conter a origem. Sendo assim, não há perda de generalidade em assumir que envoltória convexa dos f_i não contém o zero.

Novamente, chamando S a envoltória convexa de $\{e_0, \dots, e_n\}$, todo $s \in S$ pode ser escrito, unicamente, como uma combinação convexa dos e_i com coeficientes z_i , que serão colocados em um vetor $z \in \mathbb{R}^{n+1}$.

Sabendo que C_1 e C_2 são cones aproximantes para A_1 e A_2 na origem, dado $\epsilon > 0$, existem vetores $f'_0, \dots, f'_n$ e $g'_0, \dots, g'_n$ satisfazendo $\|f_i - f'_i\| < \frac{\epsilon}{2}$ e $\|g_i - g'_i\| < \frac{\epsilon}{2}$ para $i \in \{0 : n\}$, $\delta > 0$ e aplicações contínuas

$$\begin{aligned} \Xi_1 : [0, \delta]^{n+1} &\rightarrow A_1 \cup \{0\} \\ z &\mapsto \sum_{i=0}^n z_i f'_i + o(\|z\|), \end{aligned}$$

$$\begin{aligned} \Xi_2 : [0, \delta]^{n+1} &\rightarrow A_2 \cup \{0\} \\ z &\mapsto \sum_{i=0}^n z_i g'_i + o(\|z\|). \end{aligned}$$

Sendo $B[0, \alpha] \subseteq \text{int } S$, para $r > 0$ pequeno, defina $\Phi_r : S \rightarrow \mathbb{R}^n$ por

$$\Phi_r(s) = \frac{\Xi_1(rz) - \Xi_2(rz)}{r},$$

para r e ϵ pequenos o suficiente, vale

$$\begin{aligned} \|\Phi_r(s) - s\| &= \left\| \sum_{i=0}^n z_i (f'_i - g'_i) + \frac{o(r)}{r} - \sum_{i=0}^n z_i e_i \right\| = \left\| \sum_{i=0}^n z_i (f_i - f'_i) + \sum_{i=0}^n z_i (g'_i - g_i) + \frac{o(r)}{r} \right\| \\ &\leq \epsilon + \frac{o(r)}{r} \leq \frac{\alpha}{2}. \end{aligned}$$

Mais uma vez, do Teorema do Ponto Fixo de Brouwer, $\Phi_r^{-1} \left(B_{\mathbb{R}^n} \left[0, \frac{\alpha}{2} \right] \right) \subseteq S$, em particular, existe $z_r = (z_{0,r}, \dots, z_{n,r}) \in S$ satisfazendo $\Xi_1(rz_r) = \Xi_2(rz_r)$.

Afirmção. $\Xi_1(rz_r) \neq 0$.

De fato, como

$$\frac{\Xi_1(rz_r)}{r} = \sum_{i=0}^n z_i f'_i + \frac{o(r)}{r}$$

e a primeira parcela pertence à envoltória convexa dos vetores $\{f'_0, \dots, f'_n\}$, como a correspondente envoltória dos vetores $\{f_0, \dots, f_n\}$ não contém a origem (foi assumido acima sem perda de generalidade) e ambas são conjuntos compactos, escolhendo $\epsilon > 0$ pequeno o suficiente, é possível garantir que aquela não contém, também, a origem, o que, escolhendo $r > 0$ pequeno o suficiente, garante que $\Xi_1(rz_r) \neq 0$.

Desse modo, o ponto $\bar{q} : \Xi_1(rz_r) \in A_1 \cap A_2 \setminus \{0\}$ pertence a $B_{\mathbb{R}^n} \left(0, r \frac{\alpha}{2}\right)$. Da arbitrariedade de r e do fato de qualquer vizinhança da origem contém uma bola do tipo $B_{\mathbb{R}^n} \left(0, r \frac{\alpha}{2}\right)$, concluímos que q é ponto limite de $A_1 \cap A_2$. $\square$

Definição 54 (página 72 de [50]). A seguinte propriedade é denominada **Princípio da Interseção**: dados $C_1, \dots, C_m \subseteq \mathbb{R}^n$, se, para cada $i \in \{1 : m\}$, o cone K_i é aproximante de C_i em um ponto x_0 , e os cones $K_1, \dots, K_m$ são inseparáveis, então, x_0 é ponto de acumulação de $\bigcap_{i=1}^m C_i$.

Corolário 55. *Os cones aproximantes de Boltjanskij satisfazem o Princípio da Interseção.*

Demonstração. Segue da **Proposição 36** e do **Teorema 53**. $\square$

2.4 Princípios Variacionais

2.4.1 Penalização exata

Os resultados a seguir serão feitos supondo que X é espaço de Banach real.

Definição 56 (página 210 de [11]). Dado $S \subseteq X$, função $d_S : X \rightarrow [0, \infty[$ definida por $d_S(x) = \text{dist}(x, S)$ é denominada **função distância associada a S** .

Propriedade 57 (página 210 de [11]). *A função d_S é, globalmente, lipschitziana de posto um.*

Demonstração. Dados $x_1, x_2 \in X$, da desigualdade triangular, qualquer $y \in S$ satisfaz $d_S(x_1) - \|x_1 - x_2\|_X \leq \|x_1 - y\|_X - \|x_1 - x_2\|_X \leq \|x_2 - y\|_X$, assim, tomando ínfimo do lado direito e rearranjando a desigualdade, vem $d_S(x_1) - d_S(x_2) \leq \|x_1 - x_2\|_X$. $\square$

Proposição 58 (página 211 de [11]). *Dado $S \subseteq X$ um conjunto fechado, suponha que $f : X \rightarrow \mathbb{R}$ seja lipschitziano de posto M em um aberto U contendo S . Nessas condições:*

- Se $x^* \in S$ minimiza f em tal conjunto, então, o funcional $f + Md_S$ atinge mínimo em U no ponto x^* ;
- Se $x^* \in U$ minimiza $f + \tilde{M}d_S$ em tal conjunto para algum $\tilde{M} > M$, então, $x^* \in S$ e minimiza f em tal conjunto.

Demonstração. • Supondo que $x^* \in S$ minimiza f em tal conjunto, sejam $x \in U$ e $\epsilon > 0$ e escolha $s \in S$ de modo que $\|x - s\| \leq d_S(x) + \epsilon$. O fato de que x^* minimiza f em S e a continuidade lipschitziana de f garantem que

$$f(x^*) \leq f(s) \leq f(x) + M\|s - x\|_X \leq f(x) + Md_S(x) + M\epsilon.$$

Fazendo $\epsilon \downarrow 0$, obtemos que x^* minimiza $f + Md_S$ em U ;

- Supondo que $x^* \in U$ minimiza $f + \tilde{M}d_S$ em tal conjunto com $\tilde{M} > M$, se $x^* \in S$, então, o fato de que tal conjunto é fechado garante que $d_S(x^*) > 0$, o que permite escolher $s \in S$ de modo que

$$\|s - x^*\|_X < \frac{\tilde{M}}{M}d_S(x^*).$$

Ainda, da continuidade lipschitziana de f , segue que $f(s) \leq f(x^*) + M\|s - x^*\|_X$, logo, como x^* minimiza $f + \tilde{M}d_S$ em U , vale que

$$f(x^*) + \tilde{M}d_S(x^*) \leq f(s) \leq f(x^*) + M\|s - x^*\|_X < f(x^*) + \tilde{M}d_S(x^*) \quad (\text{absurdo!})$$

Portanto, $x^* \in S$ e minimiza f em tal conjunto. □

2.4.2 Princípio Variacional de Ekeland

Teorema 59 (página 112 de [48]). *Dados $g : X \rightarrow \mathbb{R}$ semicontínuo inferior e $x_0 \in X$, suponha que existem números $\lambda, \alpha > 0$ satisfazendo*

$$g(x_0) \leq \inf g(X) + \lambda\alpha. \quad (2.3)$$

Nessas condições, existe $x^ \in B_X[x_0, \lambda]$ que minimiza $x \mapsto g(x) + \alpha\|x - x^*\|_X$ e satisfaz $g(x^*) \leq g(x_0)$.*

Demonstração. Defina a multifunção $T : X \rightsquigarrow X$ por $T(x) = \{y \in X \mid g(y) + \alpha\|x - y\|_X \leq g(x)\}$. Dados $x \in X$ e $\{y_i\} \subseteq T(x)$ tendendo a $y \in X$, a semicontinuidade inferior de g garante que

$$g(y) + \alpha\|x - y\|_X = g(y) + \alpha \lim_{i \uparrow \infty} \|x - y_i\|_X \leq \liminf_{i \uparrow \infty} g(y_i) + \alpha \liminf_{i \uparrow \infty} \|x - y_i\|_X \leq g(x),$$

o que mostra que $y \in T(x)$ e, conseqüentemente, que T assume, somente, valores fechados.

Afirmção. Para todos $x, y \in X$, $y \in T(x) \Rightarrow T(y) \subseteq T(x)$.

De fato, dado $z \in T(y)$, temos que

$$g(z) + \alpha\|x - z\|_X \leq g(z) + \alpha\|y - z\|_X + \alpha\|x - y\|_X \leq g(y) + \alpha\|x - y\|_X \leq g(x),$$

o que mostra que $z \in T(x)$.

Agora, definindo $\xi : X \rightarrow \mathbb{R}$ por $\xi(x) = \inf(g \circ T)(x)$, como $\xi(x) \leq g(x) - \alpha\|x - y\|_X$ para todo $y \in T(x)$, segue que $\text{diam } T(x) \leq \frac{2}{\alpha}[g(x) - \xi(x)]$. Ainda, dessa definição, é possível construir uma sequência $\{x_i\} \subseteq X$ escolhendo, para cada $j \in \mathbb{Z}_+$, $x_{j+1} \in T(x_j)$ satisfazendo $g(x_{j+1}) - \xi(x_j) \leq \frac{1}{2^j}$. Para cada $i \in \mathbb{Z}_+$, como $T(x_{i+1}) \subseteq T(x_i)$, vale $\xi(x_i) \leq \xi(x_{i+1})$, logo, como $\xi \leq g$ em X ,

$$\xi(x_{i+1}) \leq g(x_{i+1}) \leq \xi(x_i) + \frac{1}{2^i} \leq \xi(x_{i+1}) + \frac{1}{2^i}.$$

Isso mostra que $0 \leq g(x_{i+1}) - \xi(x_{i+1}) \leq \frac{1}{2^i}$, então, qualquer $i \in \mathbb{Z}_+$ satisfaz

$$\text{diam } T(x_{i+1}) \leq \frac{1}{2^{i-1}\alpha},$$

o que acarreta que $\{T(x_i) \mid i \in \mathbb{N}\}$ é uma família de conjuntos encaixados em ordem decrescente cujos diâmetros tendem a zero para $i \uparrow \infty$. A partir disso, $\{x_i\}$ é sequência de Cauchy e, consequentemente, converge a $x^* \in X$, o que mostra que

$$\bigcap_{i=0}^{\infty} T(x_i) = \{x^*\},$$

logo, $g(x^*) + \alpha\|x_0 - x^*\|_X \leq g(x_0)$. Também, da afirmação provada, $T(x^*) = \{x^*\}$, logo, $x \neq x^* \Rightarrow g(x^*) < g(x) + \alpha\|x - x^*\|_X$. Por fim,

$$g(x^*) + \alpha\|x_0 - x^*\|_X \leq g(x_0) \leq \inf g(X) + \lambda\alpha \leq g(x^*) + \lambda\alpha,$$

o que mostra que $\|x_0 - x^*\|_X \leq \lambda$. □

3 Princípio de Pontrjagin

3.1 Problemas de controle ótimo

Nessa seção, apresentaremos os problemas de otimização que serão trabalhados neste capítulo. O trabalho, aqui, é baseado em [1].

3.1.1 Problemas sentido-estrito

Começaremos com o problema original, no qual a variável temporal é independente e, por causa disso, tal problema não apresenta impulsos.

Hipótese 1 (páginas 818 e 819 de [1]). Considere:

- Um conjunto fechado $\mathcal{T} \subseteq [0, \infty[\times \mathbb{R}^n$, que será denominado *conjunto alvo*;
- Um compacto $A \subseteq \mathbb{R}^q$ e um cone fechado $C \subseteq \mathbb{R}^m$ formado pelo produto cartesiano de um cone fechado de $\mathbb{R}^{m_1}$ contendo os eixos canônicos por um cone saliente fechado de $\mathbb{R}^{m_2}$; tanto A quanto C serão denominados *conjuntos de controle*;
- Uma aplicação $f \in \mathcal{C}^0(\mathbb{R}^n \times A; \mathbb{R}^n)$ cujas primeiras n derivadas parciais de primeira ordem são contínuas, que será denominada *dinâmica de derrapagem*;
- Campos de vetores $g_1, \dots, g_m \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$;
- Uma função $\ell : \mathbb{R}^n \times C \times A \rightarrow \mathbb{R}$ definida por $\ell(x, u, a) = \ell_0(x, a) + \ell_1(x, u)$ com ℓ_0 e a *função recessão*

$$\widehat{\ell}_1(x, w_0, w) := \lim_{r \rightarrow w_0} r \ell_1\left(x, \frac{w}{r}\right)$$

sendo contínuas e tendo as primeiras n derivadas parciais de primeira ordem contínuas; ℓ será denominada *lagrangiano*;

- A função $\Psi \in \mathcal{C}^1(\mathbb{R} \times \mathbb{R}^n, \mathbb{R})$, que será chamada *custo final*.

Antes de prosseguir, introduzimos algumas notações que podem ser úteis para o texto:

- A norma de L^p ($p \in [1, \infty]$) será denotada por $\|\cdot\|_p$;
- As bolas aberta e fechada de centro $x_0 \in L^p$ ($p \in [1, \infty]$) e raio $r > 0$ serão denotadas, respectivamente, por $B_p(x_0, r)$ e $B_p[x_0, r]$.

Definição 60 (página 819 de [1]). O conjunto U consistindo de todas as ternas (T, u, a) satisfazendo $T > 0$, $u \in L^1([0, T]; C)$ e $a \in L^1([0, T]; A)$ é denominado **conjunto de controle sentido-estrito**.

Definição 61 (página 819 de [1]). Dados $(T, u, a) \in U$ e (x, v) uma solução de Karatheodori de

$$\begin{cases} \dot{x} = f(x, a) + \sum_{i=1}^m g_i(x)u_i, \\ \dot{v} = |u|, \\ (x, v)(0) = (\check{x}, 0), \end{cases} \quad (3.1)$$

a quintúpla (T, u, a, x, v) é denominada **processo sentido-estrito**. Ainda, esse processo é denominado **factível** se $(T, x(T)) \in \mathcal{T}$ e $v(T) \in [0, K]$, em que $K \in [0, \infty]$.

Com todos esses conceitos, é possível introduzir o *problema de controle ótimo sentido-estrito*

$$\begin{cases} \text{minimizar} & \Psi(T, x(T)) + \int_0^T \ell(x, u, a)(t) dt \\ \text{sobre todos os processos sentido-estrito factíveis} & (T, u, a, x, v). \end{cases} \quad (3.2)$$

Definição 62 (página 819 de [1]). Um processo sentido-estrito factível $(\bar{T}, \bar{u}, \bar{a}, \bar{x}, \bar{v})$ é denominado **minimizador sentido-estrito local** se existe $\delta > 0$ de modo que qualquer processo sentido-estrito (T, u, a, x, v) verificando $\text{dist}((T, x, v); (\bar{T}, \bar{x}, \bar{v})) < \delta$ (a distância, aqui, é a usual para espaços-produto) satisfaz

$$\Psi(\bar{T}, \bar{x}(\bar{T})) + \int_0^{\bar{T}} \ell(\bar{x}, \bar{u}, \bar{a})(t) dt \leq \Psi(T, x(T)) + \int_0^T \ell(x, u, a)(t) dt. \quad (3.3)$$

Se esta desigualdade vale para qualquer processo sentido-estrito factível, $(\bar{T}, \bar{u}, \bar{a}, \bar{x}, \bar{v})$ é chamado **minimizador sentido-estrito global**.

3.1.2 Problemas espaço-temporais

Passaremos, agora, ao problema mais geral, que parametriza tanto as variáveis temporal e espacial e permite, com isso, um caráter impulsivo. Antes disso, porém, dado um funcional f , seu *ínfimo essencial*, ou seja, o maior $r \in \mathbb{R}$ para o qual a medida de Lebesgue de $f^{-1}(]-\infty, r[)$ é nula, será denotado por $\text{ess inf } f$.

Definição 63 (página 820 de [1]). O conjunto $\mathcal{W}$ das quadrúplas (S, w_0, w, α) satisfazendo $S > 0$, $w_0 \in L^\infty([0, S]; [0, \infty[)$, $w \in L^\infty([0, S]; C)$, $\alpha \in L^\infty([0, S]; A)$ e $\text{ess inf}(w_0 + |w|) > 0$ é denominado **conjunto de controle espaço-temporal**.

Definição 64 (página 820 de [1]). Dados $(S, w_0, w, \alpha) \in \mathcal{W}$ e (y_0, y, β) uma solução Karatheodori de

$$\begin{cases} \dot{y}_0 = w_0, \\ \dot{y} = f(y, \alpha)w_0 + \sum_{i=1}^m g_i(y)w_i, \\ \dot{\beta} = |w|, \\ (y_0, y, \beta)(0) = (0, \check{x}, 0) \end{cases} \quad (3.4)$$

o processo $(S, w_0, w, \alpha, y, y_0, \beta)$ é denominado **processo espaço-temporal**. Mais ainda, ele é chamado **factível** se $(y_0, y)(S) \in \mathfrak{T}$ e $\beta(S) \in [0, K]$.

A partir disso, é possível introduzir o *problema espaço-temporal*

$$\begin{cases} \text{minimizar} & \Psi(y_0, y)(S) + \int_0^S \ell^e(y, w_0, w, \alpha)(s) \, ds \\ \text{sobre todos os processos espaço-temporais factíveis} & (S, w_0, w, \alpha, y_0, y, \beta), \end{cases} \quad (3.5)$$

em que o *lagrangiano extendido* $\ell^e : \mathbb{R}^n \times [0, \infty[\times C \times A \rightarrow \mathbb{R}$ é definido por $\ell^e(y, w_0, w, \alpha) = \ell_0(y, \alpha)w_0 + \hat{\ell}_1(y, w_0, w)$.

Propriedade 65 (página 820 de [1]). *Dada uma função crescente, bijetiva e bilipschitziana $\varsigma : [0, S] \rightarrow [0, \tilde{S}]$, e definindo, para cada $(S, w_0, w, \alpha, y, y_0, \beta)$,*

$$\begin{aligned} (\widetilde{w}_0, \widetilde{w}) &= [(w_0, w) \circ \varsigma^{-1}] \frac{d\varsigma^{-1}}{ds}, \\ (\tilde{\alpha}, \tilde{y}_0, \tilde{y}, \tilde{\beta}) &= (\alpha, y_0, y, \beta) \circ \varsigma^{-1}, \end{aligned}$$

$(S, w_0, w, \alpha, y, y_0, \beta)$ é um processo espaço-temporal para (3.5) se, e somente se, $(\tilde{S}, \widetilde{w}_0, \widetilde{w}, \tilde{\alpha}, \tilde{y}_0, \tilde{y}, \tilde{\beta})$ também o é, e os custos coincidem.

Demonstração. Claramente, se $(S, w_0, w, \alpha, y, y_0, \beta)$ é um processo espaço-temporal para (3.5), então,

$$\begin{aligned} \frac{d\tilde{y}_0}{ds} &= (\dot{y}_0 \circ \varsigma^{-1}) \frac{d\varsigma^{-1}}{ds} = (w_0 \circ \varsigma^{-1}) \frac{d\varsigma^{-1}}{ds} = \widetilde{w}_0, \\ \frac{d\tilde{y}}{ds} &= (\dot{y} \circ \varsigma^{-1}) \frac{d\varsigma^{-1}}{ds} = \left\{ \left[f(y, \alpha)w_0 + \sum_{i=1}^m g_i(y)w_i \right] \circ \varsigma^{-1} \right\} \frac{d\varsigma^{-1}}{ds} \\ &= \left\{ f(\tilde{y}, \tilde{\alpha})(w_0 \circ \varsigma^{-1}) + \sum_{i=1}^m g_i(\tilde{y})(w_i \circ \varsigma^{-1}) \right\} \frac{d\varsigma^{-1}}{ds} = f(\tilde{y}, \tilde{\alpha})\widetilde{w}_0 \\ &\quad + \sum_{i=1}^m g_i(\tilde{y})\widetilde{w}_i, \\ \frac{d\tilde{\beta}}{ds} &= (\dot{\beta} \circ \varsigma^{-1}) \frac{d\varsigma^{-1}}{ds} = |w \circ \varsigma^{-1}| \frac{d\varsigma^{-1}}{ds} = |\widetilde{w}|, \\ (\tilde{y}_0, \tilde{y}, \tilde{\beta})(0) &= (y_0, y, \beta) \circ \varsigma^{-1}(0) = (\tilde{y}_0, \tilde{y}, \tilde{\beta})(0) = (0, \check{x}, 0), \end{aligned}$$

o que mostra que $(\tilde{S}, \tilde{w}_0, \tilde{w}, \tilde{\alpha}, \tilde{y}, \tilde{y}_0, \tilde{\beta})$ o faz também.

Agora, se $(\tilde{S}, \tilde{w}_0, \tilde{w}, \tilde{\alpha}, \tilde{y}, \tilde{y}_0, \tilde{\beta})$ é um processo espaço-temporal para (3.5), como

$$\begin{aligned} (w_0, w) &= [(\tilde{w}_0, \tilde{w}) \circ \varsigma] \dot{\varsigma} \\ (\alpha, y_0, y, \beta) &= (\tilde{\alpha}, \tilde{y}_0, \tilde{y}, \tilde{\beta}) \circ \varsigma, \end{aligned}$$

então,

$$\begin{aligned} \dot{y}_0 &= \left(\frac{d\tilde{y}_0}{ds} \circ \varsigma \right) \dot{\varsigma} = (\tilde{w}_0 \circ \varsigma) \dot{\varsigma} = w_0, \\ \frac{d\tilde{y}}{ds} &= (\dot{y} \circ \varsigma) \dot{\varsigma} = \left\{ \left[f(\tilde{y}, \tilde{\alpha}) \tilde{w}_0 + \sum_{i=1}^m g_i(\tilde{y}) \tilde{w}_i \right] \circ \varsigma \right\} \dot{\varsigma} \\ &= \left\{ f(y, \alpha) (\tilde{w}_0 \circ \varsigma) + \sum_{i=1}^m g_i(y) (\tilde{w}_i \circ \varsigma) \right\} \dot{\varsigma} = f(y, \alpha) w_0 + \sum_{i=1}^m g_i(y) w_i, \\ \dot{\beta} &= \left(\frac{d\tilde{\beta}}{ds} \circ \varsigma \right) \dot{\varsigma} = |\tilde{w} \circ \varsigma| \dot{\varsigma} = |w|, \\ (y_0, y, \beta)(0) &= (\tilde{y}_0, \tilde{y}, \tilde{\beta}) \circ \varsigma(0) = (y_0, y, \beta)(0) = (0, \check{x}, 0), \end{aligned}$$

o que mostra que $(S, w_0, w, \alpha, y, y_0, \beta)$ o é também.

Por fim,

$$\begin{aligned} \Psi(\tilde{y}_0, \tilde{y})(\tilde{S}) + \int_0^{\tilde{S}} \ell^e(\tilde{y}, \tilde{w}_0, \tilde{w}, \tilde{\alpha})(s) ds &= \Psi(y_0, y)((\varsigma^{-1} \circ \varsigma)(S)) \\ &+ \int_0^{\varsigma(S)} \ell^e(y, w_0, w, \alpha)(\varsigma^{-1}(s)) \frac{d\varsigma^{-1}}{ds} ds \\ &= \Psi(y_0, y)(S) + \int_0^S \ell^e(y, w_0, w, \alpha)(s) ds, \end{aligned}$$

o que mostra que os custos coincidem. $\square$

Definição 66 (página 820 de [1]). Um processo espaço-temporal factível $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é denominado **minimizador espaço-temporal local** se existe $\delta > 0$ tal que qualquer processo espaço-temporal factível $(S, w_0, w, \alpha, y_0, y, \beta)$ verificando

$$\text{dist}((S, y_0, y, \beta); (\bar{S}, \bar{y}_0, \bar{y}, \bar{\beta})) < \delta$$

satisfaz

$$\Psi(\bar{y}_0, \bar{y})(\bar{S}) + \int_0^{\bar{S}} \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})(s) ds \leq \Psi(y_0, y)(S) + \int_0^S \ell^e(y, w_0, w, \alpha)(s) ds. \quad (3.6)$$

Se essa desigualdade vale para qualquer processo espaço-temporal, $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é chamado **minimizador espaço-temporal global**.

A partir dessa propriedade, é possível estabelecer uma relação $\sim$ de equivalência entre dois processos: $(S, w_0, w, \alpha, y, y_0, \beta)$ é equivalente a $(\tilde{S}, \tilde{w}_0, \tilde{w}, \tilde{\alpha}, \tilde{y}, \tilde{y}_0, \tilde{\beta})$ se, e somente se, existe uma $\varsigma : [0, S] \rightarrow [0, \tilde{S}]$ com as características exibidas na **Propriedade 65**.

Combinando essa definição com o lema anterior e notando que uma aplicação bilipschitziana associa pontos de uma bola aberta a pontos de outra bola aberta, surge o seguinte resultado:

Proposição 67 (página 820 de [1]). *Um processo espaço-temporal factível $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é um minimizador local (respectivamente, global) se, e somente se, todo processo equivalente por $\sim$ também é um minimizador local (respectivamente, global).*

O resultado acima permite considerar (3.5) como um problema no espaço quociente e, consequentemente, permite substituir qualquer processo factível por sua parametrização canônica.

Definição 68 (página 821 de [1]). Dado um processo $(S, w_0, w, \alpha, y_0, y, \beta)$ e colocando $\varsigma(s) = \int_0^s [w_0(r) + |w(r)|] dr$, o processo $(S_c, w_{0,c}, w_c, \alpha_c, y_{0,c}, y_c, \beta_c)$ definido por

$$\begin{aligned} S_c &= \varsigma(S), \\ (w_{0,c}, w_c) &= [(w_0, w) \circ \varsigma^{-1}] \frac{d\varsigma^{-1}}{ds}, \\ (\alpha_c, y_{0,c}, y_c, \beta_c) &= (\alpha, y_0, y, \beta) \circ \varsigma^{-1} \end{aligned} \quad (3.7)$$

é denominado **parametrização canônica de $(S, w_0, w, \alpha, y_0, y, \beta)$** .

Definição 69 (página 821 de [1]). O conjunto a seguir é denominado **conjunto de controle espaço-temporal canônico**

$$\mathcal{W}_c = \{(S, w_0, w, \alpha) \in \mathcal{W} \mid w_0(s) + |w(s)| = 1 \quad \forall s \in [0, S]\},$$

e seus elementos são denominados **controles espaço-temporais canônicos**.

Incorporando os problemas sentido-estrito neles

Dado (T, u, a, x, v) um processo sentido-estrito, pondo

$$\begin{aligned} \varsigma(t) &= \int_0^t [1 + |u(\tau)|] d\tau, \\ S &= \varsigma(T), \\ y_0 &= \varsigma^{-1} : [0, S] \rightarrow [0, T], \end{aligned} \quad (3.8)$$

o processo

$$(S, w_0, w, \alpha, y_0, y, \beta) := (S, \dot{y}_0, (u \circ y_0)\dot{y}_0, a \circ y_0, y_0, x \circ y_0, v \circ y_0) \quad (3.9)$$

é um processo espaço-temporal canônico com $w_0 > 0$ quase sempre. Também, dado um processo $(S, w_0, w, \alpha, y_0, y, \beta)$ com $w_0 > 0$ quase sempre, a função $y_0 : [0, S] \rightarrow [0, T]$ — definida por $y_0(s) = \int_0^s w_0(\tau) d\tau$, é crescente e bijetiva — possui uma inversa $\varsigma \in \mathcal{AC}([0, T]; [0, S])$ e

$$(T, u, a, x, v) := \left(T, (w \circ \varsigma) \dot{\varsigma}, \alpha \circ \varsigma, y \circ \varsigma, \beta \circ \varsigma \right)$$

é um processo sentido-estrito. Logo, o conjunto de controle sentido-estrito pode ser identificado com o conjunto dos processos espaço-temporais satisfazendo $w_0 > 0$ quase sempre.

Proposição 70 (página 821 de [1]). *Um processo sentido-estrito $(\bar{T}, \bar{u}, \bar{a}, \bar{x}, \bar{v})$ é um minimizador local para (3.2) se, e somente se, o processo espaço-temporal $(\bar{S}, \bar{w}_0, \bar{w}, \bar{a}, \bar{y}_0, \bar{y}, \bar{\beta})$ definido como em (3.8) e (3.9) é um minimizador local para (3.5) sobre todos os processos factíveis satisfazendo $w_0 > 0$ quase sempre.*

Como visto, quando $w_0 > 0$ quase sempre, o problema (3.5) nada mais é do que uma reparametrização do (3.2). Porém, permitindo que w_0 se anule em um conjunto de medida positiva, pode-se observar o aparecimento de *impulsos*, isto é, intervalos na variável s nos quais a variável temporal permanece constante, resultando em “evolução instantânea” das variáveis de controle e de estado. No caso mais extremo que $w_0 = 0$ quase sempre, o processo correspondente só apresenta evoluções instantâneas das variáveis de controle e estado, o que o caracteriza como, *totalmente, impulsivo*.

Não é objetivo deste capítulo estudar problemas sentido-estrito e/ou, totalmente, impulsivos, porém, ainda assim, apresentamos uma versão do Princípio de Pontrjagin para aqueles — e que melhora um resultado obtido por [10] — e condições de posto para ocorrência destes.

3.2 Versão de primeira ordem

O foco desta seção é apresentar uma versão do Princípio de Pontrjagin de primeira ordem para o problema (3.5) e algumas de suas consequências, em especial, uma “concorrência” entre as funções hamiltonianas *impulsiva* e de *derrapagem*. O trabalho, aqui, é baseado em [1].

Devido à independência de taxa do problema (3.5), um minimizador deste será assumido canônico. Pondo $W = \{(w_0, w) \in [0, \infty[\times C \mid w_0 + |w| = 1\}$, considere a função *hamiltoniana não-maximizada* $H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times [0, \infty[\times C \times A \rightarrow \mathbb{R}$ definida por

$$H(x, p_0, p, \pi, \lambda, w_0, w, a) = p_0 w_0 + \left\langle p; f(x, a) w_0 + \sum_{i=1}^m g_i(x) w_i \right\rangle + \pi |w| - \lambda \ell^e(x, w_0, w, a)$$

e a *hamiltoniana regular* $\mathfrak{h} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ formada maximizando a expressão acima nas últimas três variáveis quando $(w_0, w) \in W$ e $a \in A$.

Teorema 71 (página 822 de [1]). *Se $(\bar{S}, \bar{w}_0, \bar{w}, \bar{a}, \bar{y}_0, \bar{y}, \bar{\beta})$ é um minimizador local para (3.5), então, para cada cone aproximante de Boltjanskij Γ para $\mathcal{T}$ em $(\bar{y}_0, \bar{y})(\bar{S})$, existem $p_0 \in \mathbb{R}$, $p \in \mathcal{AC}([0, S]; \mathbb{R}^n)$, $\pi \in]-\infty, 0]$ e $\lambda \in [0, \infty[$ satisfazendo:*

- (Não-trivialidade) *Nem todos entre p_0 , p e λ são nulos. Mais ainda, se $\bar{y}_0(\bar{S}) > 0$, então, p e λ não podem ser nulos ao mesmo tempo;*

- (Não-transversalidade)

$$\begin{aligned} (p_0, p(\bar{S})) &\in -\lambda \nabla \Psi(\bar{y}_0, \bar{y})(\bar{S}) \\ \pi &\in \begin{cases} \{0\} & \text{se } \bar{\beta}(\bar{S}) < K, \\]0, \infty[& \text{se } \bar{\beta}(\bar{S}) = K; \end{cases} \end{aligned} \quad (3.10)$$

- (Equação adjunta) *Para quase todo* $s \in [0, \bar{S}]$,

$$\dot{p}(s) = -\frac{\partial H}{\partial x}(\bar{y}(s), p(s), \pi, \lambda, \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s)); \quad (3.11)$$

- (Maximização de primeira ordem) *Para quase todo* $s \in [0, \bar{S}]$,

$$H(\bar{y}(s), p_0, p(s), \pi, \lambda, \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s)) = \mathfrak{h}(\bar{y}(s), p_0, p(s), \pi, \lambda); \quad (3.12)$$

- (Anulação da hamiltoniana) *Para todo* $s \in [0, \bar{S}]$,

$$\mathfrak{h}(\bar{y}(s), p_0, p(s), \pi, \lambda) = 0. \quad (3.13)$$

A demonstração desse teorema pode ser feita ignorando o raciocínio relativo às variações tipo-colchete na seção 3.5 e repetindo o restante.

Definição 72 (página 823 de [1]). Um processo $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é um **extremal** se existe algum multiplicador (p_0, p, π, λ) satisfazendo as condições. Mais ainda, tal extremal é denominado **normal** se nenhuma escolha de multiplicador satisfaz $\lambda = 0$; **anormal** caso contrário; e, **estritamente, anormal** se qualquer escolha de multiplicador satisfaz $\lambda = 0$.

Defina, agora, a *hamiltoniana impulsiva* $\mathfrak{h}_{\text{imp}} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ e a *derrapante* $\mathfrak{h}_{\text{der}} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ por

$$\begin{aligned} \mathfrak{h}_{\text{der}}(x, p_0, p, \lambda) &= \max\{p_0 + \langle p, f(x, a) \rangle - \lambda \ell_0(x, a) \mid a \in A\}, \\ \mathfrak{h}_{\text{imp}}(x, p, \pi, \lambda) &= \max\left\{\left\langle p, \sum_{i=1}^m g_i(x) w_i \right\rangle + \pi - \lambda \ell_1(x, w) \mid w \in C \cap \mathbb{S}^{m-1}\right\}. \end{aligned}$$

Corolário 73 (página 823 de [1]). *Dados um extremal $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ e (p_0, p, π, λ) , existe um conjunto nulo $\mathcal{N} \subseteq [0, \bar{S}]$ tal que, para todo $s \in [0, \bar{S}] \setminus \mathcal{N}$,*

$$0 = \mathfrak{h}(\bar{y}(s), p_0, p(s), \pi, \lambda) = \max\{\mathfrak{h}_{\text{der}}(\bar{y}(s), p_0, p(s), \lambda); \mathfrak{h}_{\text{imp}}(\bar{y}(s), p(s), \pi, \lambda)\}, \quad (3.14)$$

$$0 = \bar{w}_0(s) [p_0 + \langle p(s); f(\bar{y}(s), \bar{\alpha}(s)) \rangle - \lambda \ell_0(\bar{y}(s), \bar{\alpha}(s))], \quad (3.15)$$

$$0 = \left\langle p(s), \sum_{i=1}^m (g_i \circ \bar{y})(s) \bar{w}_i(s) \right\rangle + \pi |\bar{w}(s)| - \lambda \ell_1(\bar{y}(s), \bar{w}(s)). \quad (3.16)$$

Em particular,

- Se algum $s \in [0, \bar{S}] \setminus \mathcal{N}$ satisfaz $\mathfrak{h}_{\text{der}}(\bar{y}(s), p_0, p(s), \lambda) < 0$, então, $\bar{w}_0(s) = 0$ e

$$\left\langle p(s), \sum_{i=1}^m (g_i \circ \bar{y})(s) \bar{w}_i(s) \right\rangle + \pi - \lambda \ell_1(\bar{y}(s), \bar{w}(s)) = \mathfrak{h}_{\text{imp}}(\bar{y}(s), p(s), \pi, \lambda) = 0;$$

- Se algum $s \in [0, \bar{S}] \setminus \mathcal{N}$ satisfaz $\mathfrak{h}_{\text{imp}}(\bar{y}(s), p(s), \pi, \lambda) < 0$, então, $\bar{w}(s) = 0$ e

$$p_0 + \langle p(s); f(\bar{y}(s), \alpha(s)) \rangle - \lambda \ell_0(\bar{y}(s), \alpha(s)) = \mathfrak{h}_{\text{der}}(\bar{y}(s), p(s), \pi, \lambda) = 0.$$

Demonstração. De (3.13), qualquer $s \in [0, \bar{S}]$ satisfaz

$$p_0 w_0 + \left\langle p(s); f(\bar{y}(s), a) w_0 + \sum_{i=1}^m (g_i \circ \bar{y})(s) w_i \right\rangle + \pi |w| - \lambda [\ell_0(\bar{y}(s), a) w_0 + \ell_1(\bar{y}(s), w)] \leq 0$$

para todo $(w_0, w, a) \in W \times A$. Em particular, escolhendo $w = 0$, vale $w_0 = 1$ e

$$p_0 + \langle p(s); f(\bar{y}(s), a) \rangle - \lambda \ell_0(\bar{y}(s), a) \leq 0 \quad \forall a \in A,$$

enquanto que, para $w_0 = 0$, é obtido que

$$\left\langle p(s), \sum_{i=1}^m (g_i \circ \bar{y})(s) w_i \right\rangle \quad \forall (w, a) \in C \times A, |w| = 1.$$

A partir disso, $\mathfrak{h}_{\text{der}}(\bar{y}(s), p(s), \pi, \lambda) \leq 0$ e $\mathfrak{h}_{\text{imp}}(\bar{y}(s), \pi, \lambda) \leq 0$, o que permite afirmar, por (3.13) novamente, que o máximo dos dois é zero. Tomando $\mathcal{N} \subseteq [0, \bar{S}]$ o conjunto nulo sobre cujo complementar é verificada (3.12), fica demonstrada (3.14).

Ainda, se $s \in [0, \bar{S}] \setminus \mathcal{N}$, por (3.12) e (3.13), vale

$$\begin{aligned} 0 &= \bar{w}_0(s) [p_0 + \langle p(s); f(\bar{y}(s), \bar{\alpha}(s)) \rangle - \lambda \ell_0(\bar{y}(s), \bar{\alpha}(s))] \\ &\quad + \left\langle p(s); \sum_{i=1}^m (g_i \circ \bar{y})(s) \bar{w}_i(s) \right\rangle + \pi |\bar{w}(s)| - \lambda \ell_1(\bar{y}(s), \bar{w}(s)). \end{aligned}$$

Como cada uma das parcelas é não-negativa pelas duas desigualdades acima, ficam provadas (3.15) e (3.16).

Para provar o primeiro item, supondo $\mathfrak{h}_{\text{der}}(\bar{y}(s), p(s), \pi, \lambda) < 0$, (3.15) implica $\bar{w}_0(s) = 0$ e $|\bar{w}(s)| = 1$, o que prova o item por (3.16). Finalmente, para provar o segundo item, supondo $\mathfrak{h}_{\text{imp}}(\bar{y}(s), p(s), \pi, \lambda) < 0$, a 1-homogeneidade positiva de H com respeito a (w_0, w) e (3.16) implicam $\bar{w}(s) = 0$, logo, $\bar{w}_0(s) = 1$ e o resultado fica provado por (3.15). $\square$

Os vetores canônicos dos espaços euclidianos serão denotados por e_j .

Corolário 74 (página 824 de [1]). Sendo $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ um extremal canônico para (3.5) ligado a um multiplicador (p_0, p, π, λ) , se

$$\pi = 0, \quad \lambda \ell^e(\bar{y}(s), 0, \pm e_i, a) = 0 \quad \forall s \in [0, \bar{S}], \quad i \in \{1 : m_1\}, \quad (3.17)$$

então, $\langle p(s), (g_i \circ \bar{y})(s) \rangle = 0$ para todo $s \in [0, \bar{S}]$ e $i \in \{1 : m_1\}$.

Demonstração. Por (3.13) e da definição da hamiltoniana, dados $s \in [0, \bar{S}]$ e $(w_0, w, a) \in W \times A$, vale

$$p_0 w_0 + \left\langle p(s); f(\bar{y}(s), a) w_0 + \sum_{i=1}^m (g_i \circ \bar{y})(s) w_i \right\rangle + \pi |w| - \lambda [\ell_0(\bar{y}(s), a) w_0 + \ell_1(\bar{y}(s), w)] \leq 0.$$

Escolhendo, em particular, $w_0 = 0$ e $w = \pm e_i$ para $i \in \{1 : m_1\}$, o resultado fica provado. $\square$

Observação 75 (página 824 de [1]). Do Teorema 71, $\pi = 0$ quando $\bar{\beta}(\bar{S}) < K$. Mais ainda, a hipótese $\lambda \ell^e(\bar{y}(s), 0, \pm e_i, a) = 0$ é satisfeita quando $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é um extremal anormal ou quando $\hat{\ell}_1(x, 0, w) = 0$ para quaisquer $x \in \mathbb{R}^n$ e $w \in \mathbb{R}^{m_1} \times \{0\}^{m_2}$.

3.3 Versão de alta ordem

O foco desta seção é apresentar a versão de alta ordem do Princípio de Pontrjagin para o problema (3.5) e algumas de suas consequências.

Teorema 76 (página 825 de [1]). *Assumindo que a Hipótese 1 é satisfeita com $\hat{\ell}_1(\cdot, 0, \cdot) \equiv 0$, se $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é um minimizador para (3.5) satisfazendo $\bar{\beta}(\bar{S}) < K$, então, para cada cone aproximante de Boltjanskij Γ para o alvo $\mathcal{T}$ em $(\bar{y}_0, \bar{y})(\bar{S})$, existem $p_0 \in \mathbb{R}$, $p \in \mathcal{AC}([0, \bar{S}]; \mathbb{R}^n)$ e $\lambda \in [0, \infty[$ satisfazendo todas as condições do Teorema 71 e verificando*

$$0 = \langle p(s), (g_i \circ \bar{y})(s) \rangle \quad \forall s \in [0, \bar{S}], \quad i \in \{1 : m_1\}, \quad (3.18)$$

$$0 = \langle p(s), (B \circ \bar{y})(s) \rangle \quad \forall s \in [0, \bar{S}], \quad B \in \mathfrak{B}^0. \quad (3.19)$$

Corolário 77 (páginas 825 e 826 de [1]). *Assumindo que a Hipótese 1 é satisfeita com $\hat{\ell}_1(\cdot, 0, \cdot) \equiv 0$, se $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é um minimizador para (3.5) satisfazendo $\bar{\beta}(\bar{S}) < K$, então, para cada cone aproximante de Boltjanskij Γ para o alvo $\mathcal{T}$ em $(\bar{y}_0, \bar{y})(\bar{S})$, sejam $p_0 \in \mathbb{R}$, $p \in \mathcal{AC}([0, \bar{S}]; \mathbb{R}^n)$ e $\lambda \in [0, \infty[$ satisfazendo todas as condições do Teorema 76. Sob essas condições, para cada $B \in \mathfrak{B}^1 \cup \{g_1, \dots, g_m\}$, vale*

$$\left\langle p; [f_{\bar{\alpha}}, B](\bar{y})\bar{w}_0 + \sum_{j=m_1+1}^m [g_j, B](\bar{y})\bar{w}_j \right\rangle = -\lambda \left\langle \frac{\partial \ell^e}{\partial x}(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}); B(\bar{y}) \right\rangle \quad (3.20)$$

quase sempre em $[0, \bar{S}]$. Em particular, se $m_1 = m$ e a condição

$$\lambda \left\langle \frac{\partial \ell^e}{\partial x}(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}); B(\bar{y}) \right\rangle = 0 \quad (3.21)$$

é satisfeita quase sempre em $[0, \bar{S}]$, então,

$$\langle p; [f_{\bar{\alpha}}, B](\bar{y})\bar{w}_0 \rangle = 0 \quad (3.22)$$

quase sempre em $[0, \bar{S}]$.

Demonstração. Em primeiro lugar, fixando $i \in \{1 : m_1\}$, derivar, formalmente, a correspondente igualdade em (3.19) em relação a t fornece

$$\begin{aligned}
0 &= \langle \dot{p}, B(\bar{y}) \rangle + \left\langle p, J_x B(\bar{y}) \frac{d\bar{y}}{dt} \right\rangle \\
&= - \left\langle p, J_x f_{\bar{\alpha}}(\bar{y}) \bar{w}_0 + \sum_{i=1}^m g_i(\bar{y}) \bar{w}_i \right\rangle + \lambda \langle \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}); B(\bar{y}) \rangle \\
&\quad + \left\langle p; J_x B(\bar{y}) \left[f_{\bar{\alpha}}(\bar{y}) \bar{w}_0 + \sum_{i=1}^m g_i(\bar{y}) \bar{w}_i \right] \right\rangle \\
\Leftrightarrow -\lambda \langle \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}); B(\bar{y}) \rangle &= \langle p, J_x f_{\bar{\alpha}}(\bar{y}) B(\bar{y}) - J_x B(\bar{y}) f_{\bar{\alpha}}(\bar{y}) \rangle \bar{w}_0 \\
&\quad + \left\langle p; \sum_{i=1}^m [J_x B(\bar{y}) g_i(\bar{y}) + J_x g_i(\bar{y}) B(\bar{y})] \bar{w}_i \right\rangle \\
-\lambda \langle \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}); B(\bar{y}) \rangle &= \left\langle p; [f_{\bar{\alpha}}, B](\bar{y}) \bar{w}_0 + \sum_{i=1}^m [g_i, B] \bar{w}_i \right\rangle
\end{aligned}$$

quase sempre em $[0, \bar{5}]$. Dito isso, se $B \in \mathfrak{B}^1$, segue que $[g_i, B] \in \mathfrak{B}^0$ para $i \in \{1 : m_1\}$, logo, por (3.19), $\langle p; [g_i, B] \rangle = 0$ para $i \in \{1 : m_1\}$, o que mostra (3.20) para esse caso. Senão, para $B \in \{g_1, \dots, g_m\}$, (3.18) garante que as contas valem ainda assim, o que prova (3.20) para esse caso.

Finalmente, (3.22) segue de uma simples substituição de (3.21) em (3.20). $\square$

Observação 78 (página 826 de [1]). A condição (3.21) pode ser verificada se $\lambda = 0$ (isto é, no caso anormal) ou se ℓ_0 e ℓ_1 são independentes de x e $\hat{\ell}_1(0, \cdot) \equiv 0$ (o problema de tempo mínimo é um exemplo de onde isso se aplica).

3.3.1 Sistemas lineares

Tome C e E matrizes reais de ordem $n \times n$ e $n \times m$ respectivamente, e considere o sistema linear $\dot{x} = Cx + Eu$.

Pondo $f(x) := Cx$ e g_i como o campo de vetores constante e igual à i^{a} coluna de E , como $[g_i, g_j] = E_i \cdot E_j - E_j \cdot E_i = 0$ para todo $i \in \{1 : m\}$ e $j \in \{1 : n\}$, todos os colchetes de Lie de qualquer ordem se anulam. Mais ainda, (3.18) é reduzida a

$$pE \equiv 0, \tag{3.23}$$

enquanto, devido a (3.21), a equação adjunta é lida $\dot{p} = -pC$, o que, diferenciando (3.23) $n - 1$ vezes, fornece

$$pC^j E = 0 \quad \forall j \in \{1 : n - 1\}. \tag{3.24}$$

3.3.2 Processos, totalmente, impulsivos

O objetivo dessa subseção é destacar as condições necessárias/suficientes para que um processo ótimo $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ seja, *totalmente, impulsivo*, isto é, evolua de modo instantâneo, o que significa que $\bar{y}_0(\bar{S}) = 0$ e, conseqüentemente, $\bar{w}_0 = 0$ quase sempre em $[0, \bar{S}]$.

Para tal objetivo, duas condições de posto são dadas:

1. Condições de posto $\mathcal{C}^0$ -pontuais em $x \in \mathbb{R}^n$:

(1.1)_x Existem $r \in \mathbb{N}$ e colchetes de Lie iterados $B_1, \dots, B_r \in \mathfrak{B}^0$ satisfazendo

$$\text{span}\{B_1(x), \dots, B_r(x), g_1(x), \dots, g_{m_1}(x)\} = \mathbb{R}^n; \quad (3.25)$$

(1.2)_x Para cada $a \in A$, existem $r, k \in \mathbb{N}$ e colchetes de Lie iterados $B_1, \dots, B_r \in \mathfrak{B}^0$ e $\hat{B}_1, \dots, \hat{B}_k \in \mathfrak{B}^1$ satisfazendo

$$\begin{aligned} \text{span}\{B_1(x), \dots, B_r(x), [f_a, \hat{B}_1](x), \dots, [f_a, \hat{B}_k](x), g_1(x), \dots, g_{m_1}(x), \\ [f_a, g_1](x), \dots, [f_a, g_{m_1}](x)\} = \mathbb{R}^n. \end{aligned} \quad (3.26)$$

2. Condição de Controlabilidade de Kalman: o sistema é linear da forma $\dot{x} = Cx + Eu$ e sua matriz de controlabilidade $[E \quad CE \quad \dots \quad C^{n-1}E]$ possui posto n .

Hipótese 2 (página 827 de [1]). Além dos itens da **Hipótese 1**, suponha que

6. O alvo é invariante por tempo, isto é, da forma $\mathcal{T} = \mathbb{R} \times \hat{\mathcal{T}}$ com $\hat{\mathcal{T}} \subseteq \mathbb{R}^n$;
7. O custo final Ψ é independente do tempo;
8. O lagrangiano ℓ é, estritamente, positivo e $\hat{\ell}_1(\cdot, 0, \cdot) \equiv 0$.

Lema 79 (página 827 de [1]). *Assumindo 1 a 8 na Hipótese 2 e sendo $\pi = 0$, para cada subconjunto $\mathcal{J} \subseteq [0, \bar{S}]$ de medida positiva, não vale*

$$p = 0 \text{ e } \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}) > 0 \text{ quase sempre em } \mathcal{J} \quad (3.27)$$

nem

$$p = 0 \text{ e } \frac{\partial \ell^e}{\partial x}(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}) \neq 0 \text{ quase sempre em } \mathcal{J}. \quad (3.28)$$

Demonstração. Do item 7 da **Hipótese 2**, é obtido que $\Gamma = \mathbb{R} \times \hat{\Gamma}$ com $\hat{\Gamma}$ um cone de $\mathbb{R}^n$, assim, do item 8 desta e da identidade $\Gamma^\perp = \{0\} \times \hat{\Gamma}^\perp$, a condição de não-transversalidade fornece $p_0 = 0$.

Assumindo que (3.27) valha em um conjunto $\mathcal{J} \subseteq [0, \bar{S}]$ de medida positiva, como $(p_0, p(s), \pi) = (0, 0, 0)$ para todo $s \in \mathcal{J}$, a igualdade (3.13) fornece $\lambda \ell^e \circ (\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})(s) = 0$ para quase todo $s \in \mathcal{J}$, assim, por (3.27), $\lambda = 0$.

Agora, assumindo a validade de (3.28) em um conjunto $\mathcal{J} \subseteq [0, \bar{S}]$ de medida positiva, como $(p_0, p(s), \pi) = (0, 0, 0)$ para todo $s \in \mathcal{J}$, a igualdade (3.11) fornece $\lambda \frac{\partial \ell^e}{\partial x}(\bar{y}(s), \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s)) = 0$ para quase todo $s \in \mathcal{J}$, assim, por (3.28), $\lambda = 0$.

Escolhendo, agora, $\hat{s} \in \mathcal{I}$ de modo que $p(\hat{s}) = 0$, como, em qualquer caso, $\lambda = 0$, a equação adjunta é linear em p , o que mostra que $p \equiv 0$ em $[0, \bar{S}]$. Sendo assim, $(p_0, p, \pi, \lambda) = (0, 0, 0, 0)$, o que contradiz a condição de não-trivialidade. $\square$

Teorema 80 (página 827 de [1]). *Assumindo a **Hipótese 2**, sejam $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ um minimizador local canônico para (3.5) satisfazendo $\bar{\beta}(\bar{S}) < K$ e (p_0, p, λ) um multiplicador como no **Teorema 76**. Sob tais condições, se uma das condições abaixo é satisfeita, então, $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ é, totalmente, impulsivo:*

1. Para cada $s \in [0, \bar{S}]$, $(1.1)_{\bar{y}(s)}$ é satisfeita;
2. Para cada $s \in [0, \bar{S}]$, $(1.2)_{\bar{y}(s)}$ é satisfeita e

$$J := \{s \in [0, \bar{S}] \mid (1.1)_{\bar{y}(s)} \text{ não vale}\} \neq \emptyset.$$

Mais ainda, $m_1 = m$ e $\lambda \frac{\partial \ell^e}{\partial x}(\bar{y}(s), \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s)) = 0$ para quase todo $s \in J$;

3. O sistema é linear, a Condição de Controlabilidade de Kalman é satisfeita e $\lambda \frac{\partial \ell^e}{\partial x} \circ (\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})(s) = 0$ para quase todo $s \in [0, \bar{S}]$.

Demonstração. O primeiro passo é notar que, como $\bar{\beta}(\bar{S}) < K$, $\pi = 0$.

Supondo que o item 1 da hipótese seja verificado, para cada $s \in [0, \bar{S}]$, por $(1.1)_{\bar{y}(s)}$, existem um natural r e colchetes de Lie iterados $B_1, \dots, B_r \in \mathfrak{B}^0$ satisfazendo (3.25) e, em vista de (3.18) e (3.19), qualquer $s \in [0, \bar{S}]$ satisfaz

$$\begin{aligned} \langle p(s), (g_i \circ \bar{y})(s) \rangle &= 0, \\ \langle p(s), (B_j \circ \bar{y})(s) \rangle &= 0 \end{aligned}$$

para $i \in \{1 : m_1\}$ e $j \in \{1 : r\}$, o que prova que $p \equiv 0$ em $[0, \bar{S}]$.

Assuma, por absurdo, que existe um subconjunto de medida positiva $\mathcal{I} \subseteq [0, \bar{S}]$ no qual $\bar{w}_0 > 0$ quase sempre. Da 1-homogeneidade da função ℓ em relação a w_0 , isto implica $\ell^e(\bar{y}_0, \bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}) > 0$ quase sempre em $\mathcal{I}$, o que contradiz o **Lema 79**.

Agora, assumindo a validade do item 2 da hipótese, se $s \in [0, \bar{S}] \setminus J$, o raciocínio acima garante que $p(s) = 0$. Senão, se J é um conjunto nulo, o mesmo vale pela continuidade de p . Caso contrário, para quase todo $s \in J$ e $a := \bar{\alpha}(s)$, por $(1.2)_{\bar{y}(s)}$, existem naturais r, k e colchetes de Lie iterados $B_1, \dots, B_r \in \mathfrak{B}^0$ e $\hat{B}_1, \dots, \hat{B}_k \in \mathfrak{B}^1$ satisfazendo (3.26), assim, para quase todo $s \in J$, (3.18), (3.19) e (3.22) garantem que

$$\begin{aligned} \langle p(s), (g_i \circ \bar{y})(s) \rangle &= 0, & \langle p(s), (B_j \circ \bar{y})(s) \rangle &= 0, \\ \langle p(s); [f_{\bar{\alpha}(s)}, g_i] \rangle(\bar{y}(s)) &= 0, & \langle p(s); f_{\bar{\alpha}(s)}, \hat{B}_\ell \rangle(\bar{y}(s)) &= 0 \end{aligned}$$

para todos $i \in \{1 : m\}$, $j \in \{1 : r\}$ e $\ell \in \{1 : k\}$. Isso mostra que $p \equiv 0$ quase sempre em J , assim, a continuidade de p garante que isso vale sempre em $[0, \bar{S}]$. Novamente, assumindo a existência de um subconjunto de medida positiva $\mathcal{I} \subseteq [0, \bar{S}]$ no qual $\bar{w}_0 > 0$ quase sempre, a argumentação segue como acima.

Finalmente, assumindo a validade do item 3 da hipótese, (3.23) e (3.24) garantem que $p \equiv 0$ em $[0, \bar{S}]$, o que permite continuar a argumentação como nos casos anteriores. $\square$

3.3.3 Processos não impulsivos

Quando o processo $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ não é impulsivo, ou seja, $\bar{w}_0 > 0$ quase sempre em $[0, \bar{S}]$, o Teorema 76 pode ser reescrito em função do parâmetro temporal t introduzindo a *hamiltoniana não-maximizada* para (3.2) $\mathcal{H} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times C \times A \rightarrow \mathbb{R}$ dada por

$$\mathcal{H}(x, \mathfrak{p}, \lambda, u, a) = \left\langle \mathfrak{p}; f(x, a) + \sum_{i=1}^m g_i(x) u_i \right\rangle - \lambda \ell(x, u, a).$$

Teorema 81 (página 829 de [1]). *Assumindo que a Hipótese 1 é satisfeita com $\hat{\ell}_1(\cdot, 0, \cdot) \equiv 0$, sendo $(\bar{S}, \bar{w}_0, \bar{w}, \bar{\alpha}, \bar{y}_0, \bar{y}, \bar{\beta})$ um minimizador espaço-temporal local não-impulsivo para (3.5), isto é, que verifica $\bar{w}_0 > 0$ quase sempre em $[0, \bar{S}]$, suponha $\bar{\beta}(\bar{S}) < K$.*

Sob tais condições, se $(\bar{T}, \bar{u}, \bar{a}, \bar{x}, \bar{v})$ é o minimizador sentido-estrito para (3.2), então, para cada cone aproximante de Boltjanskij Γ para o alvo $\mathcal{T}$ em $(\bar{T}, \bar{x}(\bar{T}))$, existem $p_0 \in \mathbb{R}$, $\mathfrak{p} \in \mathcal{AC}([0, \bar{T}]; \mathbb{R}^n)$ e $\lambda \geq 0$ satisfazendo as seguintes condições:

- (Não-trivialidade) *Nem todos entre $\mathfrak{p}$ e λ são nulos;*
- (Não-transversalidade)

$$(p_0, \mathfrak{p}(\bar{T})) \in [-\lambda \nabla \Psi(\bar{T}, \bar{x}(\bar{T})) - \Gamma^\perp];$$

- (Equação adjunta)

$$\dot{\mathfrak{p}}(t) = -\frac{\partial \mathcal{H}}{\partial x}(\bar{x}(t), \mathfrak{p}(t), \lambda, \bar{u}(t), \bar{a}(t)) \quad \forall t \in [0, \bar{T}];$$

- (Maximização de primeira ordem)

$$\begin{aligned} \mathcal{H}(\bar{x}(t), \mathfrak{p}(t), \lambda, \bar{u}(t), \bar{a}(t)) &= \max \{ \mathcal{H}(\bar{x}(t), \mathfrak{p}(t), \lambda, u, a) \mid u \in C, a \in A \} \\ &= -p_0 \quad \forall t \in [0, \bar{T}]; \end{aligned}$$

- (Anulamento do produto)

$$\begin{aligned} 0 &= \langle \mathfrak{p}(t); (g_i \circ \bar{x})(t) \rangle \quad \forall t \in [0, \bar{T}], \quad i \in \{1 : m_1\}, \\ 0 &= \langle \mathfrak{p}(t); (B \circ \bar{x})(t) \rangle \quad \forall t \in [0, \bar{T}], \quad B \in \mathfrak{B}^0. \end{aligned}$$

Ainda, se T é livre, então, $p_0 = 0$.

Demonstração. Definindo $\bar{\varsigma}$ como em (3.8) e $\mathfrak{p} = p \circ \bar{\varsigma}$ e aplicando o Teorema 76:

- Como $\bar{\varsigma}$ é bilipschitziana e crescente, $\mathfrak{p}$ é nula se, e somente se, p também o faz, o que mostra a não-trivialidade;
-

$$\begin{aligned}\mathfrak{p}(\bar{T}) &= (p \circ \bar{\varsigma})(\bar{T}) = p(\bar{S}), \\ \bar{T} &= \bar{y}_0(\bar{S}) \\ \bar{x}(\bar{T}) &= (\bar{y} \circ \bar{\varsigma})(\bar{T}) = \bar{y}(\bar{S}),\end{aligned}$$

o que mostra a não-transversalidade;

- Sabendo que H é 1-homogênea em relação a w_0 e w , quase todo $t \in [0, \bar{T}]$ satisfaz

$$\begin{aligned}\dot{\mathfrak{p}}(t) &= (\dot{p} \circ \bar{\varsigma})(t) \frac{d\bar{\varsigma}}{dt} = -\frac{\partial H}{\partial x}(\bar{y}, p, 0, \lambda, \bar{w}_0, \bar{w}, \bar{\alpha}) \circ \bar{\varsigma}(t) \frac{d\bar{\varsigma}}{dt} \\ &= -\frac{\partial H}{\partial x}(\bar{x}(t), \mathfrak{p}(t), 0, \lambda, 1, \bar{u}(t), \bar{a}(t)) = -\frac{\partial \mathcal{H}}{\partial x}(\bar{x}(t), \mathfrak{p}(t), \lambda, \bar{u}(t), \bar{a}(t)) \dot{\forall} t \in [0, \bar{T}],\end{aligned}$$

o que mostra a validade da equação adjunta;

- Sabendo que $\mathcal{H}(x, \mathfrak{p}, \lambda, u, a) = H(x, p_0, \mathfrak{p}, 0, \lambda, 1, \bar{u}, \bar{a}) - p_0$ para quaisquer $x, \mathfrak{p} \in \mathbb{R}^n, \lambda \in \mathbb{R}, u \in C$ e $a \in A$,

$$\begin{aligned}\mathcal{H}(\bar{x}(t), \mathfrak{p}(t), \lambda, \bar{u}(t), \bar{a}(t)) &= H(\bar{x}(t), p_0, \mathfrak{p}(t), 0, \lambda, 1, \bar{u}(t), \bar{a}(t)) - p_0 \\ &= \max \{H(\bar{x}(t), p_0, \mathfrak{p}(t), 0, \lambda, 1, u, a) \mid u \in C, a \in A\} - p_0 \\ &= -p_0 = \max \{\mathcal{H}(\bar{x}(t), p_0, \mathfrak{p}(t), \lambda, u, a) \mid u \in C, a \in A\} \\ &\dot{\forall} t \in [0, \bar{T}];\end{aligned}$$

•

$$\begin{aligned}\langle \mathfrak{p}(t), (g_i \circ \bar{x})(t) \rangle &= \langle p, g_i(\bar{y}) \rangle \circ \bar{\varsigma}(s) \quad \forall t \in [0, \bar{T}], \quad i \in \{1 : m_1\}, \\ \langle \mathfrak{p}(t), (B \circ \bar{x})(t) \rangle &= \langle p, B(\bar{y}) \rangle \circ \bar{\varsigma}(s) \quad \forall t \in [0, \bar{T}], \quad B \in \mathfrak{B}^0,\end{aligned}$$

o que mostra, devido à bijetividade de ς , o anulamento do produto.

□

3.4 Exemplo de aplicação

O exemplo a seguir, retirado de [1], mostra como condições necessárias de alta ordem podem ser usada para filtrar extremais obtido com condições de primeira ordem.

Defina funções $g_1, g_2, g_3 : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ por

$$g_1(x) = e_1 - \frac{x_2}{2}e_3,$$

$$g_2(x) = e_2 + \frac{x_1}{2}e_3,$$

$$g_3(x) = x_3e_4,$$

e considere o problema de controle ótimo:

$$\left\{ \begin{array}{l} \text{minimizar } x_4(1) + \int_0^1 [x_3^2(t) + x_4^2(t)] dt \\ \text{sujeito a } u : [0, 1] \rightarrow \mathbb{R}^3 \text{ mensurável e } x \in \mathcal{AC}([0, 1]; \mathbb{R}^4) \text{ satisfazendo} \\ \dot{x} = g_1(x)u_1 + g_2(x)u_2 + g_3(x)u_3 \text{ quase sempre em } [0, 1], \\ x(0) = e_1, \\ x_1(1) = x_2(1) = 0 \text{ e } x_4(1) \geq -\frac{1}{2}. \end{array} \right. \quad (3.29)$$

Adicionando a equação $\dot{x}_5 = x_3^2 + x_4^2$ e redefinindo $x = (x_1, \dots, x_5)$ e definindo $\hat{g}_1, \hat{g}_2, \hat{g}_3 : \mathbb{R}^5 \rightarrow \mathbb{R}^5$ adicionando uma componente zero a g_1, g_2 e g_3 respectivamente, o correspondente problema espaço-temporal se torna

$$\left\{ \begin{array}{l} \text{minimizar } y_4(S) + y_5(S) \\ \text{sujeito a } S > 0, w_0, y_0 : [0, S] \rightarrow [0, \infty[, w : [0, S] \rightarrow \mathbb{R}^3 \text{ e } y \in [0, S] \rightarrow \mathbb{R}^5 \\ \text{satisfazendo} \\ \dot{y}_0 = w_0 \text{ e } \dot{y} = w_0 (y_3^2 + y_4^2) e_5 + \hat{g}_1(y)w_1 + \hat{g}_2(y)w_2 + \hat{g}_3(y)w_3 \\ \text{quase sempre em } [0, S], \\ y_0(0) = 0, y(0) = e_1, \\ y_0(S) = 1, y_1(S) = y_2(S) = 0 \text{ e } y_4(S) \geq -\frac{1}{2}. \end{array} \right. \quad (3.30)$$

Considerando o processo $(2, \widehat{w}_0, \widehat{w}, \widehat{y}_0, \widehat{y})$ onde

$$\widehat{w}_0(s) = \frac{1}{2} \quad \dot{y}s \in [0, 2]$$

$$\widehat{y}_0(s) = \frac{s}{2} \quad \dot{y}s \in [0, 2]$$

$$\widehat{w}(s) = -\frac{1}{2}e_1 \quad \dot{y}s \in [0, 2],$$

$$\widehat{y}(s) = \left(1 - \frac{s}{2}\right)e_1 \quad \dot{y}s \in [0, 2],$$

e o custo final ligado a ele é, claramente, zero.

Afirmção. Tal processo é um extremal para (3.30).

Claramente, escolhendo $\Gamma := \mathbb{R}^3 \times \{(0, 0, 0)\}$ como cone aproximante de Boltjanskij para o alvo $\{1\} \times \mathcal{T} \times \mathbb{R}$ no ponto final $(\hat{y}_0, \hat{y})(2)$, como

$$\begin{aligned}
 H(y, p_0, p, w_0, w) &= p_0 w_0 \\
 &+ p \left[\frac{1}{2} (y_3^2 + y_4^2) e_5 - \frac{1}{2} \left(e_1 - \frac{y_2}{2} e_3 \right) w_1 + \left(e_2 + \frac{y_1}{2} e_3 \right) w_2 + y_3 e_4 w_3 \right] \\
 &= p_0 w_0 + w_0 (y_3^2 + y_4^2) p_5 + \left(p_1 - \frac{y_2}{2} p_3 \right) w_1 + \left(p_2 + \frac{y_1}{2} p_3 \right) w_2 \\
 &+ y_3 p_4 w_3 \\
 &= (p_0 w_0 + p_1 w_1 + p_2 w_2) + \frac{p_3 w_2}{2} y_1 - \frac{p_3 w_1}{2} y_2 + p_4 w_3 y_3 + p_5 w_0 y_3^2 \\
 &+ p_5 w_0 y_4^2,
 \end{aligned}$$

é possível reescrever a condição de não-transversalidade e a equação adjunta como

$$\begin{cases} \dot{p}_0 = 0, \\ \dot{p}_1 = -\frac{p_3 \widehat{w}_2}{2} = 0, \\ \dot{p}_2 = \frac{p_3 \widehat{w}_1}{2} = -\frac{p_3}{4}, \\ \dot{p}_3 = -p_4 \widehat{w}_3 - 2p_5 \widehat{w}_0 \hat{y}_3 = 0, \\ \dot{p}_4 = -2p_5 \hat{y}_4 \widehat{w}_0 = 0, \\ \dot{p}_5 = 0, \end{cases} \quad \forall s \in [0, \bar{S}]$$

$$(p_0, p(2)) \in -\lambda(e_5 + e_6) - \mathbb{R}^3 \times \{(0, 0, 0)\}$$

com $\lambda \geq 0$, o que mostra que p_0, p_1 e p_2 são constantes reais arbitrárias, $p_3 \equiv 0$ e $p_4 = p_5 \equiv -\lambda$. Mais ainda, (3.12) e (3.13) garantem que

$$\begin{aligned}
 0 &= p_0 \widehat{w}_0 + p_1 \widehat{w}_1 + p_2 \widehat{w}_2 = \frac{p_0 - p_1}{2}, \\
 0 &\geq p_0 w_0 + p_1 w_1 + p_2 w_2
 \end{aligned}$$

para quaisquer $w_0 \geq 0$ e $w \in \mathbb{R}^3$ satisfazendo $w_0 + |w| = 1$. Em particular:

- Escolhendo $w_0 = 1$ e $w = 0$, obtemos $p_0 = 0$;
- Escolhendo $w_0 = 0$ e $w = \pm e_1$, obtemos $p_1 = 0$;
- Escolhendo $w_0 = 0$ e $w = \pm e_2$, obtemos $p_2 = 0$.

Para qualquer $\lambda \geq 0$, o processo dado e o multiplicador $(p_0, p, \lambda) = \lambda(e_6 - e_5 - e_4)$ satisfaz as condições do Teorema 71.

Entretanto, sabendo que

$$\begin{aligned} [\hat{g}_1, \hat{g}_2] &= \begin{bmatrix} \frac{e_3}{2} & 0 & 0 & 0 & 0 \end{bmatrix} \left(e_1 - \frac{x_2}{2} e_3 \right) - \begin{bmatrix} 0 & -\frac{e_3}{2} & 0 & 0 & 0 \end{bmatrix} \left(e_2 + \frac{x_1}{2} e_3 \right) \\ &= \frac{e_3}{2} + \frac{e_3}{2} = e_3 \end{aligned}$$

e, consequentemente,

$$[[\hat{g}_1, \hat{g}_2], \hat{g}_3] = \begin{bmatrix} 0 & 0 & e_4 & 0 & 0 \end{bmatrix} e_3 = e_4,$$

a condição (3.19) garante que $p_4 = 0$, o que prova que $\lambda = 0$ e, consequentemente, $(p_0, p, \lambda) = (0, 0, 0)$. Portanto, não existe um multiplicador para o qual todas as condições do Teorema 76 são satisfeitas, o que mostra que $(2, \widehat{w}_0, \widehat{w}, \widehat{y}_0, \widehat{y})$ não é um minimizador.

Por outro lado, pondo $\bar{S} = 2 + \sqrt{1 + 4\pi}$ e $\varrho = \frac{1}{\sqrt{1+4\pi}}$, o controle espaço-temporal $(\bar{S}, \bar{w}_0, \bar{w})$ definido por

$$(\bar{w}_0, \bar{w})(s) = \begin{cases} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \end{bmatrix}^T & \text{para } s \in [0, 2[, \\ \varrho \begin{bmatrix} 0 & 2\sqrt{\pi} \sin(2\pi\varrho(s-2)) & 2\sqrt{\pi} \cos(2\pi\varrho(s-2)) & 1 \end{bmatrix}^T & \text{para } s \in [2, \bar{S}] \end{cases}$$

é ótimo. De fato, as equações diferenciais em (3.30) são dadas por

$$\begin{aligned} \frac{d\bar{y}_0}{ds} &= \begin{cases} \frac{1}{2} & \text{se } s \in [0, 2], \\ 0 & \text{se } s \in [2, \bar{S}], \end{cases} \\ \frac{d\bar{y}}{ds} &= \begin{cases} \begin{bmatrix} -\frac{1}{2} & 0 & \frac{\bar{y}_2}{2} & 0 & \frac{\bar{y}_3^2 + \bar{y}_4^2}{2} \end{bmatrix}^T & \text{se } s \in [0, 2], \\ \begin{bmatrix} 2\sqrt{\pi}\varrho \sin(2\pi\varrho(s-2)) \\ 2\sqrt{\pi}\varrho \cos(2\pi\varrho(s-2)) \\ \bar{y}_1\sqrt{\pi}\varrho \cos(2\pi\varrho(s-2)) - \bar{y}_2\sqrt{\pi}\varrho \sin(2\pi\varrho(s-2)) \\ \bar{y}_3\varrho \\ 0 \end{bmatrix} & \text{se } s \in [2, \bar{S}], \end{cases} \end{aligned}$$

assim:

$$\begin{aligned}
s \in [0, 2] &\Rightarrow \begin{cases} \bar{y}_0(s) = \frac{s}{2}, \\ \bar{y}_1(s) = 1 - \frac{s}{2}, \\ \bar{y}_2(s) = 0, \\ \bar{y}_3(s) = 0, \\ \bar{y}_4(s) = 0, \\ \bar{y}_5(s) = 0, \end{cases} \\
s \in [2, \bar{S}] &\Rightarrow \begin{cases} \bar{y}_0(s) = 1, \\ \bar{y}_1(s) = \frac{1 - \cos(2\pi\varrho(s-2))}{\sqrt{\pi}}, \\ \bar{y}_2(s) = \frac{\sin(2\pi\varrho(s-2))}{\sqrt{\pi}}, \\ \bar{y}_3(s) = \frac{\sin(2\pi\varrho(s-2))}{2\pi} + \varrho(2-s), \\ \bar{y}_4(s) = \frac{1 - \cos(2\pi\varrho(s-2))}{4\pi^2} + \varrho^2\left(-2 + 2s - \frac{s^2}{2}\right), \\ \bar{y}_5(s) = 0 \end{cases}
\end{aligned}$$

o que mostra que

$$\begin{aligned}
\bar{y}_0(\bar{S}) &= 1, & \bar{y}_1(\bar{S}) &= 0, \\
\bar{y}_2(\bar{S}) &= 0, & \bar{y}_3(\bar{S}) &= -1, \\
\bar{y}_4(\bar{S}) &= -\frac{1}{2}, & \bar{y}_5(\bar{S}) &= 0,
\end{aligned}$$

e isso significa que tal processo satisfaz as condições de contorno em (3.30). Ainda, o custo final associado é $-\frac{1}{2}$, que é o *menor valor possível*.

Escolhendo $\Gamma := \mathbb{R}^3 \times \{0\} \times]-\infty, 0] \times \{0\}$ como o cone aproximante de Boltjanskij para o conjunto $\{1\} \times \mathcal{T} \times \mathbb{R}$ no ponto $(\hat{y}_0, \hat{y})(2)$, a condição de não-transversalidade é lida

$$(p_0, p(\bar{S})) \in -\lambda(e_5 + e_6) - \mathbb{R}^3 \times \{0\} \times]-\infty, 0] \times \{0\}$$

com $\lambda \geq 0$, o que fornece

$$(p_0, p(\bar{S})) = [c_0 \quad c_1 \quad c_2 \quad 0 \quad c_4 - \lambda \quad -\lambda]^T$$

com $c_0, c_1, c_2 \in \mathbb{R}$ e $c_4 \geq 0$. Escolhendo $c_0 = c_1 = c_2 = 0$, $c_4 = 1$ e $\lambda = 1$, é obtido o multiplicador não-trivial $(p_0, p, \lambda) \equiv e_7 - e_6$, que satisfaz as condições do Teorema 76. Em particular, a condição (3.19) é satisfeita pelo fato de que a quinta componente dos campos de vetores e dos colchetes de Lie gerados por eles se anulam.

3.5 Demonstrando a versão de alta ordem

Nesta seção, demonstramos o **Teorema 76** utilizando o raciocínio de [1].

Nas hipóteses do **Teorema 76**, o primeiro passo será reescalonar o problema (3.5) para fixar o instante final. Paralelo a isso, vamos usar, também, a notação mais compacta

$$F^e(x, w_0, w, a) = f(x, a)w_0 + \sum_{i=1}^m g_i(x)w_i \quad \forall (x, w_0, w, a) \in \mathbb{R}^n \times [0, \infty[\times C \times A,$$

$$\bar{F}^e(s) = F^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})(s), \quad \bar{\ell}^e(s) = \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})(s) \quad \forall s \in [0, \bar{S}].$$

Definição 82. A seção S de $\mathcal{W}$ relativa a $\bar{S}$ é denominada **conjunto de controle espaço-temporal reescalado** e denotada por $\mathcal{W}_{\bar{S}}$.

Definição 83 (página 832 de [1]). Fixando $\varrho > 0$, para cada $(w_0, w, \alpha) \in \mathcal{W}_{\bar{S}}$, $\zeta \in L^\infty([0, \bar{S}]; [-\varrho, \varrho])$ e (y_0, y, y_{n+1}, β) uma solução de Karatheodori de

$$\begin{cases} \dot{y}_0 = w_0(1 + \zeta), \\ \dot{y} = F^e(y, w_0, w, \alpha)(1 + \zeta), \\ \dot{y}_{n+1} = \ell^e(y, w_0, w, \alpha)(1 + \zeta), \\ \dot{\beta} = |w|(1 + \zeta), \\ (y_0, y, y_{n+1}, \beta) = (0, \check{x}, 0, 0), \end{cases} \quad (3.31)$$

o processo $(w_0, w, \alpha, \zeta, y_0, y, y_{n+1}, \beta)$ é denominado **processo espaço-temporal reescalado**, e é factível se $(y_0, y)(\bar{S}) \in \mathcal{T}$ e $\beta(\bar{S}) \in [0, K]$.

A partir de tal construção, é possível definir o *problema espaço-temporal reescalado* por

$$\begin{cases} \text{minimizar} & \Psi(y_0(\bar{S}), y(\bar{S})) + y_{n+1}(\bar{S}) \\ \text{sobre processos reescalados factíveis} & (w_0, w, \alpha, \zeta, y_0, y, y_{n+1}, \beta). \end{cases} \quad (3.32)$$

Definição 84. A solução $(\bar{w}_0, \bar{w}, \bar{\alpha}, 0, \bar{y}_0, \bar{y}, \bar{y}_{n+1}, \bar{\beta})$ é denominada **processo espaço-temporal reescalado referencial**.

Como, para todo $s \in [0, \bar{S}]$,

$$\begin{aligned} y_0(s) - \bar{y}_0(s) &= \int_0^s \{w_0(\tau) [1 + \zeta(\tau)] - \bar{w}_0(\tau)\} d\tau \\ &= \int_0^s [1 + \zeta(\tau)] [w_0(\tau) - \bar{w}_0(\tau)] d\tau + \int_0^s \zeta(\tau) \bar{w}_0(\tau) d\tau, \end{aligned}$$

$$\begin{aligned}
y(s) - \bar{y}(s) &= \int_0^s \{ [F^e \circ (y, w_0, w, \alpha)](\tau) [1 + \zeta(\tau)] - F^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})(\tau) \} d\tau \\
&= \int_0^s [1 + \zeta(\tau)] [F^e(y, w_0, w, \alpha) - F^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})](\tau) d\tau \\
&\quad + \int_0^s \zeta(\tau) F^e(\bar{y}(\tau), \bar{w}_0(\tau), \bar{w}(\tau), \bar{\alpha}(\tau)) d\tau, \\
y_{n+1}(s) - \bar{y}_{n+1}(s) &= \int_0^s \{ [\ell^e \circ (y, w_0, w, \alpha)](\tau) [1 + \zeta(\tau)] - \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})(\tau) \} d\tau \\
&= \int_0^s [1 + \zeta(\tau)] [\ell^e(y, w_0, w, \alpha) - \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})](\tau) d\tau \\
&\quad + \int_0^s \zeta(\tau) \ell^e(\bar{y}(\tau), \bar{w}_0(\tau), \bar{w}(\tau), \bar{\alpha}(\tau)) d\tau, \\
\beta(s) - \bar{\beta}(s) &= \int_0^s \{ |w(\tau)| [1 + \zeta(\tau)] - |\bar{w}(\tau)| \} d\tau \\
&= \int_0^s [1 + \zeta(\tau)] [|w(\tau)| - |\bar{w}(\tau)|] d\tau + \int_0^s \zeta(\tau) |\bar{w}(\tau)| d\tau,
\end{aligned}$$

segue que

$$\begin{aligned}
|y_0(s) - \bar{y}_0(s)| &\leq (1 + \varrho) \int_0^s |w_0(\tau) - \bar{w}_0(\tau)| d\tau + \varrho \int_0^s |\bar{w}_0(\tau)| d\tau, \\
|y(s) - \bar{y}(s)| &\leq (1 + \varrho) \int_0^s |[F^e(y, w_0, w, \alpha) - F^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})](\tau)| d\tau \\
&\quad + \varrho \int_0^s |F^e(\bar{y}(\tau), \bar{w}_0(\tau), \bar{w}(\tau), \bar{\alpha}(\tau))| d\tau, \\
|y_{n+1}(s) - \bar{y}_{n+1}(s)| &\leq (1 + \varrho) \int_0^s |[\ell^e(y, w_0, w, \alpha) - \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})](\tau)| d\tau \\
&\quad + \varrho \int_0^s |\ell^e(\bar{y}(\tau), \bar{w}_0(\tau), \bar{w}(\tau), \bar{\alpha}(\tau))| d\tau, \\
|\beta(s) - \bar{\beta}(s)| &\leq (1 + \varrho) \int_0^s ||w(\tau)| - |\bar{w}(\tau)|| d\tau + \varrho \int_0^s |\bar{w}(\tau)| d\tau,
\end{aligned}$$

o que mostra que, para $\varrho > 0$ pequeno o suficiente, o processo referencial é um minimizador local do problema (3.32), assim, como a demonstração considerará, somente, as trajetórias próximas à referencial $(\bar{y}_0, \bar{y})$, e os controles assumirão valores em um conjunto compacto, utilizando argumentos de truncamento e sequências de *mollifiers*, a seguinte hipótese será feita:

Hipótese 3 (página 832 de [1]). Todas as afirmações da **Hipótese 1** são válidas, a continuidade de ℓ^e, f , de todas as g_i s, das derivadas parciais de todas essas funções em x e de todos os colchetes de Lie iterados $B \in \mathfrak{B}^0$ são uniformes, e todas as funções citadas são limitadas.

3.5.1 Aproximações das variáveis de controle

Definição 85 (página 833 de [1]). O conjunto $W \times A \times [-\varrho, \varrho]$ será chamado **conjunto gerador de variações de agulha**, e o conjunto $\mathfrak{B}^0$ será chamado **conjunto gerador de variações tipo-colchete**. A união $\mathfrak{D}$ desses dois conjuntos será chamada **conjunto gerador de variações**, e seus elementos serão chamados **geradores de variação** de comprimento $h \in \mathbb{N}$.

A qualquer $\mathbf{g} \in \mathfrak{D}$ e qualquer instante $\bar{s} \in]0, \bar{S}[$, é possível associar uma variação espaço-temporal infinitesimal da trajetória referencial $(\bar{y}_0, \bar{y}, \bar{y}_{n+1}, \bar{\beta})$. Antes disso, porém, vamos denotar o conjunto do pontos de Lebesgue da função $(\bar{w}_0, \bar{F}^e, \bar{\ell}^e, |\bar{w}|)$ no intervalo $]0, \bar{S}[$ por $L_{0, \bar{s}}$.

Variações de agulha e tipo-colchete

Definição 86 (página 833 de [1]). Para cada $\bar{s} \in L_{0, \bar{s}}$ e cada $\mathbf{n} = (w_0, w, a, \zeta) \in W \times A \times [-\varrho, \varrho]$, o vetor a seguir é denominado **variação de agulha**:

$$\begin{bmatrix} \mathbf{v}_{\mathbf{n}, \bar{s}, 0} \\ \mathbf{v}_{\mathbf{n}, \bar{s}} \\ \mathbf{v}_{\mathbf{n}, \bar{s}, n+1} \\ \mathbf{v}_{\mathbf{n}, \bar{s}, \beta} \end{bmatrix} = \begin{bmatrix} w_0(1 + \zeta) - \bar{w}_0(\bar{s}) \\ F^e(\bar{y}(\bar{s}), w_0, w, a)(1 + \zeta) - \bar{F}^e(\bar{s}) \\ \ell^e(\bar{y}(\bar{s}), w_0, w, a)(1 + \zeta) - \bar{\ell}^e(\bar{s}) \\ |w|(1 + \zeta) - |\bar{w}(\bar{s})| \end{bmatrix}. \quad (3.33)$$

Já para cada $\bar{s} \in]0, \bar{S}[$ e $B \in \mathfrak{B}^0$, o vetor a seguir é denominado **variação tipo-colchete**:

$$\begin{bmatrix} \mathbf{v}_{B, \bar{s}, 0} \\ \mathbf{v}_{B, \bar{s}} \\ \mathbf{v}_{B, \bar{s}, n+1} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{(B \circ \bar{y})(\bar{s})}{r(B)} \\ 0 \end{bmatrix}. \quad (3.34)$$

Aproximações de agulha

Definição 87 (página 833 de [1]). Tome $\mathbf{n} = (w_0, w, a, \zeta)$ um gerador de variação de agulha e $\bar{s} \in]0, \bar{S}[$. Para cada $(w'_0, w', \alpha', \zeta') \in L^\infty([0, \bar{S}]; W \times A \times [-\varrho, \varrho])$, a família de funções $\{(w'_0, w', \alpha', \zeta')_{\mathbf{n}, \bar{s}, \epsilon} \mid \epsilon \in [0, \bar{s}]\}$, em que cada elemento é definido por

$$(w'_0, w', \alpha', \zeta')_{\mathbf{n}, \bar{s}, \epsilon}(s) = \begin{cases} (w_0, w, a, \zeta) & \text{se } s \in [\bar{s} - \epsilon, \bar{s}], \\ (w'_0, w', \alpha', \zeta')(s) & \text{caso contrário,} \end{cases} \quad (3.35)$$

é denominada **família de aproximações de agulha** para o controle $(w'_0, w', \alpha', \zeta')$ em $\bar{s}$ correspondente a $\mathbf{n}$.

Para o próximo resultado, denotaremos por

$$F'(y_0, y, y_{n+1}, \beta, w_0, w, \alpha) = (w_0; F^e(y, w_0, w, a); \ell^e(y, w_0, w, a); |w|)$$

e por M' a *solução matricial fundamental* da equação

$$\dot{V} = J_x F'(\bar{y}_0, \bar{y}, \bar{y}_{n+1}, \bar{\beta}, \bar{w}_0, \bar{w}, \bar{\alpha}) V.$$

- Como a primeira e última coluna de $J_x F'$ são garantidas nulas, o $(1, j)^\circ$ e $(n+3, j)^\circ$ elementos de M' são $\delta_{1,j}$ e $\delta_{n+3,j}$ (números de Kronecker) para todo $j \in \{1 : n+3\}$. Mais ainda, o mesmo pode ser dito para o $(j, 1)^\circ$ e $(j, n+3)^\circ$ elementos respectivamente;
- Como a submatriz de $J_x F'$ compreendendo da segunda à $(n+1)^\text{a}$ coluna coincide com $J_x F^e$, segue que, chamando M como solução de

$$\dot{V} = J_x F^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha})V, \quad (3.36)$$

a submatriz de M' compreendendo da segunda à $(n+1)^\text{a}$ coluna coincide com M ;

- Pela Regra da Cadeia, o $(r, n+2)^\circ$ elemento da derivada de M' (que será chamado μ_r brevemente) é o produto interno do gradiente de ℓ^e com a r^a coluna de M para todo $r \in \{1 : n\}$;
- O $(n+2, n+2)^\circ$ elemento de M' é um.

Proposição 88 (página 834 de [1]). *Assumindo $\bar{s} \in L_{0,\bar{s}}$, para qualquer gerador $\mathbf{n} = (w_0, w, a, \zeta) \in W \times A \times [-\varrho, \varrho]$ e qualquer $s \in]\bar{s}, \bar{S}]$, vale*

$$\begin{bmatrix} y_{0,\epsilon}(s) - \bar{y}_0(s) \\ y_\epsilon(s) - \bar{y}(s) \\ y_{n+1,\epsilon}(s) - \bar{y}_{n+1}(s) \\ \beta_\epsilon(s) - \bar{\beta}(s) \end{bmatrix} = \epsilon M'(s, \bar{s}) \begin{bmatrix} \mathbf{v}_{\mathbf{n},\bar{s},0} \\ \mathbf{v}_{\mathbf{n},\bar{s}} \\ \mathbf{v}_{\mathbf{n},\bar{s},n+1} \\ \mathbf{v}_{\mathbf{n},\bar{s},\beta} \end{bmatrix} + o(\epsilon),$$

em que, para cada $\epsilon \in [0, \bar{s}]$, $(y_{0,\epsilon}; y_\epsilon; y_{n+1,\epsilon}; \beta_\epsilon)$ é a solução de (3.31) correspondente à aproximação $(\bar{w}_0, \bar{w}, \bar{\alpha}, 0)_{\mathbf{n},\bar{s},\epsilon}$.

Demonstração. Em primeiro lugar, sabendo das definições de $y_{0,\epsilon}$ e β_ϵ e que $\bar{s} \in L_{0,\bar{s}}$, obtemos

$$\begin{aligned} \left. \frac{dy_{0,\epsilon}}{d\epsilon}(\bar{s}) \right|_{\epsilon=0} &= \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \int_{\bar{s}-\epsilon}^{\bar{s}} [w_0(1+\zeta) - \bar{w}_0(\tau)] d\tau = w_0(1+\zeta) + \frac{1}{\epsilon} \lim_{\epsilon \downarrow 0} \int_{\bar{s}-\epsilon}^{\epsilon} \bar{w}_0(\tau) d\tau \\ &= w_0(1+\zeta) - \bar{w}_0(\bar{s}) = \mathbf{v}_{\mathbf{n},\bar{s},0}, \\ \left. \frac{d\beta_\epsilon}{d\epsilon}(\bar{s}) \right|_{\epsilon=0} &= \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \int_{\bar{s}-\epsilon}^{\bar{s}} [|w|(1+\zeta) - |\bar{w}(\tau)|] d\tau = |w|(1+\zeta) + \frac{1}{\epsilon} \lim_{\epsilon \downarrow 0} |\bar{w}(\tau)| d\tau \\ &= |w|(1+\zeta) - |\bar{w}(\bar{s})| = \mathbf{v}_{\mathbf{n},\bar{s},\beta}. \end{aligned}$$

O fato de que $\bar{s} \in L_{0,\bar{s}}$ garante, também, que

$$\begin{aligned}
\left. \frac{dy_\epsilon(\bar{s})}{d\epsilon} \right|_{\epsilon=0} &= \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \int_{\bar{s}-\epsilon}^{\bar{s}} [F^e(y_\epsilon(\tau), w_0, w, a)(1 + \zeta) - \bar{F}^e(\tau)] d\tau \\
&= \mathbf{v}_{\mathbf{n}, \bar{s}} + \lim_{\epsilon \downarrow 0} \left\{ \frac{1 + \zeta}{\epsilon} \int_{\bar{s}-\epsilon}^{\bar{s}} [F^e(y_\epsilon(\tau), w_0, w, a) - F^e(\bar{y}(\bar{s}), w_0, w, a)] d\tau \right. \\
&\quad \left. + \frac{1}{\epsilon} \int_{\bar{s}-\epsilon}^{\bar{s}} [\bar{F}^e(\bar{s}) - \bar{F}^e(\tau)] d\tau \right\} \\
&= \mathbf{v}_{\mathbf{n}, \bar{s}} + \lim_{\epsilon \downarrow 0} \frac{1 + \zeta}{\epsilon} \int_{\bar{s}-\epsilon}^{\bar{s}} [F^e(y_\epsilon(\tau), w_0, w, a) - F^e(\bar{y}(\bar{s}), w_0, w, a)] d\tau.
\end{aligned}$$

Sabendo que existe $K > 0$ para o qual $\|y_\epsilon - \bar{y}(\bar{s})\|_\infty \leq K\epsilon$ e que F^e é lipschitziana na componente x , vale

$$|F^e(y_\epsilon(s), w_0, w, a) - F^e(\bar{y}(\bar{s}), w_0, w, a)| \leq \text{lip}_x(F^e)K\epsilon,$$

assim, como

$$\left| \int_{\bar{s}-\epsilon}^{\bar{s}} [F^e(y_\epsilon(\tau), w_0, w, a) - F^e(\bar{y}(\bar{s}), w_0, w, a)] d\tau \right| \leq \int_{\bar{s}-\epsilon}^{\bar{s}} \text{lip}_x(F^e)K\epsilon d\tau = o(\epsilon),$$

é obtido que

$$\left. \frac{dy_\epsilon(\bar{s})}{d\epsilon} \right|_{\epsilon=0} = \mathbf{v}_{\mathbf{n}, \bar{s}}.$$

Repetindo o argumento para ℓ^e , obtemos

$$\left. \frac{dy_{n+1, \epsilon}(\bar{s})}{d\epsilon} \right|_{\epsilon=0} = \mathbf{v}_{\mathbf{n}, \bar{s}, n+1}.$$

Finalmente, como M' é a derivada do fluxo de $\bar{s}$ para s , o resultado fica provado. $\square$

Lema 89 (página 834 de [1]). Assuma a *Hipótese 3* com $\hat{\ell}_1(\cdot, 0, \cdot) \equiv 0$ e fixe um ponto $(y'_0, y', y'_{n+1}, \beta') \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ e $a \in A$. Sob tais condições, dado um colchete de Lie $B \in \mathfrak{B}^0$ de comprimento h , existe $\hat{\epsilon} > 0$ para a qual, dado $s \in]0, \hat{\epsilon}^{\frac{1}{h}}[$, existe um controle constante por partes $(w_{B,s,0}; w_{B,s})$ em que a primeira componente se anula em $[0, s]$ e a segunda assume valores nos vetores canônicos e/ou em seus opostos e de modo que, tomando (y_0, y, y_{n+1}, b) como solução de (3.31) correspondente a $(w_{B,s,0}; w_{B,s}; a; 0)$ e à condição inicial $(y'_0, y', y'_{n+1}, \beta')$, segue que

$$\begin{aligned}
(y_0, y_{n+1})(s) &= (y'_0, y'_{n+1}), \quad b(s) = \beta' + s \quad \forall s \in [0, s], \\
y(s) &= y' + \left[\frac{s}{r(B)} \right]^h + o(s^h).
\end{aligned}$$

Aproximações tipo-colchete

Definição 90 (página 835 de [1]). Fixe $\bar{s} \in]0, \bar{S}[$ e seja $B \in \mathfrak{B}^0$ um gerador de variação tipo-colchete de comprimento h . Dado $\hat{\epsilon} > 0$ como no **Lema 89**, ponha o seguinte *alongamento* ϵ

$$\begin{aligned} \gamma^\epsilon : \left[\bar{s} - 2\epsilon^{\frac{1}{h}}, \bar{s} - \epsilon^{\frac{1}{h}} \right] &\rightarrow \left[\bar{s} - 2\epsilon^{\frac{1}{h}}, \bar{s} \right] \\ s &\mapsto \left(\bar{s} - 2\epsilon^{\frac{1}{h}} \right) + 2 \left[s - \left(\bar{s} - 2\epsilon^{\frac{1}{h}} \right) \right]. \end{aligned} \quad (3.37)$$

Dados $(w'_0, w', \alpha', \zeta') \in L^\infty([0, \bar{S}]; W \times A \times [-\varrho, \varrho])$, $a \in A$ e $\mathfrak{w}_{B, \epsilon^{\frac{1}{h}}}$ como no **Lema 89**, o conjunto $\left\{ (w'_0, w', \alpha', \zeta')_{B, \bar{s}, \epsilon} \mid \epsilon \in]0, \epsilon[\text{ e } 2\epsilon^{\frac{1}{h}} < \bar{s} \right\}$ dos elementos definidos por

$$(w'_0, w', \alpha', \zeta')_{B, \bar{s}, \epsilon}(s) = \begin{cases} (2w'_0, 2w', \alpha', \zeta') \circ \gamma^\epsilon(s) & \text{se } s \in \left[\bar{s} - 2\epsilon^{\frac{1}{h}}, \bar{s} \right], \\ \left(0, w_{B, \epsilon^{\frac{1}{h}}} \left(s - \left(\bar{s} - \epsilon^{\frac{1}{h}} \right) \right), a, \zeta' \right) & \text{se } s \in \left[\bar{s} - \epsilon^{\frac{1}{h}}, \bar{s} \right], \\ (w'_0, w', \alpha', \zeta')(s) & \text{nos outros casos} \end{cases} \quad (3.38)$$

é denominado **família de aproximações tipo-colchete** para o controle $(w'_0, w', \alpha', \zeta')$ em $\bar{s}$ **correspondente a B**.

Proposição 91 (página 835 de [1]). *Considere um gerador $B \in \mathfrak{B}^0$ de comprimento h e, para cada $\bar{s} \in]0, \bar{S}[$ e $\epsilon > 0$ como na **Definição 90**, sejam $(\bar{w}_0, \bar{w}, \bar{\alpha}, 0)_{B, \bar{s}, \epsilon}$ uma aproximação tipo-colchete para o controle $(\bar{w}_0, \bar{w}, \bar{\alpha}, 0)$ em $\bar{s}$ correspondente a B e $(y_{0, \epsilon}; y_\epsilon; y_{n+1, \epsilon}; \beta_\epsilon)$ a solução correspondente de (3.31). Sob tais condições, cada $s \in]\bar{s}, \bar{S}[$ satisfaz*

$$\begin{bmatrix} y_{0, \epsilon}(s) - \bar{y}_0(s) \\ y_\epsilon(s) - \bar{y}(s) \\ y_{n+1, \epsilon}(s) - \bar{y}_{n+1}(s) \end{bmatrix} = \epsilon M'(s, \bar{s}) \begin{bmatrix} \mathbf{v}_{B, \bar{s}, 0} \\ \mathbf{v}_{B, \bar{s}} \\ \mathbf{v}_{B, \bar{s}, n+1} \end{bmatrix} + o(\epsilon) \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\text{e } \beta_\epsilon(s) - \bar{\beta}(s) = \epsilon^{\frac{1}{h}}.$$

Demonstração. Em primeiro lugar,

$$\begin{aligned} \bar{y}_0(\gamma_\epsilon(s)) &= 2 \int_{\bar{s}-2\epsilon^{\frac{1}{h}}}^s \bar{w}_0(\gamma_\epsilon(\tau)) d\tau = \int_{\bar{s}-2\epsilon^{\frac{1}{h}}}^s (2\bar{w}_0) \circ \gamma_\epsilon(\tau) d\tau = y_{0, \epsilon}(s) \\ \bar{y}(\gamma_\epsilon(s)) &= 2 \int_{\bar{s}-2\epsilon^{\frac{1}{h}}}^s F^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}) \circ \gamma_\epsilon(\tau) d\tau = \int_{\bar{s}-2\epsilon^{\frac{1}{h}}}^s F^e(\bar{y}, 2\bar{w}_0, 2\bar{w}, \bar{\alpha})(\gamma_\epsilon(\tau)) d\tau \\ &= y_\epsilon(s) \\ \bar{y}_{n+1}(\gamma_\epsilon(s)) &= 2 \int_{\bar{s}-2\epsilon^{\frac{1}{h}}}^s \ell^e(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}) \circ \gamma_\epsilon(\tau) d\tau = \int_{\bar{s}-2\epsilon^{\frac{1}{h}}}^s \ell^e(\bar{y}, 2\bar{w}_0, 2\bar{w}, \bar{\alpha})(\gamma_\epsilon(\tau)) d\tau \\ &= y_{n+1, \epsilon}(s), \end{aligned}$$

para todo $s \in \left[\bar{s} - 2\epsilon^{\frac{1}{h}}, \bar{s} - \epsilon^{\frac{1}{h}} \right]$, o que, de (3.31), (90) e da definição de ℓ^e , mostra que $y_{0, \epsilon}(\bar{s}) = y_{0, \epsilon}(\bar{s} - \epsilon^{\frac{1}{h}}) = \bar{y}_0(\bar{s})$, $y_{n+1, \epsilon}(\bar{s}) = y_{n+1, \epsilon}(\bar{s} - \epsilon^{\frac{1}{h}}) = \bar{y}_{n+1}(\bar{s})$. Ainda, chamando $w_{B, \epsilon^{\frac{1}{h}}}$ o controle do **Lema 89** associado a B , vale

$$\begin{aligned}
y_\epsilon(\bar{s}) - \bar{y}(\bar{s}) &= \int_{\bar{s}-\epsilon^{\frac{1}{h}}}^{\bar{s}} \left[\sum_{i=1}^m (g_i \circ y_\epsilon)(s) w_{B, \epsilon^{\frac{1}{h}}, i} \left(s - \left(\bar{s} - \epsilon^{\frac{1}{h}} \right) \right) \right] ds \\
&= \int_0^{\epsilon^{\frac{1}{h}}} (g_i \circ y_\epsilon) \left(s + \left(\bar{s} - \epsilon^{\frac{1}{h}} \right) \right) w_{B, \epsilon^{\frac{1}{h}}, i}(s) ds.
\end{aligned}$$

Ainda, denominando y_ϵ a solução de

$$\begin{aligned}
\dot{y} &= \sum_{i=1}^m g_i(y) w_{B, \epsilon^{\frac{1}{h}}, i}, \\
y(0) &= \bar{y}(\bar{s}),
\end{aligned}$$

do **Lema 89**, segue

$$y_\epsilon(\bar{s}) - \bar{y}(\bar{s}) - \left[\frac{\epsilon^{\frac{1}{h}}}{r(B)} \right]^h (B \circ \bar{y})(\bar{s}) = y_\epsilon \left(\epsilon^{\frac{1}{h}} \right) - \bar{y}(\bar{s}) - \epsilon \frac{(B \circ \bar{y})(\bar{s})}{[r(B)]^h} = o(\epsilon),$$

o que mostra que $y_\epsilon(\bar{s}) - \bar{y}(\bar{s}) = \epsilon \frac{(B \circ \bar{y})(\bar{s})}{[r(B)]^h} + o(\epsilon)$, assim, como M' é a derivada do fluxo de $\bar{s}$ para s , a segunda igualdade fica provada.

Finalmente, $\beta_\epsilon(s) - \bar{\beta}(s) = \beta_\epsilon(\bar{s}) - \bar{\beta}(\bar{s}) = \epsilon^{\frac{1}{h}}$. □

3.5.2 Aproximações compostas

Sendo $\mathbf{g} \in \mathfrak{D}$ e $\bar{s} \in]0, \bar{S}[$, para cada $\epsilon > 0$, coloque o seguinte operador:

$$\begin{aligned}
\mathcal{A}_{\mathbf{g}, \bar{s}, \epsilon} : L^\infty \left([0, \bar{S}]; [0, \infty[\times C \times A \times [-\varrho, \varrho] \right) &\rightarrow L^\infty \left([0, \bar{S}]; [0, \infty[\times C \times A \times [-\varrho, \varrho] \right) \\
(w_0, w, \alpha, \zeta) &\mapsto (w_0, w, \alpha, \zeta)_{\mathbf{g}, \bar{s}, \epsilon}.
\end{aligned}$$

Proposição 92 (página 836 de [1]). *Fixe $N \in \mathbb{N}$. Sendo $\vec{\mathbf{g}} := (\mathbf{g}_1, \dots, \mathbf{g}_N)$ uma N -úpla de geradores de variação com respectivos comprimentos $\vec{h} := (h_1, \dots, h_N)$, tome $\vec{s} := (\bar{s}_1, \dots, \bar{s}_N) \in]0, \bar{S}[^N$ com coordenadas crescentes e tal que, quando $h_j = 1$, o respectivo $s_j \in L_{0, \bar{s}}$. Ainda, para cada $\vec{\epsilon} := (\epsilon_1, \dots, \epsilon_N) \in]0, \infty[^N$ pequeno o suficiente, ponha*

$$(w_{0, \vec{\epsilon}}, w_\epsilon, \alpha_\epsilon, \zeta_\epsilon) = \mathcal{A}_{\mathbf{g}_N, \bar{s}_N, \epsilon_N} \circ \dots \circ \mathcal{A}_{\mathbf{g}_1, \bar{s}_1, \epsilon_1}(\bar{w}_0, \bar{w}, \bar{\alpha}, 0) \quad (3.39)$$

e denote por $(w_{0, \vec{\epsilon}}, w_\epsilon, \alpha_\epsilon, \zeta_\epsilon, y_{0, \vec{\epsilon}}, y_\epsilon, y_{n+1, \vec{\epsilon}}, \beta_\epsilon)$ seu processo correspondente.

Sob tais condições, para todo $s \in]\bar{s}_N, \bar{S}]$, vale

$$\begin{bmatrix} y_{0,\epsilon}(s) - \bar{y}_0(s) \\ y_\epsilon(s) - \bar{y}(s) \\ y_{n+1,\epsilon}(s) - \bar{y}_{n+1}(s) \end{bmatrix} = \sum_{j=1}^N \epsilon_j M'(s, \bar{s}_j) \begin{bmatrix} \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, 0} \\ \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \\ \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \end{bmatrix} + o(|\vec{\epsilon}|) \quad (3.40)$$

$$\beta_\epsilon(s) - \bar{\beta}(s) = \sum_{j \in I_1} \epsilon_j [|w_j|(1 + \zeta_j) - |\bar{w}(\bar{s}_j)|] + o(|\vec{\epsilon}|) + \sum_{j \in \{1 : N\} \setminus I_1} \epsilon_j^{\frac{1}{h_j}}, \quad (3.41)$$

em que I_1 é o conjunto de todos os j tais que $h_j = 1$.

Demonstração. Para $N = 1$, o resultado segue das Proposições 88 e 91.

Assumindo, agora, que o resultado vale para $N - 1$ com $N \in \{2 : \infty\}$, sendo $(y_0, y, y_{n+1}, \beta)_N$ e $(y_0, y, y_{n+1}, \beta)_{N-1}$ as trajetórias associadas com todas as variações e com as $N - 1$ primeiras variações respectivamente, vale

$$\begin{bmatrix} y_{0,N-1}(\bar{s}_N) - \bar{y}_0(\bar{s}_N) \\ y_{N-1}(\bar{s}_N) - \bar{y}(\bar{s}_N) \\ y_{n+1,N-1}(\bar{s}_N) - \bar{y}_{n+1}(\bar{s}_N) \end{bmatrix} = \sum_{j=1}^{N-1} \epsilon_j \begin{bmatrix} \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, 0} \\ M(\bar{s}_N, \bar{s}_j) \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \\ \mu(s, \bar{s}_j) \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} + \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \end{bmatrix} + o\left(\left\| \begin{bmatrix} \epsilon_1 \\ \dots \\ \epsilon_{N-1} \end{bmatrix} \right\|\right) \quad (3.42)$$

e, chamando $I_{1,N-1}$ o conjunto de índices de $\{1 : N - 1\}$ para o qual $h_j = 1$,

$$\beta_{N-1}(\bar{s}_N) - \bar{\beta}(\bar{s}_N) = \sum_{j \in I_{1,N-1}} \epsilon_j [|w_j|(1 + \zeta_j) - |\bar{w}(\bar{s}_j)|] + o\left(\left\| \begin{bmatrix} \epsilon_1 \\ \dots \\ \epsilon_{N-1} \end{bmatrix} \right\|\right) + \sum_{j \in \{1 : N-1\} \setminus I_{1,N-1}} \epsilon_j^{\frac{1}{h_j}}. \quad (3.43)$$

Mostrando que

$$\begin{bmatrix} y_{0,N}(\bar{s}_N) - y_{0,N-1}(\bar{s}_N) \\ y_N(\bar{s}_N) - y_{N-1}(\bar{s}_N) \\ y_{n+1,N}(\bar{s}_N) - y_{n+1,N-1}(\bar{s}_N) \end{bmatrix} = \epsilon_N \begin{bmatrix} \mathbf{v}_{\mathbf{g}_N, \bar{s}_N, 0} \\ \mathbf{v}_{\mathbf{g}_N, \bar{s}_N} \\ \mathbf{v}_{\mathbf{g}_N, \bar{s}_N, n+1} \end{bmatrix} + o(|\vec{\epsilon}|) \quad (3.44)$$

e que

$$\beta_N(\bar{s}_N) - \beta_{N-1}(\bar{s}_N) = \begin{cases} \epsilon_N [|w_N|(1 + \zeta_N) - |\bar{w}(\bar{s}_N)|] + o(|\epsilon_N|) & \text{se } h_N = 1, \\ \epsilon_N^{\frac{1}{h_N}} & \text{se } h_N \geq 2, \end{cases} \quad (3.45)$$

o resultado ficará provado para N .

Começamos com o caso em que $h_N \geq 2$. Aqui, $\mathbf{g}_N = B_N \in \mathfrak{B}^0$, e, utilizando argumentos análogos a alguns utilizados na demonstração da **Proposição 91**,

$$\begin{bmatrix} y_{0,N} \\ y_N \\ y_{n+1,N} \\ \beta_N \end{bmatrix} \left(\bar{s}_N - \epsilon_N^{\frac{1}{h_N}} \right) = \begin{bmatrix} y_{0,N-1} \\ y_{N-1} \\ y_{n+1,N-1} \\ \beta_{N-1} \end{bmatrix} (\bar{s}_N),$$

o que mostra que $y_{0,N}(\bar{s}_N) - y_{0,N-1}(\bar{s}_N) = 0$, $y_{n+1,N}(\bar{s}_N) - y_{n+1,N-1}(\bar{s}_N) = 0$, $\beta_N(\bar{s}_N) - \beta_{N-1}(\bar{s}_N) = \frac{1}{\epsilon_N^{h_N}}$, enquanto, chamando $w_{B_N, \epsilon_N^{\frac{1}{h_N}}}$ o controle do **Lema 89**,

$$\begin{aligned} y_N(\bar{s}_N) - y_{N-1}(\bar{s}_N) &= \int_{\bar{s}_N - \epsilon_N^{\frac{1}{h_N}}}^{\bar{s}_N} \left[\sum_{i=1}^m (g_i \circ y_N)(s) w_{B_N, \epsilon_N^{\frac{1}{h_N}}, i} \left(s - \left(\bar{s}_N - \epsilon_N^{\frac{1}{h_N}} \right) \right) \right] ds \\ &= \int_0^{\epsilon_N^{\frac{1}{h_N}}} \left[(g_i \circ y_N) \left(s - \left(\bar{s}_N - \epsilon_N^{\frac{1}{h_N}} \right) \right) w_{B_N, \epsilon_N^{\frac{1}{h_N}}, i} (s) \right] ds. \end{aligned}$$

Agora, chamando y_N a solução de

$$\begin{aligned} \dot{y} &= \sum_{i=1}^m g_i(y) w_{B_N, \epsilon_N^{\frac{1}{h_N}}, i}, \\ y(0) &= y_{N-1}(\bar{s}_N), \end{aligned}$$

vale $y_N(\bar{s}_N) = y_N\left(\epsilon_N^{\frac{1}{h_N}}\right)$ e, pelo **Lema 89**,

$$\begin{aligned} y_N(\bar{s}_N) - y_{N-1}(\bar{s}_N) &- \left[\frac{\epsilon_N^{\frac{1}{h_N}}}{r(B_N)} \right]^{h_N} (B_N \circ \bar{y})(\bar{s}_N) \\ &= y_N\left(\epsilon_N^{\frac{1}{h_N}}\right) - y_{N-1}(\bar{s}_N) - \frac{\epsilon_N}{[r(B_N)]^{h_N}} (B_N \circ y_{N-1})(\bar{s}_N) \\ &+ \frac{\epsilon_N}{[r(B_N)]^{h_N}} [(B_N \circ y_{N-1})(\bar{s}_N) - (B_N \circ y_N)(\bar{s}_N)] \\ &= o(\epsilon_N) + \frac{\epsilon_N}{[r(B_N)]^{h_N}} [(B_N \circ y_{N-1})(\bar{s}_N) - (B_N \circ y_N)(\bar{s}_N)]. \end{aligned}$$

Da continuidade de B_N e da hipótese de indução, segue

$$\begin{aligned} \left| \frac{\epsilon_N}{[r(B_N)]^{h_N}} [(B_N \circ y_{N-1})(\bar{s}_N) - (B_N \circ y_N)(\bar{s}_N)] \right| &\leq \frac{\epsilon_N}{[r(B_N)]^{h_N}} \omega_{B_N} (|y_{N-1}(\bar{s}_N) - \bar{y}(\bar{s}_N)|) \\ &\leq \frac{\epsilon_N}{[r(B_N)]^{h_N}} O\left(\sum_{i=1}^{N-1} \epsilon_i\right), \end{aligned}$$

em que ω_{B_N} denota o módulo de continuidade de B_N . Portanto,

$$y_N(\bar{s}_N) - y_{N-1}(\bar{s}_N) = \frac{\epsilon_N}{[r(B_N)]^{h_N}} (B_N \circ \bar{y})(\bar{s}_N) + o(|\bar{\epsilon}|),$$

o que mostra (3.44) e (3.45).

Agora, passamos para o caso $h_N = 1$. Nesse caso, $\mathbf{g}_N = (w_{0,N}, w_N, a_N, \zeta_N)$, e vale

$$\begin{aligned}
& y_N(\bar{s}_N) - y_{N-1}(\bar{s}_N) \\
&= \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} [F^e(y_N(s), w_{0,N}, w_N, a_N)(1 + \zeta_N) - F^e(y_{N-1}(s), \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s))] ds \\
&= \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} [r_1(s) + r_2 + r_3(s)] ds,
\end{aligned}$$

com

$$\begin{aligned}
r_1(s) &= (1 + \zeta_N) [F^e(y_N(s), w_{0,N}, w_N, a_N) - F^e(\bar{y}(\bar{s}_N), w_{0,N}, w_N, a_N)], \\
r_2 &= F^e(\bar{y}(\bar{s}_N), w_{0,N}, w_N, a_N)(1 + \zeta_N) - \bar{F}^e(\bar{s}_N), \\
r_3(s) &= \bar{F}^e(\bar{s}_N) - F^e(y_{N-1}(s), \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s)).
\end{aligned}$$

O primeiro passo é estimar r_1 . Para $s \in [\bar{s}_N - \epsilon_N, \bar{s}_N]$, a continuidade lipschitziana da aplicação entrada-saída garante que $\|y_N - y_{N-1}\|_\infty = O(\epsilon_N)$, a hipótese de indução, que $\|y_{N-1} - \bar{y}\|_\infty = O\left(\sum_{i=1}^{N-1} \epsilon_i\right)$, e a continuidade lipschitziana da trajetória referencial, que $\|\bar{y} - \bar{y}(\bar{s}_N)\|_\infty = O(\epsilon_N)$, assim, $\|y_N - \bar{y}(\bar{s}_N)\|_\infty = O(|\vec{\ell}|)$, o que mostra que

$$\left| \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} r_1(s) ds \right| \leq \text{lip}(F^e) \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} |y_N(s) - \bar{y}(\bar{s}_N)| ds = O(|\vec{\ell}|).$$

Sabendo que $\bar{s}_N \in L_{0,\bar{s}}$, vale

$$\begin{aligned}
\left| \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} r_3(s) ds \right| &\leq \left| \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} [\bar{F}^e(\bar{s}_N) - \bar{F}^e(s)] ds \right| \\
&\quad + \left| \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} [\bar{F}^e(s) - F^e(y_{N-1}(s), \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s))] ds \right| \\
&\leq o(\epsilon_N) + \epsilon_N O\left(\left|\sum_{i=1}^{N-1} \epsilon_i e_i\right|\right),
\end{aligned}$$

assim, $y_N(\bar{s}_N) - y_{N-1}(\bar{s}_N) = \epsilon_N r_2 + o(|\vec{\ell}|)$, o que mostra a segunda igualdade em (3.44). Analogamente, a terceira igualdade em (3.44) segue. Por fim, a primeira igualdade em (3.44) e a igualdade em (3.45) são triviais. $\square$

3.5.3 Separação de conjuntos

Conjuntos atingíveis

Dado $(w_0, w, \alpha, \zeta, y_0, y, y_{n+1}, \beta)$ um processo factível de (3.32), introduzimos a componente de custo total $y_c := \Psi(y_0, y) + y_{n+1}$ e colocamos $\bar{y}_c := \Psi(\bar{y}_0, \bar{y}) + \bar{y}_{n+1}$.

Definição 93 (página 839 de [1]). Para qualquer $\delta > 0$, definimos o **conjunto δ atingível** $\mathcal{R}_\delta$ como o conjunto das imagens, em $\bar{S}$, de quadrúplas (y_0, y, y_c, β) cuja distância até a quadrúpla referencial $(\bar{y}_0, \bar{y}, \bar{y}_c, \bar{\beta})$ é menor que δ . Ainda, definimos $\mathcal{R}'_\delta$ como sua projeção nas três primeiras coordenadas.

Agora, sendo $N \in \mathbb{N}$, $\vec{g} \in \mathfrak{D}^N$ e $\vec{s}$ como na **Proposição 92**, defina o conjunto E' dos elementos da forma $M(s, \bar{s}_j) \begin{bmatrix} \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, 0} & \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} & \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \end{bmatrix}^T$ com $j \in \{1 : N\}$, e, no caso em que todos os $\mathbf{g}_j$ são geradores de variações de agulha, o conjunto E dos elementos da forma

$$M(s, \bar{s}_j) \begin{bmatrix} \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, 0} \\ \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \\ \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \\ \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, \beta} \end{bmatrix}$$

com $j \in \{1 : N\}$. Aqui, E' é a projeção de E em $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$.

Por último, defina os cones convexos R e R' gerados por E e E' .

Proposição 94 (página 839 de [1]). Nas condições da **Definição 93**, são válidas as seguintes afirmações:

- O conjunto R' é um cone aproximante de Boltjanskij para o conjunto $\mathcal{R}'_\delta$ no ponto $(\bar{y}_0, \bar{y}, \bar{y}_c)(\bar{S})$;
- Quando todos os $\mathbf{g}_j$ são variações de agulha, o conjunto R é um cone aproximante de Boltjanskij para $\mathcal{R}$ em $(\bar{y}_0, \bar{y}, \bar{y}_c, \bar{\beta})(\bar{S})$.

Demonstração. Pondo $y_{c, \vec{\epsilon}} = \Psi \circ (y_{0, \vec{\epsilon}}, y_{\vec{\epsilon}})(s) + y_{\vec{\epsilon}, \vec{\epsilon}}$, de (3.40), segue que

$$\begin{aligned} y_{c, \vec{\epsilon}}(\bar{S}) - \bar{y}_c(\bar{S}) &= \Psi \circ (y_{0, \vec{\epsilon}}, y_{\vec{\epsilon}})(s) + y_{n+1, \vec{\epsilon}} - \Psi \circ (\bar{y}_0, \bar{y})(s) + \bar{y}_{n+1} \\ &= \sum_{j=1}^N \epsilon_j \left[\frac{\partial \bar{\Psi}}{\partial t}(\bar{S}) \mathbf{v}_{0, \mathbf{c}_j, \bar{s}_j} + \frac{\partial \bar{\Psi}}{\partial x}(\bar{S}) \cdot M(\bar{S}, \bar{s}_j) + \langle \mu(\bar{S}, \bar{s}_j); \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \rangle + \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \right] \\ &\quad + o(|\vec{\epsilon}|), \end{aligned}$$

o que mostra, pela **Proposição 92** e seguindo um raciocínio análogo à **Propriedade 51**, o segundo item.

Para provar o primeiro, para $\epsilon > 0$ pequeno o suficiente, defina a função $F : [0, \infty]^N \cap B[0, \epsilon] \rightarrow \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}$ colocando $F(\vec{\epsilon}) = (y_{0, \vec{\epsilon}}(\bar{S}), y_{\vec{\epsilon}}(\bar{S}), y_{c, \vec{\epsilon}}(\bar{S}))$. De maneira análoga ao que foi feito acima e pela **Proposição 92**, temos

$$\begin{aligned} F(\vec{\epsilon}) &= (\bar{y}_0(\bar{S}), \bar{y}(\bar{S}), \bar{y}_c(\bar{S})) \\ &\quad + \sum_{j=1}^N \left[\begin{array}{c} \mathbf{v}_{0, \mathbf{g}_j, \bar{s}_j} \\ M(\bar{S}, \bar{s}_j) \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \\ \frac{\partial \bar{\Psi}}{\partial t}(\bar{S}) \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, 0} + \frac{\partial \bar{\Psi}}{\partial x}(\bar{S}) \cdot M(\bar{S}, \bar{s}_j) \cdot \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} + \langle \mu(\bar{S}, \bar{s}_j); \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \rangle + \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \end{array} \right], \end{aligned}$$

o que termina a demonstração pela **Proposição 92** e seguindo um raciocínio análogo à **Propriedade 51**. $\square$

Agora, considere o conjunto lucrável $\mathcal{P}$ e sua projeção $\mathcal{P}'$ definido como

$$\begin{aligned}\mathcal{P} &= \mathcal{T} \times]-\infty, \bar{y}_c(\bar{S})[\times [0, K] \cup \{(\bar{y}_0, \bar{y}, \bar{y}_c, \bar{\beta})(\bar{S})\}, \\ \mathcal{P}' &= \mathcal{T} \times]-\infty, \bar{y}_c(\bar{S})[\cup \{(\bar{y}_0, \bar{y}, \bar{y}_c)(\bar{S})\},\end{aligned}$$

e seja Γ um cone aproximante de Boltjanskij para o alvo $\mathfrak{T}$ em $(\bar{y}_0, \bar{y})(\bar{S})$. É possível verificar que

$$\begin{aligned}P &:= \Gamma \times]-\infty, 0] \times \{0\}, \\ P' &:= \Gamma \times]-\infty, 0]\end{aligned}$$

são cones aproximantes de Boltjanskij para $\mathcal{P}$ e $\mathcal{P}'$ em $(\bar{y}_0, \bar{y}, \bar{y}_c)(\bar{S})$ respectivamente.

Proposição 95 (página 840 de [1]). *Existe $\delta > 0$ tal que $\mathcal{P}'$ e $\mathcal{R}'_\delta$ são, localmente, separados em $(\bar{y}_0, \bar{y}, \bar{y}_c)(\bar{S})$.*

Demonstração. Supondo que $\mathcal{P}$ e $\mathcal{R}'_\delta$ não são, localmente, separados em $(\bar{y}_0, \bar{y}, \bar{y}_c)(\bar{S})$, dado $\delta \in]0, K - \bar{\beta}(\bar{S})[$, existe um processo $(\bar{S}, w_0, w, \alpha, \zeta, y_0, y, y_{n+1}, \beta)$ de (3.31) satisfazendo

$$(y_0, y, y_c)(\bar{S}) \in \mathcal{R}'_\delta \cap \mathcal{P}', \quad \text{dist}((y_0, y, y_c, \beta); (\bar{y}_0, \bar{y}, \bar{y}_c, \bar{\beta})) < \delta.$$

Isto mostra que $\beta(\bar{S}) \leq \delta + \bar{\beta}(\bar{S}) < K$, assim, o ponto final $(y_0, y, y_c, \beta)(\bar{S}) \in \mathcal{R}_\delta \cap \mathcal{P}$, o que implica que, para cada $\delta \in]0, K - \bar{\beta}(\bar{S})[$, os conjuntos $\mathcal{P}$ e $\mathcal{R}_\delta$ não são, localmente, separados, o que contradiz a otimalidade local do processo referencial. $\square$

A partir desse resultado, como P' não é subespaço, a **Proposição 36** e o **Teorema 53** garantem a existência de $(\xi_0, \xi, \xi_c) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \setminus \{(0, 0, 0)\}$ satisfazendo

$$\begin{aligned}\langle (v_0, v, v_c); (\xi_0, \xi, \xi_c) \rangle &\leq 0 \quad \forall (v_0, v, v_c) \in R' \\ \langle (v_0, v, v_c); (\xi_0, \xi, \xi_c) \rangle &\geq 0 \quad \forall (v_0, v, v_c) \in P' .\end{aligned}$$

Em particular, escolhendo $(0, 0, v_c) \in P'$, segue que $\xi_c \leq 0$, o que permite escolher $\lambda = -\xi_c$ e, escolhendo $(\xi_0, \xi, 0) \in P'$, segue que $(\xi_0, \xi) \in (-\Gamma')^\perp$ e $(\xi_0, \xi, \xi_c) \cdot (v_0, v, v_c) \leq 0$ para todo $(v_0, v, v_c) \in R'$, que, pela definição de R' , é equivalente a

$$\begin{aligned}\left[\xi_0 - \lambda \frac{\partial \bar{\Psi}}{\partial t}(\bar{S}) \right] \mathbf{v}_{0, \mathbf{g}_j, \bar{s}_j} + \left[\xi - \lambda \frac{\partial \bar{\Psi}}{\partial x}(\bar{S}) \right] \cdot M(\bar{S}, \bar{s}_j) \cdot \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \\ - \lambda \left[\langle \mu(\bar{S}, \bar{s}_j); \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \rangle + \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \right] \leq 0 \quad \forall j \in \{1 : N\} .\end{aligned}$$

Agora, colocando

$$(p_0, p)(s) = \left(\xi_0 - \lambda \frac{\partial \bar{\Psi}}{\partial t}(\bar{S}), M(\bar{S}, s) \left[\xi - \lambda \frac{\partial \bar{\Psi}}{\partial x}(\bar{S}) \right] - \lambda \mu(\bar{S}, s) \right),$$

o multiplicador (p_0, p, λ) com $p_0 \in \mathbb{R}$, $p \in \mathcal{AC}([0, \bar{S}]; \mathbb{R}^n)$ e $\lambda \in [0, \infty[$ satisfaz a condição de não-trivialidade e (3.10). Mais ainda, pela definição de $M(\bar{S}, \cdot)$ e $\mu(\bar{S}, \cdot)$,

$$\begin{aligned} \dot{p} &= \frac{\partial M}{\partial s}(\bar{S}, \cdot) \left[\xi - \lambda \frac{\partial \bar{\Psi}}{\partial x}(\bar{S}) \right] - \lambda \frac{\partial \mu}{\partial s}(\bar{S}, \cdot) \\ &= \frac{\partial F^e}{\partial x}(\bar{y}, \bar{w}_0, \bar{w}, \bar{\alpha}) \cdot M(\bar{S}, \cdot) \cdot \left[\xi - \lambda \frac{\partial \bar{\Psi}}{\partial x}(\bar{S}) \right] - \lambda \nabla \ell^e \cdot M(\bar{S}, \cdot) \\ &= -\frac{\partial H}{\partial x}(\bar{y}, p, 0, \lambda, \bar{w}_0, \bar{w}, \bar{\alpha}), \end{aligned}$$

o que significa que p soluciona (3.11). Finalmente, como $p_0 \mathbf{v}_{0, \mathbf{g}_j, \bar{s}_j} + \langle p(\bar{s}_j); \mathbf{v}_{\mathbf{g}_j, \bar{s}_j} \rangle - \lambda \mathbf{v}_{\mathbf{g}_j, \bar{s}_j, n+1} \leq 0$ para todo $j \in \{1 : N\}$, para um gerador de variação de agulha $\mathbf{c}_j = (w_{0,j}, w_j, a_j, \zeta_j)$, vale

$$\begin{aligned} &H(\bar{y}(\bar{s}_j), p_0, p(\bar{s}_j), 0, \lambda, w_{0,j}(1 + \zeta_j), w_j(1 + \zeta_j), a_j) \\ &\leq H(\bar{y}(\bar{s}_j), p_0, p(\bar{s}_j), 0, \lambda, \bar{w}_0(\bar{s}_j), \bar{w}(\bar{s}_j), \bar{\alpha}(\bar{s}_j)), \end{aligned}$$

enquanto que, para um gerador de variação tipo-colchete $\mathbf{c}_j = B_j$, vale

$$\langle p(\bar{s}_j); (B_j \circ \bar{y})(\bar{s}_j) \rangle \leq 0.$$

3.5.4 Extensão por Interseção Finita

Do raciocínio acima, foi obtido que as desigualdades

$$p_0 \mathbf{v}_{\mathbf{g}, s} + \langle p(s); \mathbf{v}_{\mathbf{g}, s} \rangle - \lambda \mathbf{v}_{\mathbf{g}, s} \leq 0 \quad (3.46)$$

$$\begin{aligned} &H(\bar{y}(s), p_0, p(s), 0, \lambda, w_0(1 + \zeta), w(1 + \zeta), a) \\ &\leq H(\bar{y}(s), p_0, p(s), 0, \lambda, \bar{w}_0(s), \bar{w}(s), \bar{\alpha}(s)) \end{aligned} \quad (3.47)$$

valem para todos os conjuntos da forma $F = \{(\bar{s}_1, \mathbf{g}_1); \dots; (\bar{s}_N, \mathbf{g}_N)\}$ em que todos os $\bar{s}_i$ s pertencem a $]0, \bar{S}[$ e crescem e todos os $\mathbf{g}_i$ s pertencem a $\mathfrak{D}$.

Agora, sabendo que a função $(\bar{w}_0, \bar{F}^e, \bar{\ell}^e, |\bar{w}|)$ é mensurável em $L_{0, \bar{S}}$, pelo Teorema de Lusin, existe uma sequência $\{E_k\}$ de conjuntos disjuntos cuja união possui medida cheia em $L_{0, \bar{S}}$ e de modo que a restrição de uma função a qualquer um dos conjuntos fornece uma função contínua.

Tomando, para cada $k \in \mathbb{N}$, D_k como o conjunto de pontos de densidade de E_k , isto é, pontos s satisfazendo o limite

$$\lim_{\delta \downarrow 0} \frac{\text{med}([s - \delta, s + \delta] \cap E_k)}{2\delta} = 1,$$

do Teorema da Densidade de Lebesgue, cada D_k possui a mesma medida correspondente E_k , assim,

o conjunto $D := \bigcup_{k=1}^{\infty} D_k$ possui a mesma medida de $L_{0,\bar{S}}$ — que é $\bar{S}$ pelo Teorema do Conjunto dos Pontos de Lebesgue.

Definição 96 (página 841 de [1]). Dado F um subconjunto arbitrário de $D \times \mathfrak{D}$, chamamos $\Lambda(F)$ o conjunto de todas as triplas unitárias $(\bar{p}_0, \bar{p}, \lambda) \in \mathbb{R} \times \mathbb{R}^n \times [0, \infty[$ tais que, ao colocar $p(\cdot) = \bar{p}M(\bar{S}, \cdot)$:

- $(\bar{p}_0, p(\bar{S})) + \lambda \nabla \bar{\Psi}(\bar{S}) \in -\Gamma^\perp$;
- Dado $(s, \mathbf{g}) \in F$, para $\mathbf{g} = (w_0, w, a, \zeta)$, vale

$$H(\bar{y}(s), \bar{p}_0, p(s), 0, \lambda, w_0(1 + \zeta), w(1 + \zeta), a) \leq H(\bar{y}(s), \bar{p}_0, p(s), 0, \lambda, \bar{w}_0(s), \bar{w}(s), \bar{a}(s)),$$

enquanto que, para $\mathbf{c} = B \in \mathfrak{B}^0$, vale $p(s) \cdot (B \circ \bar{y})(s) \leq 0$.

Proposição 97 (página 841 de [1]). Para cada $F \subseteq D \times \mathfrak{D}$, o conjunto $\Lambda(F)$ é compacto, e, para cada $F_1, F_2 \subseteq D \times \mathfrak{D}$, $\Lambda(F_1 \cup F_2) = \Lambda(F_1) \cap \Lambda(F_2)$.

Demonstração. Para mostrar que $\Lambda(F)$ é compacto para cada $F \subseteq D \times \mathfrak{D}$, basta mostrar que tal conjunto é fechado, já que se trata de um subconjunto da esfera unitária.

Tomando $\{(\bar{p}_{0,k}, \bar{p}_k, \lambda_k) \mid k \in \mathbb{N}\} \subseteq \Lambda(F)$ e pondo $p_k(\cdot) = \bar{p}_k M(\bar{S}, \cdot)$ para cada $k \in \mathbb{N}$, é possível assumir, sem perda de generalidade, que essa sequência converge para algum ponto $(\bar{p}_0, \bar{p}, \lambda) \in \mathbb{S}^{n+1}$, o que implica, entre outros, que $p_k \rightarrow p$ uniformemente. Dito isso:

- Sabendo que $\Gamma^\perp$ é um conjunto fechado, como $(\bar{p}_{0,k}, \bar{p}_k(\bar{S})) + \lambda_k \nabla \bar{\Psi}(\bar{S}) \in -\Gamma^\perp$ para cada $k \in \mathbb{N}$, temos $(\bar{p}_0, p(\bar{S})) + \lambda \nabla \bar{\Psi}(\bar{S}) \in -\Gamma^\perp$;
- Dado $(s, \mathbf{g}) \in F$, para $\mathbf{g} = (w_0, w, a, \zeta)$, vale

$$\begin{aligned} & H(\bar{y}(s), \bar{p}_{0,k}, p_k(s), 0, \lambda_k, w_0(1 + \zeta), w(1 + \zeta), a) \\ & \leq H(\bar{y}(s), \bar{p}_{0,k}, p_k(s), 0, \lambda_k, \bar{w}_0(s), \bar{w}(s), \bar{a}(s)) \end{aligned}$$

para todo $k \in \mathbb{N}$, assim, como H é contínua, fica provado que

$$\begin{aligned} & H(\bar{y}(s), \bar{p}_0, p(s), 0, \lambda, w_0(1 + \zeta), w(1 + \zeta), a) \\ & \leq H(\bar{y}(s), \bar{p}_0, p(s), 0, \lambda, \bar{w}_0(s), \bar{w}(s), \bar{a}(s)). \end{aligned}$$

Já para $\mathbf{g} = B \in \mathfrak{B}^0$, vale $\langle p_k(s); (B \circ \bar{y})(s) \rangle \leq 0$ para todo $k \in \mathbb{N}$, assim, como o produto interno é contínuo, vale $\mathbf{g} = B \in \mathfrak{B}^0$, vale $\langle p(s); (B \circ \bar{y})(s) \rangle \leq 0$.

Isso mostra que $(\bar{p}_0, \bar{p}, \lambda) \in \Lambda(F)$, como desejado.

Agora, por definição, $\Lambda(F_1 \cup F_2) \subseteq \Lambda(F_1)$ e $\Lambda(F_1 \cup F_2) \subseteq \Lambda(F_2)$, o que mostra que $\Lambda(F_1 \cup F_2) \subseteq \Lambda(F_1) \cap \Lambda(F_2)$. Reciprocamente, supondo que $(\bar{p}_0, \bar{p}, \lambda) \in \Lambda(F_1) \cap \Lambda(F_2)$:

- É claro que $(\bar{p}_0, p(\bar{S})) + \lambda \nabla \bar{\Psi}(\bar{S}) \in -\Gamma^\perp$;
- Dado $(s, \mathbf{g}) \in F_1$, para $\mathbf{g} = (w_0, w, a, \zeta)$, vale

$$\begin{aligned} & H(\bar{y}(s), \bar{p}_0, p(s), 0, \lambda, w_0(1 + \zeta), w(1 + \zeta), a) \\ & \leq H(\bar{y}(s), \bar{p}_0, p(s), 0, \lambda, \bar{w}_0(s), \bar{w}(s), \bar{a}(s)), \end{aligned}$$

enquanto que, para $\mathbf{g} = B \in \mathfrak{B}^0$, vale $p(s) \cdot (B \circ \bar{y})(s) \leq 0$. O mesmo se aplica para qualquer $(s, \mathbf{g}) \in F_2$, portanto, se aplica a $(s, \mathbf{g}) \in F_1 \cup F_2$. Isso mostra que $(\bar{p}_0, \bar{p}, \lambda) \in \Lambda(F_1 \cup F_2)$.

□

Como visto, $\Lambda(F) \neq \emptyset$ para todos conjuntos finitos F com instantes ascendentes. Para mostrar que $\Lambda(F) \neq \emptyset$ para todos conjuntos finitos $F = \{(\bar{s}_1, \mathbf{c}_1); \dots; (\bar{s}_N, \mathbf{c}_N)\} \subseteq D \times \mathfrak{D}$ com instantes não-descendentes, começamos notando que, como cada $\bar{s}_j$ pertence a algum $D_{k(j)}$, podemos construir, para cada $j \in \{1 : N\}$, uma sequência $\{\bar{s}_{j,i}\} \subseteq D_{k(j)}$ convergindo para $\bar{s}_j$ tal que, para todo $i \in \mathbb{N}$, $\bar{s}_{1,i} < \dots < \bar{s}_{N,i}$.

Colocando $F_i := \{(\bar{s}_{1,i}, \mathbf{g}_1); \dots; (\bar{s}_{N,i}, \mathbf{g}_N)\}$, segue $\lambda(F_i) \neq \emptyset$, o que permite escolher algum $(\bar{p}_{0,i}, \bar{p}_i, \lambda_i) \in \Lambda(F_i)$. Sabendo que $\{(\bar{p}_{0,i}, \bar{p}_i, \lambda_i)\} \subseteq \mathbb{S}^{n+1}$ para todo $i \in \mathbb{N}$, tomando uma sub-sequência se necessário, é possível assumir $(\bar{p}_{0,i}, \bar{p}_i, \lambda_i) \rightarrow (\bar{p}_0, \bar{p}, \lambda) \in \mathbb{S}^{n+1}$, assim, usando a convergência da correspondente sequência p_i para a correspondente função p em conjunto com propriedades de fechamento de cones aproximantes de Boltjanskij, pode ser visto que $(\bar{p}_0, \bar{p}, \lambda) \in \Lambda(F)$.

Com este conhecimento e utilizando a **Proposição 97**, pode ser concluído que a coleção $\Lambda(F)$ para $F \subseteq D \times \mathfrak{D}$ finito é uma coleção de conjuntos compactos com a propriedade da interseção finita, assim, a interseção de todos os elementos da família não é vazia, o que implica

$$\Lambda(D \times \mathfrak{D}) = \Lambda\left(\bigcup_{\text{card } F < \infty} F\right) = \bigcap_{\text{card } F < \infty} \Lambda(F) \neq \emptyset$$

(a última igualdade pode ser mostrada da mesma forma que a **Proposição 97**). Isto significa que existe um multiplicador não-trivial tal que, ao colocar $p(\cdot) = \bar{p}M(\bar{S}, \cdot)$, para todos os instantes $s \in D$ (um conjunto de medida cheia), vale

$$\begin{aligned} & \mathfrak{h}(\bar{y}(s), p_0, p(s), 0, \lambda) \\ & = H(\bar{y}(s), p_0, p(s), 0, \lambda, \bar{w}_0(s), \bar{w}(s), \bar{a}(s)) \\ & = \max \left\{ H(\bar{y}(s), p_0, p(s), 0, \lambda, w_0(1 + \zeta), w(1 + \zeta), a) \mid (w_0, w, a, \zeta) \in W \times A \times \left[-\frac{1}{2}, \frac{1}{2}\right] \right\} \\ & = \max \left\{ (1 + \zeta) \mathfrak{h}(\bar{y}(s), p_0, p(s), 0, \lambda) \mid \zeta \in \left[-\frac{1}{2}, \frac{1}{2}\right] \right\}, \\ & 0 \geq \langle p, B \circ \bar{y} \rangle(s) \quad \forall B \in \mathfrak{B}^0. \end{aligned}$$

A primeira igualdade coincide com (3.12) enquanto que a última implica (3.13). Finalmente, dado que $B \in \mathfrak{B}^0 \Leftrightarrow -B \in \mathfrak{B}^0$, a desigualdade se torna uma igualdade para todo $B \in \mathfrak{B}^0$, implicando (3.19). Isto conclui a prova do Teorema 76.

4 Formalismo de Dubovickij-Miljutin com cone tangente

4.1 Introdução

Neste capítulo, abordaremos uma versão do *Formalismo de Dubovickij-Miljutin* que utiliza o cone tangente de Clarke, introduzido em [12]. Ela é baseada, principalmente, em conceitos e resultados encontrados em [13], porém, com o principal resultado inspirado em [50]. Como já comentado, embora tal artigo apresente um resultado desse caráter, ele pode ser estendido, a princípio, para espaços de Hilbert, pois utiliza conceitos de ângulo e compacidade da bola unitária. No entanto, dado que o conjunto de controles do problema (1.3) é localizado em um espaço L^∞ , que não é espaço de Hilbert nem apresenta propriedades “boas” (como a reflexividade, para saber mais, consulte [9] e [11]).

4.2 Cones especiais

Os resultados dessa seção serão formulados supondo que X é um espaço de Banach real e $f: X \rightarrow \mathbb{R}$ é lipschitziano de posto $K > 0$ em uma vizinhança de um ponto $x_0 \in X$.

4.2.1 Cone contingente

Definição 98 (página 21 de [11]). Sejam $Q \subseteq X$ e $x_0 \in Q$. Um vetor $v \in X$ é denominado **direção contingente a Q em x_0** se

$$\liminf_{t \downarrow 0} \frac{d_Q(x_0 + tv)}{t} = 0.$$

O conjunto delas será denotado por $\mathfrak{C}(Q; x_0)$.

Propriedade 99. $\mathfrak{C}(Q; x_0)$ é um cone, fortemente, fechado com vértice na origem.

Demonstração. O fato de que $\mathfrak{C}(Q; x_0)$ é, fortemente, fechado segue das Proposições 183 e 184. Agora, fixados $v \in \mathfrak{C}(Q; x_0)$ e $\alpha \geq 0$: se $\alpha \neq 0$, então,

$$\liminf_{t \downarrow 0} \frac{d_Q(x_0 + t(\alpha v))}{t} = \alpha \liminf_{t \downarrow 0} \frac{d_Q(x_0 + (at)v)}{at} = 0$$

o que mostra que $\alpha v \in \mathfrak{C}(Q; x_0)$; já no caso em que $\alpha = 0$, temos que $0 \in \mathfrak{C}(Q; x_0)$, portanto, tal conjunto é um cone com vértice na origem. $\square$

Proposição 100 (página 30 de [11]). Se $Q \subseteq X$ é convexo e $x_0 \in Q$, então, $\mathfrak{C}(Q; x_0)$ é convexo e coincide com o fecho de

$$\bigcup_{t>0} \frac{Q - x_0}{t}.$$

Demonstração. Chamando a união em questão de Σ , é claro que esta é um cone com vértice na origem. Ainda, dados $u_1, u_2 \in \Sigma$, existem $t_1, t_2 > 0$ de modo que $x_0 + t_i u_i \in Q$ para todo $i \in \{1, 2\}$, então, a convexidade de Q garante que

$$x_0 + \frac{t_1 t_2}{t_1 + t_2} (u_1 + u_2) = \left(1 - \frac{t_1}{t_1 + t_2}\right) (x_0 + t_1 u_1) + \frac{t_1}{t_1 + t_2} (x_0 + t_2 u_2) \in Q,$$

o que mostra que $u_1 + u_2 \in \Sigma$ e, consequentemente, que tal cone é convexo.

Dito isso, da **Definição 98**, é claro que $\mathfrak{C}(Q; x_0) \subseteq \bar{\Sigma}$. Agora, fixando $v \in \Sigma$, existem $q \in Q$ e $t > 0$ de modo que $x_0 + tv = q$, e, dada $t_i \downarrow 0$, existe $\iota \in \mathbb{N}$ de modo que $i \geq \iota \Rightarrow x_i := (1 - t_i)x_0 + t_i q \in Q$, logo

$$i \geq \iota \Rightarrow \frac{x_i - x_0}{t t_i} = \frac{q - x_0}{t} = v,$$

o que mostra que $v \in \mathfrak{C}(Q; x_0)$. Da **Propriedade 99**, só pode ocorrer que $\mathfrak{C}(Q; x_0) = \bar{\Sigma}$. □

Corolário 101. Se $Q \subseteq X$ é convexo e $x_0 \in Q$, então, $Q \subseteq x_0 + \mathfrak{C}(Q; x_0)$.

Propriedade 102 (página 124 de [4]). • Dados $Q_1, \dots, Q_n \subseteq X$ e $x_0 \in Q := \bigcap_{i=1}^n Q_i$,

$$\mathfrak{C}(Q; x_0) \subseteq \bigcap_{i=1}^n \mathfrak{C}(Q_i; x_0);$$

• Dados $Q_1, \dots, Q_n \subseteq X$ e $x_0 \in Q := \bigcup_{i=1}^n Q_i$,

$$\mathfrak{C}(Q; x_0) = \bigcup_{i=1}^n \mathfrak{C}(Q_i; x_0);$$

• Dados $Q_i \subseteq X_i$ e $x_{i,0} \in Q_i$ para $i \in \{1 : n\}$ e pondo $Q = \prod_{i=1}^n Q_i$ e $x_0 = (x_{1,0}, \dots, x_{n,0})$,

$$\mathfrak{C}(Q; x_0) \subseteq \prod_{i=1}^n \mathfrak{C}(Q_i; x_{i,0}).$$

4.2.2 Cone tangente

Definição 103 (página 212 de [11]). Sendo $Q \subseteq X$ e $x_0 \in Q$, um vetor $v \in X$ é denominado **direção tangente a Q em x_0** se $d_Q^\circ(x_0; v) = 0$. O conjunto delas será denotado por $\mathfrak{T}(Q; x_0)$.

Propriedade 104. $\mathfrak{T}(Q; x_0)$ é um cone convexo e, fortemente, fechado com vértice na origem.

Demonstração. Dados $v_1, v_2 \in \mathfrak{T}(Q; x_0)$ e $\alpha_1, \alpha_2 \geq 0$, da **Propriedade 15**, temos que $d_Q^\circ(x_0; \alpha_1 v_1 + \alpha_2 v_2) \leq \alpha_1 d_Q^\circ(x_0; v_1) + \alpha_2 d_Q^\circ(x_0; v_2) = 0$, logo, como $d_Q^\circ(x_0; \cdot)$ nunca assume valores negativos dado que d_Q atinge mínimo em x_0 , só pode ocorrer $d_Q^\circ(x_0; \alpha_1 v_1 + \alpha_2 v_2) = 0$, o que mostra que $\alpha_1 v_1 + \alpha_2 v_2 \in \mathfrak{T}(Q; x_0)$ e, consequentemente, que tal conjunto é um cone convexo com vértice na origem.

Ainda, dada $\{v_i\} \subseteq \mathfrak{T}(Q; x_0)$ tendendo a $v \in X$, também da **Propriedade 15**, temos que $d_Q^\circ(x_0; v) = \lim d_Q^\circ(x_0; v_i) = 0$, o que mostra que $v \in \mathfrak{T}(Q; x_0)$ e, conseqüentemente, que tal conjunto é, fortemente, fechado. $\square$

Proposição 105. *Se $Q \subseteq X$, para todo $x_0 \in Q$, vale*

$$\mathfrak{T}(Q; x_0) = \liminf_{x \rightarrow x_0, t \downarrow 0} \frac{Q - x}{t}.$$

Demonstração. Supondo que $v \in \mathfrak{T}(Q; x_0)$, sejam $x_i \rightarrow x_0$ e $t_i \downarrow 0$. Nessas condições,

$$0 = d_Q^\circ(x_0; v) = \lim_{i \uparrow \infty} \frac{d_Q(x_i + t_i v) - d_Q(x_i)}{t_i} = \lim_{i \uparrow \infty} \frac{d_Q(x_i + t_i v)}{t_i}$$

e, usando a definição da função distância, para cada $i \in \mathbb{N}$, é possível encontrar $s_i \in Q$ satisfazendo $\|x_i + t_i v - s_i\|_X \leq d_Q(x_i + t_i v) - \frac{1}{i}$, então, pondo $v_i = \frac{s_i - x_i}{t_i}$, vem

$$\lim_{i \uparrow \infty} \|v_i - v\|_X = \lim_{i \uparrow \infty} t_i \|x_i + t_i v - s_i\|_X = 0.$$

Reciprocamente, suponha que v seja tal que: para cada $x_i \rightarrow x_0$ e $t_i \downarrow 0$, existe $v_i \rightarrow v$ tal que $x_i + t_i v_i \in Q$ para todo $i \in \mathbb{N}$. Dadas $y_i \rightarrow x_0$ e $t_i \downarrow 0$ tais que

$$\lim_{i \uparrow \infty} \frac{d_Q(y_i + t_i v) - d_Q(y_i)}{t_i} = d_Q^\circ(x_0; v),$$

escolha, para cada $i \in \mathbb{N}$, um $s_i \in Q$ satisfazendo $\|s_i - y_i\|_X \leq d_Q(y_i) + \frac{t_i}{i}$. Nessas condições, é fácil ver que $s_i \rightarrow x_0$, o que, por hipótese, garante a existência de uma sequência $v_i \rightarrow v$ tal que $s_i + t_i v_i \in Q$ para todo $i \in \mathbb{N}$, assim, da continuidade lipschitziana de d_Q com posto 1, segue

$$d_Q(y_i + t_i v) \leq d_Q(s_i + t_i v_i) + \|y_i - s_i\|_X + t_i \|v - v_i\|_X \leq d_Q(y_i) + t_i \left(\|v - v_i\|_X + \frac{1}{i} \right),$$

o que, fazendo $i \uparrow \infty$, acarreta que $d_Q^\circ(x_0; v) = 0$, ou seja, que $v \in \mathfrak{T}(Q; x_0)$. $\square$

Proposição 106. *Se $Q \subseteq X$, para todo $x_0 \in Q$, $\mathfrak{T}(Q; x_0)$ é o cone polar de $[0, \infty[\partial d_Q(x_0)$.*

Demonstração. Chame $\Sigma = [0, \infty[\partial d_Q(x_0)$. Fixado $v \in \mathfrak{T}(Q; x_0)$, para quaisquer $\zeta \in \partial d_Q(x_0)$ e $\alpha \geq 0$, vale $\langle \alpha \zeta, v \rangle = \alpha \langle \zeta, v \rangle \leq \alpha d_Q^\circ(x_0; v) = 0$, o que mostra que $v \in \Sigma^\Delta$. Reciprocamente, fixado $v \in \Sigma^\Delta$, em particular, $\langle \zeta, v \rangle \leq 0$ para todo $\zeta \in \partial d_Q(x_0)$, o que mostra que $d_Q^\circ(x_0; v) \leq 0$, então, como tal derivada é não-negativa, ela deve se anular, o que mostra que $v \in \mathfrak{T}(Q; x_0)$. $\square$

Proposição 107 (página 212 de [11]). *Dados X e Y espaços de Banach, $Q_1 \subseteq X$ e $Q_2 \subseteq Y$, para qualquer $(x_0, y_0) \in X \times Y$, $\mathfrak{T}(Q_1 \times Q_2; (x_0, y_0)) = \mathfrak{T}(Q_1; x_0) \times \mathfrak{T}(Q_2; y_0)$.*

Demonstração. Fixado $(v, w) \in \mathfrak{T}(Q_1 \times Q_2; (x_0, y_0))$, tome sequências $x_i \rightarrow x_0$, $y_i \rightarrow y_0$ e $t_i \downarrow 0$. Da **Proposição 105**, existe $v_i \rightarrow v$ e $w_i \rightarrow w$ tal que $x_i + t_i v_i \in Q_1$ e $y_i + t_i w_i \in Q_2$ para todo $i \in \mathbb{N}$, logo, a mesma proposição garante que $(v, w) \in \mathfrak{T}(Q_1; x_0) \times \mathfrak{T}(Q_2; y_0)$. Agora, fixado $(v, w) \in \mathfrak{T}(Q_1; x_0) \times \mathfrak{T}(Q_2; y_0)$, tome sequências $(x_i, y_i) \rightarrow (x_0, y_0)$ e $t_i \downarrow 0$. Da **Proposição 105**, existe $(v_i, w_i) \rightarrow (v, w)$ de modo que $(x_i + t_i v_i, y_i + t_i w_i) \in Q_1 \times Q_2$ para todo $i \in \mathbb{N}$, portanto, a mesma proposição garante que $(v, w) \in \mathfrak{T}(Q_1 \times Q_2; (x_0, y_0))$. $\square$

Teorema 108 (página 59 de [13]). Sendo $f: X \rightarrow \mathbb{R}$ lipschitziano em uma vizinhança de x_0 , o epígrafo de $f^\circ(x_0; \cdot)$ é $\mathfrak{T}(\text{epi} f; (x_0, f(x_0)))$.

Demonstração. Supondo que $(v, r) \in \mathfrak{T}(\text{epi} f; (x_0, f(x_0)))$, escolha sequências $x_i \rightarrow x_0$ e $t_i \downarrow 0$ de modo que

$$\lim_{i \uparrow \infty} \frac{f(x_i + t_i v) - f(x_i)}{t_i} = f^\circ(x_0; v).$$

Sabendo que $(x_i, f(x_i))$ é uma sequência no epígrafo de f tendendo a $(x_0, f(x_0))$, da **Proposição 105**, existe uma sequência $(v_i, r_i) \rightarrow (v, r)$ de modo que $(x_i, f(x_i)) + t_i(v_i, r_i) \in \text{epi} f$ para todo $i \in \mathbb{N}$, o que acarreta que

$$\frac{f(x_i + t_i v_i) - f(x_i)}{t_i} \leq r_i \quad \forall i \in \mathbb{N}.$$

Fazendo $i \uparrow \infty$, obtemos que $f^\circ(x_0; v) \leq r$.

Reciprocamente, supondo que $v \in X$ e $r \geq f^\circ(x_0; v)$, sejam $\{(x_i, y_i)\} \subseteq \text{epi} f$ tendendo a $(x_0, f(x_0))$ e $t_i \downarrow 0$. Pondo, para cada $i \in \mathbb{N}$,

$$r_i = \sup \left\{ r, \frac{f(x_i + t_i v) - f(x_i)}{t_i} \right\},$$

do Teorema do Confronto, $r_i \rightarrow r$ para $i \rightarrow \infty$ e $y_i + t_i r_i \geq y_i + [f(x_i + t_i v) - f(x_i)] \geq f(x_i + t_i v)$ para todo $i \in \mathbb{N}$, o que mostra que a sequência $(v, r_i) \rightarrow (v, r)$ satisfaz $(x_i, y_i) + t_i(v, r_i) \in \text{epi} f$ para todo $i \in \mathbb{N}$. Portanto, pela **Proposição 105**, $(v, r) \in \mathfrak{T}(\text{epi} f; (x_0, f(x_0)))$. $\square$

4.2.3 Cone hipertangente

Definição 109 (página 57 de [13]). Sendo $Q \subseteq X$ e $x_0 \in Q$, um vetor $v \in X$ é denominado **direção hipertangente** a Q em x_0 se existe $\epsilon > 0$ para o qual $B_Q(x_0, \epsilon) +]0, \epsilon[B_X(v, \epsilon) \subseteq Q$. O conjunto delas será denotado por $\mathfrak{H}(Q; x_0)$.

Propriedade 110. $\mathfrak{H}(Q; x_0)$ é um cone convexo e, fortemente, aberto com vértice na origem que está incluso em $\mathfrak{T}(Q; x_0)$.

Demonstração. Dados $v \in \mathfrak{H}(Q; x_0)$ e $\alpha > 0$, existe $\epsilon > 0$ tal como na **Definição 109**, desse modo,

$$\begin{aligned} \alpha > 1 &\Rightarrow B_Q\left(x_0, \frac{\epsilon}{\alpha}\right) +]0, \frac{\epsilon}{\alpha}[B_X\left(\alpha v, \frac{\epsilon}{\alpha}\right) \subseteq B_Q(x_0, \epsilon) +]0, \frac{\epsilon}{\alpha}[B_X(\alpha v, \alpha\epsilon) \subseteq Q, \\ \alpha < 1 &\Rightarrow B_Q(x_0, \alpha\epsilon) +]0, \alpha\epsilon[B_X(\alpha v, \alpha\epsilon) \subseteq B_Q(x_0, \epsilon) +]0, \frac{\epsilon}{\alpha}[B_X(\alpha v, \alpha\epsilon) \subseteq Q, \end{aligned}$$

o que mostra que $\alpha v \in \mathfrak{H}(Q; x_0)$ e, consequentemente, que tal conjunto é um cone convexo com vértice na origem. Ainda, dado $v \in \mathfrak{H}(Q; x_0)$ e sendo $\epsilon > 0$ como na **Definição 109**, não é difícil ver que, para todo $w \in B_X(v, \epsilon)$, existe um $\delta \in]0, \epsilon[$ para o qual $B_X(w, \delta) \subseteq B_X(v, \epsilon)$ e, consequentemente, para o qual $B_Q(x_0, \delta) +]0, \delta[B_X(w, \delta) \subseteq Q$, o que mostra que $B_X(v, \epsilon) \subseteq \mathfrak{H}(Q; x_0)$ e, consequentemente, que tal conjunto é aberto forte. Por fim, dado $v \in \mathfrak{H}(Q; x_0)$, da **Definição 109**, vemos que $d_Q^\circ(x_0; v) = 0$, o que mostra, pela **Definição 103**, que $v \in \mathfrak{T}(Q; x_0)$. $\square$

Proposição 111 (página 57 de [13]). Se $\mathfrak{H}(Q; x_0)$ não é vazio, então, ele coincide com o interior forte de $\mathfrak{Z}(Q; x_0)$.

Demonstração. Pelo que foi mostrado na **Propriedade 110**, basta mostrar que $\text{int } \mathfrak{Z}(Q; x_0) \subseteq \mathfrak{H}(Q; x_0)$. Dados $v_1 \in \mathfrak{H}(Q; x_0)$ e $v_2 \in \mathfrak{Z}(Q; x_0)$, da **Definição 109** e como consequência da **Proposição 105**, existem $\epsilon_1, \epsilon_2 > 0$ para os quais

$$B_Q(x_0, \epsilon_1) +]0, \epsilon_1[B_X(v_1, \epsilon_1) \subseteq Q \quad (4.1)$$

$$B_Q(x_0, \epsilon_2) +]0, \epsilon_2[B_X\left(v_2, \frac{\epsilon_1}{2}\right) \subseteq Q, \quad (4.2)$$

logo, ponha $\epsilon = \inf\left\{\epsilon_2, \frac{\epsilon_1}{2}, \frac{\epsilon_1}{1+\epsilon_1+\|v_2\|}\right\}$.

Afirmção. $B_Q(x_0, \epsilon) +]0, \epsilon[B_X(v_1 + v_2, \epsilon) \subseteq Q$.

De fato, sendo v um elemento do lado esquerdo da inclusão, existem $x \in B_Q(x_0, \epsilon)$, $t \in]0, \epsilon[$ e $u \in B_X(0, 1)$ para os quais $v = x + t(v_1 + v_2 + \epsilon u)$, o fato de que $\epsilon \leq \epsilon_2$ permite afirmar, de (4.2), que existe $w \in B_X(0, 1)$ para o qual $y + t\left(v_2 - \frac{\epsilon_1}{2}w\right) \in Q$, e esse vetor pertence a $B_Q(x_0, \epsilon_1)$, pois

$$\left|x + t\left(v_2 - \frac{\epsilon_1}{2}w\right) - x_0\right| \leq \epsilon + \epsilon(\epsilon_1 + \|v_2\|) < \epsilon_1,$$

consequentemente, (4.1) garante que $x + t\left[\left(v_2 - \frac{\epsilon_1}{2}w\right) + B_X(v_1, \epsilon_1)\right] \subseteq Q$. A partir disso, como é possível escrever

$$v = x + t\left[\left(v_2 - \frac{\epsilon_1}{2}w\right) + \underbrace{v_1 + \epsilon u + \frac{\epsilon_1}{2}w}_{\in B_X(v_1, \epsilon_1)}\right],$$

a conclusão é que $v \in Q$.

A partir disso, segue que $\mathfrak{H}(Q; x_0) + \mathfrak{Z}(Q; x_0) \subseteq \mathfrak{H}(Q; x_0)$. Finalmente, dado $v \in \text{int } \mathfrak{Z}(Q; x_0)$, qualquer $w \in \mathfrak{H}(Q; x_0)$ admite um $\lambda > 0$ para o qual $v - \lambda w \in \mathfrak{Z}(Q; x_0)$, portanto, como $\lambda \in \mathfrak{H}(Q; x_0)$, vale $v \in \mathfrak{H}(Q; x_0)$. $\square$

4.2.4 Cone normal estrito

Definição 112 (página 21 de [11]). Dados $Q \subseteq X$ e $x_0 \in Q$, o **cone normal estrito a Q em x_0** é o polar de $\mathfrak{C}(Q; x_0)$. Ele será denotado por $\hat{\mathfrak{N}}(Q; x_0)$.

Proposição 113 (página 30 de [11]). Se $Q \subseteq X$ é convexo e $x_0 \in Q$, então, $\hat{\mathfrak{N}}(Q; x_0) = (Q - x_0)^\Delta$.

Demonstração. Sabendo que $(Q - x_0)^\Delta$ é um cone, a **Proposição 100** e as Propriedades 42 e 44 garantem que

$$\hat{\mathfrak{N}}(Q; x_0) = \bigcap_{t>0} \frac{(Q - x_0)^\Delta}{t} = (Q - x_0)^\Delta.$$

$\square$

4.2.5 Cone normal

Definição 114 (página 212 de [11]). Sendo $Q \subseteq X$ e $x_0 \in Q$, o **cone normal a Q em x_0** é o polar de $\mathfrak{Z}(Q; x_0)$. Ele será denotado por $\mathfrak{N}(Q; x_0)$.

Proposição 115 (página 213 de [11]). *Dados $Q \subseteq X$ e $x_0 \in Q$, vale que $\mathfrak{N}(Q; x_0) = \overline{[0, \infty[\partial d_Q(x_0)}$ e $\mathfrak{Z}(Q; x_0) = [\mathfrak{N}(Q; x_0)]^\Delta$.*

Demonstração. Sabendo que $\partial d_Q(x_0)$ é, por definição, convexo, o conjunto $\Sigma := [0, \infty[\partial d_Q(x_0)$ é um cone convexo com vértice na origem, então, a primeira afirmação segue da **Proposição 106** e da **Propriedade 45**. Já a segunda segue da primeira e da **Propriedade 42**. $\square$

Proposição 116. *Dados X e Y espaços de Banach, $Q_1 \subseteq X$ e $Q_2 \subseteq Y$, para qualquer $(x_0, y_0) \in X \times Y$, $\mathfrak{N}(Q_1 \times Q_2; (x_0, y_0)) = \mathfrak{N}(Q_1; x_0) \times \mathfrak{N}(Q_2; y_0)$.*

Proposição 117 (página 61 de [13]). *Um elemento $\zeta \in X^*$ pertence a $\partial f(x_0)$ se, e somente se, $(\zeta, -1) \in \mathfrak{N}(\text{epi } f; (x_0, f(x_0)))$.*

Demonstração. Por definição, $\zeta \in \partial f(x_0)$ se, e somente se, $f^\circ(x_0; v) \geq \langle v, \zeta \rangle$ para todo $v \in X$, e isto equivale a dizer que, para todos $v \in X$ e $r \geq f^\circ(x_0; v)$, vale $\langle (v, r); (\zeta, -1) \rangle \leq 0$. Do **Teorema 108**, isto equivale a dizer que, para todos $(v, r) \in \mathfrak{Z}(\text{epi } f; (x_0, f(x_0)))$, $\langle (v, r); (\zeta, -1) \rangle \leq 0$, o que equivale, pela **Definição 114**, a dizer que

$$(\zeta, -1) \in \mathfrak{N}(\text{epi } f; (x_0, f(x_0))).$$

$\square$

Definição 118. Dado $Q \subseteq X$, a **função indicadora de Q** é definida como

$$\psi_Q(x) = \begin{cases} 0 & \text{se } x \in Q, \\ \infty & \text{caso contrário.} \end{cases}$$

Proposição 119. *Sendo $Q \subseteq X$ e $x_0 \in Q$, o cone normal a Q em x_0 é o gradiente generalizado de ψ_Q em x_0 .*

Demonstração. Pela **Proposição 117**, $\zeta \in \partial \psi_Q(x_0)$ se, e somente se, $(\zeta, -1) \in \mathfrak{N}(\text{epi } \psi_Q; (x_0, 0))$, desse modo, como $\text{epi } \psi_Q = Q \times [0, \infty[$, a **Proposição 116** garante a equivalência entre este pertencimento e $\zeta \in \mathfrak{N}(Q; x_0)$ e $-1 \in \mathfrak{N}([0, \infty[; 0) =]-\infty, 0]$, o que prova o resultado. $\square$

4.2.6 Relações entre os cones

Proposição 120. *Se $Q \subseteq X$ e $x_0 \in Q$, então, $\mathfrak{Z}(Q; x_0) \subseteq \mathfrak{C}(Q; x_0)$.*

Demonstração. Segue da **Definição 98** e das Proposições 105 e 184. $\square$

Corolário 121. *Se $Q \subseteq X$ e $x_0 \in Q$, então, $\mathfrak{N}(Q; x_0) \supseteq \hat{\mathfrak{N}}(Q; x_0)$.*

Teorema 122 (adaptado da página 130 de [4]). *Se Q é fechado e $x_0 \in Q$, então,*

$$\mathfrak{T}(Q; x_0) = \liminf_{x \rightarrow x_0} \mathfrak{C}(Q; x).$$

Demonstração. Dado $v \in \mathfrak{T}(Q; x_0)$, seja $x_i \rightarrow x_0$. Dada $t_i \downarrow 0$, é possível encontrar $\epsilon_i \downarrow 0$ de modo que $t \in]0, t_i] \Rightarrow d_Q(x_i + tv) < \epsilon_i t$, logo, para todo $i \in \mathbb{N}$,

$$\liminf_{\delta \downarrow 0} \left\{ \frac{d_Q(x_i + tv)}{t} \mid t \in]0, \delta] \right\} = \liminf_{t \downarrow 0} \frac{d_Q(x_i + tv)}{t} < \epsilon_i.$$

Sabendo que, para todo $j \in \mathbb{N}$,

$$\inf \left\{ \frac{d_Q(x_i + tv)}{t} \mid i \geq j, \quad t \in]0, \delta] \right\} \leq \inf \left\{ \frac{d_Q(x_j + tv)}{t} \mid t \in]0, \delta] \right\},$$

segue, do Teorema do Confronto, que

$$\liminf_{i \uparrow \infty, t \downarrow 0} \frac{d_Q(x_i + tv)}{t} = 0,$$

portanto, pela **Proposição 183**, vale que $v \in \liminf_{x \rightarrow x_0} \mathfrak{C}(Q; x)$.

Para mostrar a inclusão recíproca, tome $v \in \liminf_{x \rightarrow x_0} \mathfrak{C}(Q; x)$ unitário. Para todo $\epsilon > 0$, existe $i_0 \in \mathbb{N}$ tal que

$$i \geq i_0, \quad x \in B_Q \left[x_0, \frac{1}{i} \right] \Rightarrow B_X \left[v, \frac{\epsilon}{2} \right] \cap \mathfrak{C}(Q; x) \neq \emptyset. \quad (4.3)$$

Supondo que $v \notin \mathfrak{T}(Q; x_0)$, é possível encontrar $\epsilon_0 \in]0, \frac{1}{4}[$ e sequências $\{x_i\} \subseteq B_Q[0, \frac{1}{2i}]$ e $\{t_i\} \subseteq]0, \frac{1}{4i}[$ de modo que cada conjunto $B_X[x_i + t_i v, t_i \epsilon_0]$ é disjunto de Q . Agora, fixando $\iota \in \mathbb{N}$ de modo que (4.3) seja satisfeita para ι e ϵ_0 , ponha

$$K := Q \cap \left(\bigcup_{a \in [0, t_\iota]} B_X[x_\iota + av, a\epsilon_0] \right).$$

Do **Teorema 59** aplicado a $x \mapsto -\|x - x_\iota\|_X$, existe $z \in K$ de modo que qualquer $x \in K$ satisfaz

$$\|x - x_\iota\|_X \leq \|z - x_\iota\|_X + \frac{\epsilon_0}{1 + \epsilon_0} \|z - x\|_X. \quad (4.4)$$

Da disjunção entre $B_X[x_\iota + t_\iota v, t_\iota \epsilon_0]$ e K , é possível afirmar que z pode ser escrito como $z = x_\iota + (1 - \alpha)t_\iota w_\iota$ com $\alpha \in]0, 1]$ e $\|w_\iota - v\|_X \leq \epsilon_0$.

Afirmação.

$$\bigcup_{a \in [0, \alpha t_\iota]} B_X[z + av, a\epsilon_0] \subseteq \bigcup_{a \in [0, t_\iota]} B_X[x_\iota + av, a\epsilon_0].$$

Claramente, dados $a \in [0, \alpha t_\iota]$ e $y \in B_X[z + av, a\epsilon_0]$, pondo $b = a + (1 - \alpha)t_\iota \in [0, t_\iota]$, vale que

$$\begin{aligned} \|y - (x_\iota + bv)\|_X &\leq \|(z + av) - (x_\iota + bv)\|_X + \|y - (z + av)\|_X \\ &\leq \|(1 - \alpha)t_\iota w_\iota + (a - b)v\|_X + a\epsilon_0 \leq a\epsilon_0 + (1 - \alpha)t_\iota \epsilon_0 = b\epsilon_0, \end{aligned}$$

o que mostra que $y \in B_X[x_i + bv, b\epsilon_0]$.

Ainda, como $\|z - x_i\|_X \leq (1 - \alpha)t_i(1 + \epsilon_0)$, segue que

$$\|z - x_0\|_X \leq (1 - \alpha)t_i(1 + \epsilon_0) + \|x_i - x_0\|_X \leq \frac{1}{4t} \left(1 + \frac{1}{4}\right) + \frac{1}{2t} = \frac{13}{16t} < \frac{1}{t},$$

então, (4.3) garante que a existência de $u \in \mathfrak{C}(Q; z) \cap B\left[v, \frac{\epsilon_0}{2}\right]$. Da definição do cone contingente, existem sequências $u_j \rightarrow u$ e $s_j \downarrow 0$ de modo que $z + k_j u_j \in Q$ para todo $j \in \mathbb{N}$, e, da afirmação provada, $z + k_j u_j \in B_X[x_i + t_i v, t_i \epsilon_0]$ para j grande o suficiente, assim, existe $p \in \mathbb{N}$ para o qual $z + k_p u_p \in K$ e $\|u_p - v\|_X \leq \epsilon_0$. Substituindo $x = z + k_p u_p$ em (4.4), vem

$$\begin{aligned} [(1 - \alpha)t_i + k_p]\|w_i\|_X - k_p\|u_p - w_i\|_X &\leq \|[(1 - \alpha)t_i + k_p]w_i - k_p(u_p - w_i)\|_X \\ &= \|(z + k_p u_p) - x_i\|_X \\ &\leq (1 - \alpha)t_i\|w_i\|_X + \frac{\epsilon_0}{1 + \epsilon_0}k_p\|u_p\|_X, \end{aligned}$$

ou seja, $k_p(1 + \epsilon_0)(\|w_i\|_X - \|u_p - w_i\|_X) \leq k_p\epsilon_0\|u_p\|_X$. Sabendo que $\|u_p - w_i\|_X \leq 2\epsilon_0$ e que $\|u_p\|_X \leq 1 + \epsilon_0$, obtemos que $1 - \epsilon_0 \leq \|w_i\|_X \leq 3\epsilon_0$, porém, isso contradiz o fato de que $4\epsilon_0 < 1$. Com a contradição, fica mostrado que $v \in \mathfrak{Z}(Q; x_0)$. $\square$

Definição 123 (página 215 de [11]). Dados $Q \subseteq X$ e $x_0 \in Q$, o conjunto Q é dito **regular em x_0** se $\mathfrak{Z}(Q; x_0) = \mathfrak{C}(Q; x_0)$.

Proposição 124 (página 215 de [11]). Dados $Q \subseteq X$ e $x_0 \in Q$, Q é regular em x_0 se, e somente se, $\mathfrak{N}(Q; x_0) = \widehat{\mathfrak{N}}(Q; x_0)$.

Demonstração. Supondo que Q seja regular em x_0 , segue que $\mathfrak{Z}(Q; x_0) = \mathfrak{C}(Q; x_0)$, portanto, tomar polaridade implica que $\mathfrak{N}(Q; x_0) = \widehat{\mathfrak{N}}(Q; x_0)$. Reciprocamente, se $\mathfrak{N}(Q; x_0) = \widehat{\mathfrak{N}}(Q; x_0)$, as Proposições 115 e 120 e a **Propriedade 45** garantem que

$$\mathfrak{Z}(Q; x_0) = [\mathfrak{N}(Q; x_0)]^\Delta = [\widehat{\mathfrak{N}}(Q; x_0)]^\Delta = \{[\mathfrak{C}(Q; x_0)]^\Delta\}^\Delta \supseteq \mathfrak{C}(Q; x_0) \supseteq \mathfrak{Z}(Q; x_0),$$

o que mostra que $\mathfrak{Z}(Q; x_0) = \mathfrak{C}(Q; x_0)$ e, consequentemente, que Q é regular em x_0 . $\square$

Propriedade 125 (página 215 de [11]). Se Q é convexo e fechado e $x_0 \in Q$, então, Q é regular em x_0 e

$$\begin{aligned} \mathfrak{Z}(Q; x_0) &= \mathfrak{C}(Q; x_0) = \overline{\bigcup_{t>0} \frac{Q - x_0}{t}}, \\ \mathfrak{N}(Q; x_0) &= \widehat{\mathfrak{N}}(Q; x_0) = (Q - x_0)^\Delta. \end{aligned}$$

Demonstração. Começamos provando que $\partial d_Q(x_0) \subseteq \widehat{\mathfrak{N}}(Q; x_0)$. De fato, sabendo que d_Q é convexa, dado $\zeta \in \partial d_Q(x_0)$, temos que $\langle x - x_0, \zeta \rangle \leq d_Q(x) - d_Q(x_0) = 0$ para todo $x \in Q$, assim, da **Proposição 113**, $\zeta \in \widehat{\mathfrak{N}}(Q; x_0)$. A partir disso, como $\widehat{\mathfrak{N}}(Q; x_0)$ é cone fechado, a Proposição 115

garante que $\mathfrak{N}(Q; x_0) \subseteq \widehat{\mathfrak{N}}(Q; x_0)$, o que, em conjunto com o **Corolário 121**, mostra que ambos são iguais, portanto, da **Proposição 124**, Q é regular em x_0 .

As demais igualdades seguem por polaridade e das Proposições 113 e 100. $\square$

4.3 Formulando o formalismo

4.3.1 Diferenciais gerais de funcionais com valores na reta estendida

Nesta subseção, os resultados serão formulados supondo que X é um espaço de Banach real, e será usada a notação $(x, \alpha) \downarrow_f x_0$ para indicar que $(x, \alpha) \in \text{epi} f$ e $(x, \alpha) \rightarrow (x_0, f(x_0))$.

Teorema 126 (página 96 de [13]). *Seja $f: X \rightarrow [-\infty, \infty]$ finito em x_0 , o cone tangente ao epígrafo de f em $(x_0, f(x_0))$ é o epígrafo do funcional $f^\circ(x_0; \cdot): X \rightarrow [-\infty, \infty]$ definido por*

$$f^\circ(x_0; v) = \lim_{\epsilon \downarrow 0} \limsup_{\substack{(x, \alpha) \downarrow_f x_0 \\ t \downarrow 0}} \inf \left\{ \frac{f(x + tw) - \alpha}{t} \mid w \in B_X(v, \epsilon) \right\}. \quad (4.5)$$

Demonstração. Em primeiro lugar, seja $(v, r) \in \mathfrak{T}(\text{epi} f, (x_0, f(x_0)))$. Dados $(x_i, \alpha_i) \downarrow_f x_0$ e $t_i \downarrow 0$, da **Proposição 105**, existe uma sequência (v_i, r_i) tendendo a (v, r) para a qual $(x_i, \alpha_i) + t_i(v_i, r_i) \in \text{epi} f$ para todo $i \in \mathbb{N}$, isto é, satisfazendo $\alpha_i + t_i r_i \geq f$ para todo $i \in \mathbb{N}$. A partir disso, dado $\epsilon > 0$, como existe $i_0 \in \mathbb{N}$ para o qual $i > i_0 \Rightarrow v_i \in B_X(v, \epsilon)$, $r_i \in]r - \epsilon, r + \epsilon[$, segue que

$$i > i_0 \Rightarrow \inf \left\{ \frac{f(x_i) + t_i w) - \alpha_i}{t_i} \mid w \in B_X(v, \epsilon) \right\} \leq r_i < r + \epsilon.$$

Isso mostra que, para todo $\epsilon > 0$,

$$\limsup_{\substack{(x, \alpha) \downarrow_f x_0 \\ t \downarrow 0}} \inf \left\{ \frac{f(x + tw) - \alpha}{t} \mid w \in B_X(v, \epsilon) \right\} \leq r,$$

portanto, $f^\circ(x_0; v) \leq r$.

Reciprocamente, supondo que $v \in X$ e $r \in \mathbb{R}$ sejam tais que $f^\circ(x_0; v) \leq r$, sejam $(x_i, \alpha_i) \downarrow_f x_0$ e $t_i \downarrow 0$. Para cada $n \in \mathbb{N}$, existe um $\epsilon_n < \frac{1}{n}$ para o qual

$$\limsup_{\substack{(x, \alpha) \downarrow_f x_0 \\ t \downarrow 0}} \inf \left\{ \frac{f(x + tw) - \alpha}{t} \mid w \in B_X(v, \epsilon_n) \right\} \leq r + \frac{1}{n},$$

desse modo, como $(x_i, \alpha_i) \downarrow_f x_0$ e $t_i \downarrow 0$, deve existir um índice $i(n)$ para o qual

$$\inf \left\{ \frac{f(x_{i(n)} + t_{i(n)} w) - \alpha_{i(n)}}{t_{i(n)}} \mid w \in B_X(v, \epsilon_n) \right\} \leq r + \frac{2}{n},$$

e, conseqüentemente, um vetor $v_{i(n)} \in B_X(v, \epsilon_n)$ para o qual

$$\frac{f(x_{i(n)} + t_{i(n)} v_{i(n)}) - \alpha_{i(n)}}{t_{i(n)}} \leq r + \frac{3}{n}.$$

Definindo, agora,

$$r_{i(n)} = \sup \left\{ r, \frac{f(x_{i(n)} + t_{i(n)}v_{i(n)}) - \alpha_{i(n)}}{t_{i(n)}} \right\},$$

segue que $\alpha_{i(n)} + t_{i(n)}r_{i(n)} \geq f(x_{i(n)} + t_{i(n)}v_{i(n)})$ para todo $n \in \mathbb{N}$ e que $(v_{i(n)}, r_{i(n)}) \rightarrow (v, r)$, portanto, a **Proposição 105** garante que $(v, r) \in \mathfrak{Z}(\text{epi } f; (x_0, f(x_0)))$. $\square$

Definição 127 (página 61 de [13]). Dados $f: X \rightarrow [-\infty, \infty]$ e $x_0 \in X$ um ponto no qual f assume valor finito, a **diferencial geral de f em x_0** , denotada por $\partial f(x_0)$, é o conjunto dos $\zeta \in X^*$ para os quais $(\zeta, -1) \in \mathfrak{N}(\text{epi } f; (x_0, f(x_0)))$.

Corolário 128 (página 97 de [13]). $\partial f(x_0) = \emptyset$ se, e somente se, $f^\circ(x_0; 0) = -\infty$. Caso contrário,

$$\begin{aligned} \partial f(x_0) &= \{\zeta \in X^* \mid \langle v, \zeta \rangle \leq f^\circ(x_0; v) \ \forall v \in X\}, \\ f^\circ(x_0; v) &= \sup \langle v, \partial f(x_0) \rangle. \end{aligned}$$

Demonstração. Do teorema e definição anteriores, é possível afirmar que $\zeta \in \partial f(x_0)$ se, e somente, $\langle v, \zeta \rangle \leq r$ para todos $(v, r) \in \mathfrak{Z}(\text{epi } f; (x_0, f(x_0))) = \text{epi } f^\circ(x_0; \cdot)$, ou seja, para todos $(v, r) \in X \times [-\infty, \infty]$ para os quais $f^\circ(x_0; v) \leq r$. No caso em que $f^\circ(x_0; v) = -\infty$, não existe ζ que satisfaça tais condições e, conseqüentemente, $\partial f(x_0) = \emptyset$, e, como $f^\circ(x_0; \cdot)$ é, positivamente, 1-homogênea e semicontínua inferior, ela vale $-\infty$ em algum vetor $v \in X$ se, e somente se, o faz no vetor nulo, o que mostra que $\partial f(x_0) = \emptyset \Leftrightarrow f^\circ(x_0; 0) = -\infty$.

Agora, quando $f^\circ(x_0; \cdot) > g \equiv -\infty$, em particular, temos que $\zeta \in \partial f(x_0)$ se, e somente, $\langle v, \zeta \rangle \leq f^\circ(x_0; v)$ para todo $v \in X$, e isso equivale a dizer que $f^\circ(x_0; \cdot)$ é o funcional de suporte de $\partial f(x_0)$. $\square$

Definição 129 (página 98 de [13]). O funcional f é dito, **lipschitziano em x_0 e na direção $v_0 \in X$** se é finito em x_0 e a quantidade

$$f^+(x_0; v_0) := \limsup_{\substack{(x, \alpha) \downarrow f x_0 \\ v \rightarrow v_0 \\ t \downarrow 0}} \frac{f(x + tv) - \alpha}{t}$$

não é ∞ . Caso exista uma direção $v \in X$ para o qual isso acontece, f será dito, **direcionalmente, lipschitziano em x_0** .

O conjunto das direções $v \in X$ de modo que f é lipschitziano em x_0 e em v será denotado por $\mathfrak{L}(f; x_0)$.

Proposição 130 (página 98 de [13]). *O funcional f é, direcionalmente, lipschitziano em x_0 em uma direção v_0 se, e somente se, existe $\beta \in \mathbb{R}$ para o qual (v_0, β) é hipertangente a $\text{epi } f$ em $(x_0, f(x_0))$.*

Demonstração. Supondo, em primeiro lugar, que (v_0, β) seja hipertangente a $\text{epi } f$ em $(x_0, f(x_0))$, por definição, existe $\epsilon > 0$ para o qual

$$B_{\text{epi } f}((x_0, f(x_0)); \epsilon) +]0, \epsilon[B_{X \times \mathbb{R}}((v_0, \beta); \epsilon) \subseteq \text{epi } f,$$

assim, sempre que $(x, \alpha) \in \text{epi } f$ está próximo a $(x_0, f(x_0))$ e v está próximo a v_0 , vale $\alpha + t\beta \geq f(x + tv)$, o que mostra que $f^+(x_0; v_0) < \beta$ e, consequentemente, que f é, direcionalmente, lipschitziano em x_0 na direção v_0 .

Reciprocamente, supondo que f seja, direcionalmente, lipschitziano em x_0 em uma direção v_0 e sendo $\beta > f^+(x_0; v_0)$, para todo $(x, \alpha) \in \text{epi } f$ próximo a $(x_0, f(x_0))$, para todo v próximo a v_0 e para todo $t > 0$ próximo a zero, vale

$$\frac{f(x + tv) - \alpha}{t} < \beta,$$

o que mostra que (v_0, β) é hipertangente a $\text{epi } f$ em $(x_0, f(x_0))$. $\square$

Corolário 131 (página 99 de [13]). *Se $f^+(x_0; v) < \infty$, então, essa quantidade coincide com o ínfimo dos $\beta \in \mathbb{R}$ para os quais (v, β) é hipertangente a $\text{epi } f$ em $(x_0, f(x_0))$.*

Teorema 132 (página 100 de [13]). *Sendo f finito em x_0 e supondo que $\mathfrak{L}(f; x_0) \neq \emptyset$, este é o interior das direções para as quais a derivada direcional geral de f em x_0 é finita e $f^\circ(x_0; \cdot)$ é contínua e coincide com $f^+(x_0; \cdot)$ em tal conjunto.*

Demonstração. Como $\mathfrak{H}(\text{epi } f; (x_0, f(x_0)))$ é não-vazio por suposição, a **Proposição 130** garante que $\mathfrak{L}(f; x_0) = \Pi_X(\mathfrak{H}(\text{epi } f; (x_0, f(x_0))))$, e o **Teorema 126** garante que

$$\mathfrak{H}(\text{epi } f; (x_0, f(x_0))) = \text{int } \mathfrak{Z}(\text{epi } f; (x_0, f(x_0))),$$

o que, significa que, para mostrar a primeira afirmação do teorema, é suficiente mostrar que

$$\Pi_X(\mathfrak{H}(\text{epi } f; (x_0, f(x_0)))) = \text{int}\{v \in X \mid f^\circ(x_0; v) < \infty\}.$$

Se v_0 pertence à projeção, existe $\beta \in \mathbb{R}$ para o qual $(v_0, \beta) \in \mathfrak{H}(\text{epi } f; (x_0, f(x_0)))$, assim, o **Teorema 126** garante que $(v_0, \beta) \in \text{int } \text{epi } f^\circ(x_0; \cdot)$. A partir disso, $f^\circ(x_0; v) < \infty$ para todo v próximo o suficiente a v_0 , o que mostra que v_0 pertence ao interior no lado direito.

Reciprocamente, se v_0 pertence ao interior no lado direito, existem $\epsilon > 0$ e $\beta_0 \in \mathbb{R}$ para os quais $v \in B_X[v_0, \epsilon] \Rightarrow f^\circ(x_0; v) < \beta_0 - 1$, o que mostra que todo (v, β) próximo a (v_0, β_0) pertence a $\text{epi } f^\circ(x_0; \cdot) = \mathfrak{Z}(\text{epi } f; (x_0, f(x_0)))$, o que mostra que $(v_0, \beta_0) \in \mathfrak{H}(\text{epi } f; (x_0, f(x_0)))$.

Agora, sabendo que existe v_0 para o qual $f^\circ(x_0; v_0) < \infty$, trata-se de um funcional convexo que é limitado em uma vizinhança de v_0 , logo, a **Proposição 153** garante que $f^\circ(x_0; \cdot)$ é contínuo em $\mathfrak{L}(f; x_0)$.

Finalmente, sabendo que $\mathfrak{Z}(\text{epi } f; (x_0, f(x_0))) = \text{epi } f^\circ(x_0; \cdot)$ e usando o **Corolário 131**, obtemos que

$$\begin{aligned} f^\circ(x_0; v) &= \inf \left\{ \beta \in \mathbb{R} \mid (v, \beta) \in \mathfrak{Z}(\text{epi } f; (x_0, f(x_0))) \right\}, \\ f^+(x_0; v) &= \inf \left\{ \beta \in \mathbb{R} \mid (v, \beta) \in \mathfrak{H}(\text{epi } f; (x_0, f(x_0))) \right\}, \end{aligned}$$

logo, a igualdade $\mathfrak{H}(\text{epi} f, (x_0, f(x_0))) = \text{int} \mathfrak{T}(\text{epi} f, (x_0, f(x_0)))$ garante que $f^+(x_0; v) = f^\circ(x_0; v)$. $\square$

Propriedade 133 (página 103 de [13]). *Seja f_1, f_2 finitos em x_0 e f_2 , direcionalmente, lipschitzianos neste, se*

$$\{v \mid f_1^\circ(x_0; v) < \infty\} \cap \text{int}\{v \mid f_2^\circ(x_0; v) < \infty\} \neq \emptyset, \quad (4.6)$$

então, $(f_1 + f_2)^\circ(x_0; v) \leq f_1^\circ(x_0; v) + f_2^\circ(x_0; v)$.

Demonstração. Em primeiro lugar, denotaremos por $f_0 = f_1 + f_2$ e $\ell_i = f_i^\circ(x_0; \cdot)$ para todo $i \in \{0 : 2\}$. Dito isso, sendo $D_i = \ell_i^{-1}([-\infty, \infty[)$ para todo $i \in \{0 : 2\}$, $v_0 \in D_1 \cap \text{int} D_2$ e $\beta > \ell_2(v_0)$, do **Teorema 132**, para algum $\delta > 0$,

$$\left. \begin{array}{l} t \in]0, \delta[, \\ x \in B_X[x_0, \delta], \\ v \in B_X[v, \delta], \quad \alpha_2 \in [f_2(v), f_2(v) + \delta[\end{array} \right\} \Rightarrow \frac{f_2(x + tv) - \alpha_2}{t} < \beta.$$

Sabendo que, da **Definição 129**,

$$\ell_0(v_0) = \lim_{\epsilon \downarrow 0} \left(\limsup_{\substack{(x, \alpha_1) \downarrow_{f_1} x_0 \\ (x, \alpha_2) \downarrow_{f_2} x_0 \\ t \downarrow 0}} \inf \left\{ \frac{f_1(x + tv) - \alpha_1}{t} + \frac{f_2(x + tv) - \alpha_2}{t} \mid w \in B_X[v, \epsilon] \right\} \right),$$

é possível afirmar que

$$\ell_0(x_0; v_0) \leq \lim_{\epsilon \downarrow 0} \left(\limsup_{\substack{(x, \alpha_1) \downarrow_{f_1} x_0 \\ (x, \alpha_2) \downarrow_{f_2} x_0 \\ t \downarrow 0}} \inf \left\{ \frac{f_1(x + tv) - \alpha_1}{t} + \beta \mid w \in B_X[v, \epsilon] \right\} \right) = \ell_1(v_0) + \beta,$$

logo, fazer $\beta \downarrow \ell_2(v_0)$ fornece $\ell_0(v_0) \leq \ell_1(v_0) + \ell_2(v_0)$.

Agora, retiramos a hipótese de que v_0 é elemento da interseção em (4.6). Se $\ell_1(v_0)$ ou $\ell_2(v_0)$ assumem valor ∞ , a desigualdade é, claramente, válida. Já no caso em que ambas as derivadas são finitas, $v_0 \in D_1 \cap D_2$ e, por (4.6), vale $D_1 \cap \text{int} D_2 \ni \tilde{v}$, assim, da **Propriedade 151**, $(1 - \epsilon)v_0 + \tilde{v}$ para $\epsilon \in]0, 1[$, o que mostra que

$$\ell_0((1 - \epsilon)v_0 + \tilde{v}) \leq \ell_1((1 - \epsilon)v_0 + \tilde{v}) + \ell_2((1 - \epsilon)v_0 + \tilde{v}).$$

Por fim, a convexidade e semicontinuidade inferior da derivada direcional permite fazer $\epsilon \downarrow 0$ e concluir que $\ell_0(v_0) \leq \ell_1(v_0) + \ell_2(v_0)$. $\square$

Teorema 134 (página 102 de [13]). *Nas condições da Propriedade 133, $\partial(f_1 + f_2)(x_0) \subseteq \partial f_1(x_0) + \partial f_2(x_0)$.*

Demonstração. Se $\ell_1(0)$ ou $\ell_2(0)$ assume valor $-\infty$, então, o mesmo vale para $\ell_0(0)$, o que mostra, pelo **Corolário 128**, que os gradientes generalizados de f_1, f_2 e da soma em x_0 são vazios, o que garante o resultado. Senão, sabendo que as derivadas direcionais gerais são sublineares, vale $\ell_1(0) = \ell_2(0) = 0$, o que mostra que

$$\begin{aligned}\partial f_0(x_0) &= \{\zeta \in X^* \mid \langle v, \zeta \rangle \leq \ell_0(v) \forall v \in X\} = \{\zeta \in X^* \mid \langle v, \zeta \rangle \leq (\ell_1 + \ell_2)(v) \forall v \in X\} \\ &= \partial(\ell_1 + \ell_2)(0).\end{aligned}$$

Utilizando (4.6), é possível afirmar que $\partial(\ell_1 + \ell_2)(0) = \partial\ell_1(0) + \partial\ell_2(0) = \partial f_1(x_0) + \partial f_2(x_0)$, o que, em conjunto com a cadeia de desigualdades acima, mostra que $\partial(f_1 + f_2)(x_0) \subseteq \partial f_1(x_0) + \partial f_2(x_0)$. $\square$

Corolário 135 (página 105 de [13]). *Seja $Q_1, Q_2 \subseteq X$ e $x_0 \in Q_1 \cap Q_2$, suponha que Q_2 admita, ao menos, uma direção v hipertangente em x_0 e que $\mathfrak{T}(Q_1; x_0)$ intercepte $\mathfrak{H}(Q_2; x_0)$. Nessas condições,*

$$\mathfrak{T}(Q_1 \cap Q_2; x_0) \supseteq \mathfrak{T}(Q_1; x_0) \cap \mathfrak{T}(Q_2; x_0), \quad (4.7)$$

$$\mathfrak{N}(Q_1 \cap Q_2; x_0) \subseteq \mathfrak{N}(Q_1; x_0) + \mathfrak{N}(Q_2; x_0). \quad (4.8)$$

Demonstração. Sabemos que

- Pela **Proposição 130**, ψ_{Q_2} é, direcionalmente, lipschitziano;
- Pela **Propriedade 110**, o interior forte de $\mathfrak{T}(Q_2; x_0)$ é $\mathfrak{H}(Q_2; x_0)$;
- Pelo **Teorema 126**, $\psi_{Q_i}^\circ(x_0; v) < 0$ se, e somente se, $v \in \mathfrak{T}(Q_i; x_0)$ para todo $i \in \{1, 2\}$.

A partir disso, aplicando o **Teorema 134** a ψ_{Q_1} e ψ_{Q_2} , a **Proposição 119** garante (4.8). Já (4.7) é garantida aplicando polaridade a (4.8) e usando a **Proposição 115** e a **Propriedade 41**. $\square$

4.3.2 Resultado principal

Proposição 136 (adaptado da página 76 de [50]). *Seja $Q_1, \dots, Q_m \subseteq X$, $x_0 \in Q := \bigcap_{i=1}^m Q_i$ e supondo que cada um dos conjuntos Q_i para $i \in \{2 : m\}$ admita, ao menos, uma direção hipertangente em x_0 e*

$$\mathfrak{T}(Q_1; x_0) \cap \left[\bigcap_{i=2}^m \mathfrak{H}(Q_i; x_0) \right] \neq \emptyset,$$

se $\mathfrak{T}(Q_1; x_0), \dots, \mathfrak{T}(Q_m; x_0)$ são inseparáveis e, ao menos, um deles não é subespaço, então, x_0 é ponto de acumulação de Q .

Demonstração. Se x_0 for ponto isolado de Q , então, $\mathfrak{T}(Q; x_0) = \{0\}$, e a inseparabilidade garantirá, pela **Propriedade 47**, que $\bigcap_{i=1}^m \text{ri } \mathfrak{T}(Q_i; x_0) \neq \emptyset$. Mais ainda, tal interseção não contém a origem pelo fato de que, ao menos, um cone tangente não é subespaço, porém, do **Corolário 135**, o vetor

nulo é o único elemento comum aos cones $\mathfrak{Z}(Q_1; x_0), \dots, \mathfrak{Z}(Q_m; x_0)$. Essa contradição mostra que x_0 deve ser ponto de acumulação de Q_{m+1} . $\square$

Como consequência desse resultado e do **Teorema 49**, temos o último e principal resultado do capítulo. Para tal, consideraremos um funcional $f: X \rightarrow [-\infty, \infty]$, conjuntos $C_1, \dots, C_n \subseteq X$ e $x_0 \in C := \bigcap_{i=1}^n C_i$.

Teorema 137. *Se f é direcionalmente, lipschitziano em x_0 , suponha que cada um dos conjuntos C_i para $i \in \{1 : n\}$ admita, ao menos, uma direção hipertangente em x_0 e*

$$\mathfrak{Z}(f^{-1}([-\infty, f(x_0)]); x_0) \cap \left[\bigcap_{i=1}^n \mathfrak{H}(C_i; x_0) \right] \neq \emptyset.$$

Nessas condições, se x_0 minimiza, local e estritamente, f em C , então, existem $y_0, \dots, y_m$ não todos nulos cuja somatória é nula e de modo que $y_0 \in \mathfrak{N}(f^{-1}([-\infty, f(x_0)]); x_0)$ e $y_i \in \mathfrak{N}(C_i; x_0)$ para todo $i \in \{1 : n\}$.

Demonstração. O fato de que x_0 é minimizador local estrito de f em C mostra que tal ponto é ponto isolado de $f^{-1}([-\infty, f(x_0)]) \cap C$, desse modo, a **Proposição 136** garante que os cones tangentes em x_0 são separáveis. Do **Teorema 49**, isso significa que existem $y_0, \dots, y_m$ não todos nulos cuja somatória é nula e de modo que $y_0 \in [\mathfrak{Z}(f^{-1}([-\infty, f(x_0)]); x_0)]^\Delta$ e $y_i \in [\mathfrak{Z}(C_i; x_0)]^\Delta$ para todo $i \in \{1 : n\}$. Por fim, a **Definição 114** garante o resultado. $\square$

5 Considerações finais e próximos passos

Diante do exposto neste trabalho, de acordo com o nosso conhecimento, o resultado mais geral em relação ao Princípio de Pontrjagin para o problema (1.3), enunciado no Teorema 76 desta dissertação, apresenta condições de otimalidade usando os colchetes de Lie para filtrar extremais, previamente, obtidos, o que foi bem demonstrado através do exemplo da seção 3.4. Entretanto, o problema (1.3) não contém restrições extras nas variáveis de estado — por exemplo, uma do tipo $A(t, x(t), u(t)) \leq x(t) \leq B(t, x(t), u(t))$. Restrições como estas são abordadas em [48] e [11], porém, apenas condições de primeira ordem são formuladas utilizando a abordagem de inclusões diferenciais, de modo que se torna interessante adaptar a abordagem geométrica para restrições deste caráter. Também, como visto, a dinâmica é afim e o crescimento do lagrangiano é sublinear nas variáveis de controle, embora este último facilite o aparecimento de impulsos¹, é interessante investigar o resultado apresentado permitindo dinâmicas quadráticas e lagrangiano de crescimento linear. Por fim, é interessante investigar condições de normalidade e não-degeneração como as encontradas em [33] e [34].

Já em relação ao formalismo de Dubovickij-Miljutin, o Teorema 137 é, possivelmente, um resultado original construído a partir de resultados encontrados em [11, 13, 21, 50, 42, 51] e de espírito similar à encontrada em [27], pois não é de conhecimento do autor a existência de resultados similares ou mais gerais em outras obras literárias. No entanto, o número de referências consultado é baixo, sendo necessário maior investigação (não realizada por falta de tempo) para resolver essa questão — por exemplo, o artigo [37] trata condições de otimalidade de alta ordem utilizando o formalismo de Dubovickij-Miljutin através das chamadas *pré-derivadas estritas*.. Ainda, é interessante investigar mais propriedades dos cones hipertangentes para tentar eliminar essa hipótese ou enfraquecê-la, pois o requerimento de hipertangência pode se mostrar demasiado restritivo, sem mencionar a necessidade de não-disjunção de diversos conjuntos. Mesmo assim, acredito que esta represente uma melhora — ainda que pequena — em relação à proposta de [15].

Referências

- [1] ARONNA, M. S.; MOTTA, M.; RAMPAZZO, F. A Higher Order Maximum Principle for Impulsive Optimal Controls. **SIAM Journal of Control and Optimization**, n. 58, p. 814–844, 2020. ISSN 0363-0129.
- [2] ARUTJUNOV, A. V.; DYXTA, V. A.; PEREIRA, F. L. Necessary conditions for impulsive nonlinear optimal control problems without *ā priori* normality assumptions. **Journal of Optimization Theory and Applications**, n. 124, p. 55–77, 2005. ISSN 0022-3239.
- [3] ARUTJUNOV, A. V. et al. Second-order necessary optimality conditions in optimal impulsive control problems. **Differential Equations**, n. 54, p. 1083–1101, 2018. ISSN 0012-2661.
- [4] AUBIN, J.-P.; FRANKOWSKA, H. **Set-Valued Analysis**. Boston: Springer Science & Business Media, 2009. xx+462. (Modern Birkhäuser Classics). ISBN 0-8176-3478-9.
- [5] BAUMEISTER, J.; LEITÃO, A. **Introdução à Teoria de Controle e Programação Dinâmica**. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2014. xiv+399. (Projeto Euclides). ISBN 85-244-0271-7.
- [6] BELLMAN, R. E. The theory of dynamic programming. **Bulletin of the American Mathematical Society**, n. 60, p. 503–515, 1954. ISSN 0273-0979.
- [7] BOLTJANSKIJ, V. G. The Method of Tents in the theory of extremal problems. **Russian Mathematical Surveys**, n. 30, p. 1–54, 1975. ISSN 0042-1316.
- [8] BRESSAN, A.; RAMPAZZO, F. On differential systems with vector-valued impulsive controls. **Bolletino dell'Unione Matematica Italiana B**, n. 2, p. 647–656, 1988. ISSN 0392-4041.
- [9] BREZIS, H. **Functional Analysis, Sobolev Spaces and Partial Differential Equations**. Nova York: Springer Science & Business Media, 2011. xiv+599. (Universitext). ISBN 0-387-70913-4.
- [10] CHITTARO, F.; STEFANI, G. Minimum-time strong optimality of a singular arc: The multi-input non involutive case. **ESAIM: Control, Optimisation and Calculus of Variations**, n. 22, p. 786–810, 2016. ISSN 1292-8119.
- [11] CLARKE, F. H. **Functional Analysis, Calculus of Variations and Optimal Control**. Londres: Springer Science & Business Media, 2013. xiv+591. (Graduate Texts in Mathematics, 264). ISBN 1-4471-4819-3.
- [12] CLARKE, F. H. Generalized Gradients and Applications. **Transactions of the Americal Mathematical Society**, n. 205, p. 247–262, 1975. ISSN 0002-9947.

- [13] CLARKE, F. H. **Optimization and Nonsmooth Analysis**. Filadélfia: Society for Industrial e Applied Mathematics, 1990. xii+308. (Classics in Applied Mathematics, 5). ISBN 0-89871-256-4.
- [14] DANTZIG, G. B. **Linear Programming and Extensions**. Princeton: Princeton University Press, 1963. xvi+640. (Princeton Landmarks in Mathematics and Physics). ISBN 0-691-05913-6.
- [15] DUBOVICKIJ, A. J.; MILJUTIN, A. A. Extremum problems with certain constraints. **Proceedings of the USSR Academy of Sciences**, n. 149, p. 759–762, 1963. ISSN 0002-3264.
- [16] DYXTA, V. A. The variational maximum principle and quadratic conditions for impulse processes and singular processes. **Siberian Mathematical Journal**, n. 35, p. 70–82, 1994. ISSN 0037-4474.
- [17] EKELAND, I. I.; TÉMAM, R. M. **Convex Analysis and Variational Problems**. Filadélfia: Society for Industrial e Applied Mathematics, 1999. xiv+402. (Classics in Applied Mathematics, 28). ISBN 0-89871-450-8.
- [18] EULER, L. **A method for finding curved lines enjoying properties of maximum or minimum, or solution of isoperimetric problems in the broadest accepted sense**. Lausanne, 1744. 328 p.
- [19] EULER, L. Elements of the Calculus of Variations. **New Comentaries of The Imperial Academy of Sciences and Arts**, n. 10, p. 51–93, 1766.
- [20] GEL'FAND, I. M.; FOMIN, S. V. **Calculus of Variations**. Edição e tradução: Richard A. Silverman. Eaglewood Cliffs: Prentice-Hall, Inc., 1963. vii+232. (Selected Russian Publications in Mathematical Sciences).
- [21] GIRSANOV, I. V. **Lectures on Mathematical Theory of Extremum Problems**. Berlim: Springer Science & Business Media, 1972. ii+136. (Lecture Notes in Economics and Mathematical Systems, 67). ISBN 3-540-05857-5.
- [22] GOLDSTINE, H. H. **A History of the Calculus of Variations from the 17th through the 19th Century**. Nova York: Springer Science & Business Media, 1980. xviii+410. (Studies in the history of Mathematics and Physical Sciences, 5). ISBN 1-4613-8108-8.
- [23] HESTENES, M. R. **A General Problem in the Calculus of Variations with Applications to Paths of Least Time**. Santa Monica: RAND Corporation, 1950. 46 p. (Research Memoranda, 100).
- [24] HUYGENS, C. **The Pendulum Clock: or Geometrical Demonstrations Concerning the Motion of Pendula as Applied to Clocks**. Paris: F. Muguet, 1673. xiv+165.
- [25] KANTOROVICH, L. V. **Mathematical Methods of Organizing and Planing Production**. Leningrado: Leningrad State University Publishing House, 1939. 68 p.

- [26] LAGRANGE, J.-L. On tautochrone curves. **Memories of the Prussian Academy of Sciences**, n. 21, 1765.
- [27] LASIECKA, I. Conical approximations of sets in optimization theory. **Control and Cybernetics**, n. 4, p. 39–58, 1975. ISSN 0324-8569.
- [28] LAW DEN, D. F. **Optimal Trajectories for Space Navigation**. Londres: Butterworth & Company Limited, 1963. vii+126.
- [29] LIMA, E. L. **Variedades Diferenciáveis**. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2011. x+290. (Publicações Matemáticas, 25). ISBN 85-244-0267-9.
- [30] MAUREC, J.-P. **Optimal Space Trajectories**. Elsevier B. V., 1979. (Studies in Astronautics, 1). ISBN 0-444-60107-4.
- [31] MILLER, B. M. The generalized solutions of nonlinear optimization problems with impulse control. **SIAM Journal on Control and Optimization**, n. 34, p. 1420–1440, 1996. ISSN 0363-0129.
- [32] MOTTA, M.; RAMPAZZO, F. Space-time trajectories of nonlinear systems driven by ordinary and impulsive controls. **Differential Integral Equations**, n. 8, p. 269–288, 1995. ISSN 0893-4983.
- [33] MOTTA, M.; RAMPAZZO, F.; VINTER, R. Normality and Gap Phenomena in Optimal Unbounded Control. **ESAIM: Control, Optimisation and Calculus of Variations**, n. 24, p. 1645–1673, 2018. ISSN 1292-8119.
- [34] MOTTA, M.; SARTORI, C. Normality and Nondegeneracy of the Maximum Principle in Optimal Impulsive Control Under State Constraints. **Journal of Optimization Theory and Applications**, n. 185, p. 44–71, 2020. ISSN 0022-3239.
- [35] MUNKRES, J. R. **Topology**. 2. ed. Upper Saddle River: Prentice-Hall, Inc., 2000. ISBN 0-13-181629-2.
- [36] NEWTON, I. **The Mathematical Principles of Natural Philosophy**. Londres: Royal Society, 1687. 306 p.
- [37] PÁLES, Z. Optimum problems with nonsmooth equality constraints. **Nonlinear Analysis**, n. 63, p. 2575–2581, 2005. ISSN 0362-546X.
- [38] PEREIRA, F. L.; SILVA, G. N. Necessary conditions of optimality for vector-valued impulsive control problems. **Systems & Control Letters**, n. 40, p. 205–215, 2000. ISSN 0167-6911.
- [39] PESCH, H. J.; PLAIL, M. The Cold War and the Maximum Principle of Optimal Control. **Documenta Mathēmatica**, n. 17, p. 331–343, 2012. ISSN 1431-0635.
- [40] PONTRJAGIN, L. S.; BOLTJANSKIJ, V. G.; GAMKRELIDZE, R. V. Theory of optimal processes. I. The maximum principle. **Proceedings of the USSR Academy of Sciences. Mathematical Series**, n. 24, p. 3–42, 1960. ISSN 1064-5632.

- [41] RISHEL, R. W. An extended Pontrjagin principle for control systems whose control laws contain measures. **Journal of the Society for Industrial and Applied Mathematics**, n. 3, p. 191–205, 1965. ISSN 0036-1399.
- [42] ROCKAFELLAR, R. T. **Convex Analysis**. Princeton: Princeton University Press, 1970. xiv+435. (Princeton Landmarks in Mathematics and Physics, 18). ISBN 0-691-61586-1.
- [43] SCHÄTTLER, H.; LEDZEWICZ, U. **Geometric Optimal Control: Theory, Methods and Examples**. Berlin: Springer Science & Business Media, 2012. xx+640. (Interdisciplinary Applied Mathematics, 38). ISBN 1-4614-3834-9.
- [44] SILVA, G. N.; VINTER, R. Necessary Conditions for Optimal Impulsive Control Problems. **SIAM Journal on Control and Optimization**, n. 35, p. 1829–1846, 1997. ISSN 0363-0129.
- [45] SOKAL, A. D. A really simple elementary proof of the uniform boundedness theorem. **American Mathematical Monthly**, n. 118, p. 450–452, 2011. ISSN 0002-9890.
- [46] SUSSMANN, H. J. Geometry and Optimal Control. In: BAILLIEUL, J. B.; WILLEMS, J. C. (Ed.). **Mathematical Control Theory**. Nova York: Springer Science & Business Media, 1998. P. 140–198. ISBN 0-387-98317-1.
- [47] SUSSMANN, H. J. Multidifferential calculus: chain rule, open mapping and transversal intersection theorems. In: HAGER, W. H.; PARDALÓS, P. M. (Ed.). **Optimal Control: Theory, Algorithms and Applications**. Nova York: Springer Science & Business Media, 1998. (Applied Optimization, 15). P. 436–487. ISBN 0-7923-5067-7.
- [48] VINTER, R. **Optimal Control**. Boston: Springer Science & Business Media, 2010. xx+507. (Modern Birkhäuser Classics). ISBN 0-8176-8086-1.
- [49] WARGA, J. **Optimal Control of Differential and Functional Equations**. Nova York: Elsevier B. V., 1972. ISBN 1-4832-4815-1.
- [50] WATKINS, G. G. Nonsmooth Miljutin-Dubovickij Theory and Clarke's Tangent Cone. **Mathematics of Operations Research**, n. 11, p. 70–80, 1986. ISSN 0364-765X.
- [51] ZĂLINESCU, C. **Convex Analysis in General Vector Spaces**. River Edge: World Scientific Publishing Company, 2002. xx+367. ISBN 981-238-067-1.

A Conceitos de Análise Funcional

Este capítulo apresenta conceitos e resultados de Análise Funcional necessários para compreender melhor os conteúdos da dissertação. Mais especificamente, apresenta o Teorema de Hahn-Banach de extensão de funcionais lineares dominados por sublineares, uma visão geral sobre topologias fracas encontrada em [11] com alguns resultados de [9], [17] e [21], dentre os quais se destaca o **Teorema 159** de separação de conjuntos convexos, e resultados básicos sobre espaços vetoriais normados e suas topologias — forte, fraca e fraca^{*}.

A.1 Teorema de Hahn-Banach

Definição 138. Seja P um conjunto não-vazio. Uma relação $\leq$ é dita **relação de ordem parcial** se:

- $x \in P \Rightarrow x \leq x$ — essa propriedade é denominada **reflexividade**;
- Se $x, y \in P$ são tais que $x \leq y$ e $y \leq x$, então, $x = y$ — essa propriedade é denominada **antissimetria**;
- Se $x, y, z \in P$ são tais que $x \leq y$ e $y \leq z$, então, $x \leq z$ — essa propriedade é denominada **transitividade**.

Nesse caso, o par $(P, \leq)$ é denominado **conjunto, parcialmente, ordenado** — quando não houver confusão sobre a relação de ordem usada, será dito, simplesmente, que P é um conjunto, parcialmente, ordenado.

Quando a relação de ordem parcial também conseguir comparar quaisquer dois elementos de P , ela será dita **total** e P será chamado **conjunto, totalmente, ordenado**.

Definição 139 (páginas 1 e 2 de [9]). Seja P , um conjunto parcialmente, ordenado.

- Dado $Q \subseteq P$, um elemento $c \in P$ é **cota superior** de Q se qualquer $a \in Q$ satisfaz $a \leq c$;
- Um elemento $m \in P$ é **maximal de P** se não precede nenhum elemento distinto dele;
- O conjunto P é denominado **indutivo** se todo subconjunto $Q \subseteq P$, totalmente, ordenado admite cota superior.

Lema 140 (página 2 de [9]). *Todo conjunto não-vazio e indutivo admite elemento maximal.*

Para os próximos resultados, X é um espaço vetorial real.

Definição 141 (página 1 de [9]). Um funcional $p : X \rightarrow \mathbb{R}$ é dito **sublinear** se satisfaz

- $x, y \in X \Rightarrow p(x + y) \leq p(x) + p(y)$ — essa propriedade é denominada **subaditividade**;
- $x \in X, \alpha \geq 0 \Rightarrow p(\alpha x) = \alpha p(x)$ — essa propriedade é denominada **1-homogeneidade não-negativa**.

Se a primeira desigualdade for uma igualdade (a propriedade em questão passa a ser chamada **aditividade**) e a segunda igualdade valer para todo $\alpha \in \mathbb{R}$ (nesse caso, retira-se o “não-negativa” e denomina-se a propriedade de 1-homogeneidade), então, o funcional será chamado **linear**.

Definição 142. O conjunto de todos os funcionais lineares com domínio em X é denominado **dual algébrico de X** e denotado por X' .

Dessa definição, é claro que os funcionais sublineares constituem um conjunto “mais amplo” do que os lineares.

Teorema 143 (página 1 de [9]). *Sendo $p : X \rightarrow \mathbb{R}$, L um subespaço vetorial de X e $g : L \rightarrow \mathbb{R}$ um funcional dominado, superiormente, por p em L , existe um funcional linear $f : X \rightarrow \mathbb{R}$ estendendo g para todo o espaço X e dominado, superiormente, por p em X .*

Demonstração. Considere o conjunto P dos funcionais lineares h que estendem g para um certo subespaço D_h de X e são dominados, superiormente, por p em seus domínios. Por construção, tal conjunto contém g e, conseqüentemente, não é vazio. Ainda, é possível definir a seguinte relação de ordem parcial: $h_1 \leq h_2$ se, e somente se, $D_{h_1} \subseteq D_{h_2}$ e h_2 estende h_1 — a verificação de que esta é uma relação de ordem parcial, embora não seja difícil, será omitida por não ser o foco do estudo.

Afirmção. O conjunto P é indutivo com a relação $\leq$.

Tomando $Q = \{h_i \mid i \in I\} \subseteq P$ um conjunto, totalmente, ordenado e definindo D_h como a soma dos D_{h_i} para $i \in I$ e $h : D_h \rightarrow \mathbb{R}$ de modo que $h = h_i$ em D_{h_i} , não é difícil verificar que o funcional linear h construído é um elemento de P e estende cada um dos h_i para D_h , assim, como $D_{h_i} \subseteq D_h$, tal funcional é cota superior de Q em P .

Pelo **Lema 140**, P admite um elemento maximal f .

Afirmção. O domínio D_f de f é todo o espaço X .

Supondo que isso não aconteça e tomando $x_0 \in X \setminus D_f$, para $x_1, x_2 \in D_f$, vale

$$f(x_1) + f(x_2) = f(x_1 + x_2) \leq p(x_1 + x_2) \leq p(x_1 - x_0) + p(x_0 + x_2),$$

ou, de maneira equivalente,

$$f(x_1) - p(x_1 - x_0) \leq p(x_0 + x_2) - f(x_2). \quad (\text{A.1})$$

Agora, considerando os conjuntos W e Z formados, respectivamente, por elementos do lado esquerdo e direito de (A.1), é possível afirmar que $r_1 := \sup W \leq \inf Z =: r_2$, o que permite encontrar $\alpha \in [r_1, r_2]$ e definir

$$h : D_f + [x_0] \rightarrow \mathbb{R}$$

$$x + \lambda x_0 \mapsto f(x) + \lambda \alpha.$$

Claramente, h é um funcional linear e estende f para $D_f + [x_0]$, o que implica que h é dominado, superiormente, por p em D_f . Ainda, para $\lambda > 0$ e $x \in D_f$,

$$h(x + \lambda x_0) = f(x) + \lambda \alpha \leq f(x) + \lambda p\left(x_0 + \frac{x}{\lambda}\right) - f(x) = p(x + \lambda x_0),$$

enquanto, para $\lambda = -\mu < 0$,

$$h(x + \lambda x_0) = f(x) - \mu \alpha \leq f(x) - f(x) + \mu p\left(\frac{y}{\mu} - x_0\right) = p(y - \mu x_0) = p(x + \lambda x_0),$$

o que mostra que h é dominado, superiormente, por p em $D_f + [x_0]$, consequentemente, $f < h$, o que contradiz o fato de que f é maximal.

Com essa afirmação mostrada, o resultado segue. $\square$

A.2 Topologias fracas

Além de supor que X é espaço vetorial real, os resultados daqui serão feitos supondo que $\Phi \subseteq X'$ é um espaço vetorial separante, isto é, para $x \in X \setminus \{0\}$, existe $\varphi \in \Phi$ que não se anula em x .

Nessas condições:

- Para $\varphi \in \Phi$ e $x \in X$, o valor $\varphi(x)$ será denotado por $\langle x, \varphi \rangle$;
- Para $\varphi \in \Phi$ e $Y \subseteq X$, o conjunto $\varphi(Y)$ será denotado por $\langle Y, \varphi \rangle$;
- Para $\Psi \subseteq \Phi$ e $x \in X$, o conjunto das imagens dos elementos de Ψ por x será denotado por $\langle x, \Psi \rangle$;

Lema 144 (página 48 de [11]). *A coleção de todos os conjuntos $V(x_0, \varphi, \epsilon) := \{x \in X : |\langle x - x_0, \varphi \rangle| < \epsilon\}$ para $x_0 \in X$, $\varphi \in \Phi$ e $\epsilon > 0$ forma uma subbase para uma topologia em X .*

Demonstração. Dado $x_0 \in X$, não é difícil ver que

$$x_0 \in \bigcap_{\varphi \in \Phi, \epsilon > 0} V(x_0, \varphi, \epsilon).$$

$\square$

A topologia gerada pelos conjuntos $V(x_0, \varphi, \epsilon) := \{x \in X : |\langle x - x_0, \varphi \rangle| < \epsilon\}$ com $x_0 \in X$, $\varphi \in \Phi$ e $\epsilon > 0$ será denotada $\varsigma(X, \Phi)$.

Proposição 145 (página 48 de [11]). *A topologia $\varsigma(X, \Phi)$ é a mais fraca na qual todos os funcionais $\varphi \in \Phi$ são contínuos.*

Demonstração. Seja τ uma topologia na qual todos os funcionais $\varphi \in \Phi$ são contínuos. Dados $x_0 \in X$, $\varphi \in \Phi$ e $\epsilon > 0$, temos que $V(x_0, \varphi, \epsilon) = \varphi^{-1}(]\langle x_0, \varphi \rangle - \epsilon, \langle x_0, \varphi \rangle + \epsilon[) \in \tau$ pela continuidade de φ , portanto, $\varsigma(X, \Phi) \subseteq \tau$. $\square$

Definição 146 (página 48 de [11]). A topologia $\varsigma(X, \Phi)$ é denominada **topologia fraca de X induzida por Φ** .

Propriedade 147 (página 48 de [11]). Fixados $x_0 \in X$, $\varphi \in \Phi$ e $\epsilon > 0$, vale que $V(x_0, \varphi, \epsilon) = x_0 + V(0, \varphi, \epsilon)$, e qualquer $t > 0$ satisfaz $V(0, \varphi, t\epsilon) = tV(0, \varphi, \epsilon)$.

Demonstração. Para todo $x \in X$, temos que

$$\begin{aligned} x \in V(x_0, \varphi, \epsilon) &\Leftrightarrow |\langle x - x_0, \varphi \rangle| < \epsilon \Leftrightarrow x - x_0 \in V(0, \varphi, \epsilon) \\ x \in V(0, \varphi, t\epsilon) &\Leftrightarrow t \left| \left\langle \frac{x}{t}, \varphi \right\rangle \right| = |\langle x, \varphi \rangle| < t\epsilon \Leftrightarrow x \in tV(0, \varphi, \epsilon). \end{aligned}$$

$\square$

Para o próximo teorema e outros resultados, $x_i \xrightarrow{\Phi} x$ indicará que $\{x_i\}$ converge para x na topologia fraca de X induzida por Φ .

Teorema 148 (páginas 48 e 49 de [11]). • A coleção de interseções

$$\bigcap_{i \in I} V(0, \varphi_i, \epsilon)$$

para $I \subseteq \mathbb{N}$ finito, $\{\varphi_i \mid i \in I\} \subseteq \Phi$ e $\epsilon > 0$ é uma base local de $\varsigma(X, \Phi)$ na origem;

- $(X, \varsigma(X, \Phi))$ é um espaço de Hausdorff, e — para as respectivas topologias-produto — as aplicações $(x, u) \mapsto x + u$ e $(t, x) \mapsto tx$ são contínuas;
- Dados $x_0 \in X$, $U \in \varsigma(X, \Phi)$ e $t \neq 0$, tanto $x_0 + U$ quanto tU são elementos de $\varsigma(X, \Phi)$;
- Φ é o conjunto de todos os funcionais lineares contínuos em $(X, \varsigma(X, \Phi))$;
- Dados $\{x_i\} \subseteq X$ e $x_0 \in X$, $x_i \xrightarrow{\Phi} x_0$ se, e somente se, $\langle x_i, \varphi \rangle \rightarrow \langle x_0, \varphi \rangle$ para todo $\varphi \in \Phi$;
- Sendo Y um espaço topológico, a aplicação $\psi: Y \rightarrow (X, \varsigma(X, \Phi))$ é contínua se, e somente, $\varphi \circ \psi: Y \rightarrow \mathbb{R}$ é contínua para todo $\varphi \in \Phi$.

Demonstração. • Tomando $I \subseteq \mathbb{N}$ finito, sejam $\{x_i \mid i \in I\} \subseteq X$, $\{\varphi_i \mid i \in I\} \subseteq \Phi$ e $\{\epsilon_i \mid i \in I\} \subseteq]0, \infty[$ de modo que

$$0 \in \bigcap_{i \in I} V(x_i, \varphi_i, \epsilon_i).$$

A partir disso, temos que $|\langle x_i, \varphi_i \rangle| < \epsilon_i$ para todo $i \in I$, assim, escolhendo $\epsilon = \min\{\epsilon_i - |\langle x_i, \varphi_i \rangle| \mid i \in I\}$, vemos que

$$\bigcap_{i \in I} V(0, \varphi_i, \epsilon) \subseteq \bigcap_{i \in I} V(x_i, \varphi_i, \epsilon_i);$$

- Tome $x, y \in X$ pontos distintos. Sabendo que Φ é separante, existe $\varphi \in \Phi$ tal que $\varphi(x) \neq \varphi(y)$, então, como a reta usual é espaço de Hausdorff, existe $\epsilon > 0$ de modo que $] \langle x, \varphi \rangle - \epsilon, \langle x, \varphi \rangle + \epsilon[$ é disjunto de $] \langle y, \varphi \rangle - \epsilon, \langle y, \varphi \rangle + \epsilon[$, o que mostra que $V(x, \varphi, \epsilon) \cap V(y, \varphi, \epsilon) = \emptyset$. Agora, sendo $W \in \varsigma(X, \Phi)$ uma vizinhança de 0, do primeiro item, existem $I \subseteq \mathbb{N}$ finito, $\{\varphi_i \mid i \in I\} \subseteq \Phi$ e $\epsilon > 0$ de modo que

$$\bigcap_{i \in I} V(0, \varphi_i, \epsilon) \subseteq W,$$

assim,

$$\begin{aligned} x, u \in \bigcap_{i \in I} V\left(0, \varphi_i, \frac{\epsilon}{2}\right) &\Rightarrow x + u \in \bigcap_{i \in I} V(0, \varphi_i, \epsilon), \\ t > 0, x \in \bigcap_{i \in I} V\left(0, \varphi_i, \frac{\epsilon}{t}\right) &\Rightarrow tx \in \bigcap_{i \in I} V(0, \varphi_i, \epsilon). \end{aligned}$$

o que confirma a continuidade das aplicações;

- Do primeiro item e do **Lema 144**, para cada $x \in U$, existem $I(x) \subseteq \mathbb{N}$, $\{\varphi_i \mid i \in I(x)\} \subseteq \Phi$ e $\epsilon(x) > 0$ de modo que

$$U = \bigcup_{x \in U, \epsilon(x) > 0} \left[\bigcap_{i \in I(x)} V(x, \varphi, \epsilon(x)) \right],$$

assim, a **Propriedade 147** garante que

$$\begin{aligned} x_0 + U &= \bigcup_{x \in U, \epsilon(x) > 0} \left[\bigcap_{i \in I(x)} V(x + x_0, \varphi, \epsilon(x)) \right], \\ tU &= \bigcup_{x \in U, \epsilon(x) > 0} \left[\bigcap_{i \in I(x)} V(tx, \varphi, |t|\epsilon(x)) \right], \end{aligned}$$

o que mostra que tanto $x_0 + U$ quanto tU são abertos;

- Os elementos de Φ são contínuos na topologia $\varsigma(X, \Phi)$ por construção. Agora, sendo f_0 um funcional linear contínuo em $\varsigma(X, \Phi)$, tal funcional é limitado em algum conjunto da forma

$$\bigcap_{i \in I} V(0, \varphi_i, \epsilon)$$

com $I \subseteq \mathbb{N}$ finito, $\{\varphi_i \mid i \in I\} \subseteq \Phi$ e $\epsilon > 0$, então, existe um $M > 0$ para o qual, a princípio,

$$x \in \bigcap_{i \in I} \ker \varphi_i \Rightarrow |f_0(x)| \leq M.$$

Porém, sabendo que x pode ser substituído por kx para qualquer $k \in \mathbb{Z}$, é verdade que f se anula na interseção dos núcleos dos φ_i , logo, f_0 deve ser uma combinação linear destes, o que mostra que $f_0 \in \Phi$;

- Se $x_i \xrightarrow{\Phi} x_0$, então, como Φ é o conjunto dos funcionais lineares contínuos, qualquer $\varphi \in \Phi$ satisfaz $\langle x_i, \varphi \rangle \rightarrow \langle x_0, \varphi \rangle$. Agora, se $\{x_i\} \subseteq X$ é tal que $\langle x_i, \varphi \rangle \rightarrow \langle x_0, \varphi \rangle$ para todo $\varphi \in \Phi$, seja W uma vizinhança de x_0 em $\varsigma(X, \Phi)$. Sabendo que ele contém um conjunto da forma

$$\bigcap_{j \in J} V(x_0, \varphi_j, \epsilon)$$

com $J \subseteq \mathbb{N}$ finito, $\{\varphi_j \mid j \in J\} \subseteq \Phi$ e $\epsilon > 0$, existe $\nu: J \rightarrow \mathbb{N}$ para a qual $i \geq \nu(j) \Rightarrow |\langle x_i - x_0, \varphi_j \rangle| < \epsilon$, então, pondo $\iota = \max \nu(J)$, vem

$$i \geq \iota \Rightarrow x_i \in \bigcap_{j \in J} V(x_0, \varphi_j, \epsilon) \subseteq W,$$

o que confirma que $x_i \xrightarrow{\Phi} x_0$;

- Se $\psi: Y \rightarrow (X, \varsigma(X, \Phi))$ é contínua, então, é claro que $\varphi \circ \psi$ é contínua para todo $\varphi \in \Phi$. Reciprocamente, sendo $\varphi \circ \psi$ é contínua para todo $\varphi \in \Phi$, seja U um aberto em X . Do primeiro item e do **Lema 144**, para cada $x \in U$, existem $I(x) \subseteq \mathbb{N}$, $\{\varphi_i \mid i \in I(x)\} \subseteq \Phi$ e $\epsilon(x) > 0$ de modo que

$$U = \bigcup_{x \in U, \epsilon(x) > 0} \left[\bigcap_{i \in I(x)} V(x, \varphi_i, \epsilon(x)) \right],$$

assim,

$$\psi^{-1}(U) = \bigcup_{x \in U, \epsilon(x) > 0} \left[\bigcap_{i \in I(x)} \varphi_i^{-1}(V(x, \varphi_i, \epsilon(x))) \right],$$

o que mostra que $\psi^{-1}(U)$ é aberto. Portanto, ψ é contínua. □

Definição 149 (páginas 16 e 44 de [21]). • Um funcional $f: X \rightarrow [-\infty, \infty]$ é dito **convexo** se, fixados $x, y \in X$, qualquer $t \in [0, 1]$ satisfaz

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y); \quad (\text{A.2})$$

- Um subconjunto $S \subseteq X$ é dito **convexo** se, dados $x, y \in S$, $\{(1-t)x + ty \mid t \in [0, 1]\} \subseteq S$.

Propriedade 150 (páginas 44 e 45 de [21]). *Toda combinação linear positiva de funcionais convexos é um funcional convexo, e a envoltória superior de funcionais convexos é, também, um funcional convexo. Ainda, a imagem inversa de uma semirreta por um funcional convexo é convexa, e a interseção de conjuntos convexos é convexa.*

Propriedade 151 (página 16 de [21]). *Se A é convexo, $x \in A$ e $y \in \mathring{A}$, então, dado $\lambda \in [0, 1[$, o ponto $\lambda x + (1 - \lambda)y \in \mathring{A}$.*

Demonstração. Sabendo que $y \in \mathring{A}$, é possível encontrar uma vizinhança $V \subseteq A$ de y , assim, a convexidade de A garante que cada um dos abertos $\lambda x + (1 - \lambda)V \subseteq A$ para $\lambda \in [0, 1[$, o que, sabendo que cada um deles é uma vizinhança do correspondente ponto $\lambda x + (1 - \lambda)y$, prova o resultado. $\square$

Lema 152 (página 11 de [17]). *Se um funcional f convexo é limitado, superiormente, em uma vizinhança de um ponto $x_0 \in X$, então, ele é contínuo em tal ponto.*

Demonstração. Definindo o funcional $g : X \rightarrow [-\infty, \infty]$ por $g(v) = f(x_0 + v) - f(x_0)$, trata-se de um funcional convexo, já que, para $v_1, v_2 \in X$ e $t \in [0, 1]$, vale

$$\begin{aligned} g((1 - t)v_1 + tv_2) &= f((1 - t)(x_0 + v_1) + t(x_0 + v_2)) - f(x_0) \\ &\leq (1 - t)[f(x_0 + v_1) - f(x_0)] + t[f(x_0 + v_2) - f(x_0)] \\ &= (1 - t)g(v_1) + tg(v_2), \end{aligned}$$

e que satisfaz $g(0) = 0$. Dito isso, seja $V \in \mathcal{S}(X, \Phi)$ uma vizinhança da origem para a qual $f(V) \subseteq]-\infty, a]$ com $a \in \mathbb{R}$. Definindo $W = V \cap (-V)$, trata-se de uma vizinhança simétrica da origem e, dado $\epsilon \in]0, 1[$, a convexidade de g garante que

$$v \in \epsilon W \Rightarrow \begin{cases} g(v) \leq (1 - \epsilon)g(0) + \epsilon g\left(\frac{v}{\epsilon}\right) \leq \epsilon a \\ g(-v) \leq (1 - \epsilon)g(0) + \epsilon g\left(-\frac{v}{\epsilon}\right) \leq \epsilon a \end{cases} \xrightarrow{-g(v) \leq g(-v)} g(v) \in [-\epsilon a, \epsilon a],$$

o que mostra a continuidade deste na origem, portanto, f é contínuo em x . $\square$

Proposição 153 (página 12 de [17]). *Seja $f : X \rightarrow [-\infty, \infty]$ um funcional convexo, as seguintes afirmações são equivalentes:*

- *Existe um aberto não-vazio A no qual f não coincide, totalmente, com $-\infty$ e está limitado, superiormente, por uma constante $a < \infty$;*
- *f não assume valor $-\infty$ em um ponto sequer, é finita em, ao menos um ponto e contínua no interior do conjunto $D = X \setminus f^{-1}(\infty)$, que não é vazio.*

Demonstração. Se a segunda afirmação é válida, então, por definição, existe uma vizinhança A da origem satisfazendo $f(A) \subseteq [f(0) - 1, f(0) + 1]$, o que mostra a primeira afirmação.

Reciprocamente, sendo a primeira afirmação válida, $A \subseteq \mathring{D}$ e, ao fixar $u \in A \setminus f^{-1}(-\infty)$, o **Lema 152** garantirá que f é contínuo em u e, consequentemente, finito em uma vizinhança deste, o que mostra que f não assume valor $-\infty$ em um ponto sequer. Ainda, para qualquer $v \in \text{int } D$, existe um $\varrho > 1$ para o qual $w := u + \varrho(v - u) \in \mathring{D}$, assim, definindo $h : A \rightarrow X$ por $h(x) = w + \frac{\varrho-1}{\varrho}(x - w)$, vale que

$$h(u) = w + \frac{\varrho-1}{\varrho}(u - w) = u + \varrho(v - u) + (1 - \varrho)(v - u) = v$$

e que $h(A)$ é uma vizinhança de v tal que todo elemento v' dela satisfaz

$$f(v') = f\left(\frac{w}{\varrho} + \frac{\varrho-1}{\varrho}h^{-1}(v')\right) \leq \frac{f(w)}{\varrho} + \frac{\varrho-1}{\varrho}f \circ h^{-1}(v') \leq a + \frac{F(w) - a}{\varrho}.$$

Do **Lema 152**, a conclusão é que f é contínua em v e, consequentemente, no interior de seu domínio efetivo, o que mostra a segunda afirmação. $\square$

Definição 154 (página 4 de [9]). Dados $g \in X'$ não-nulo e $\alpha \in \mathbb{R}$, o conjunto $g^{-1}(\alpha) = \{x \in X \mid g(x) = \alpha\}$ é denominado **hiperplano afim de equação $g = \alpha$** .

Definição 155 (adaptado das páginas 5 de [9], 41 de [11] e 72 de [50]). Sendo $A, B \subseteq X$ não-vazios, o hiperplano de equação $\zeta = \alpha$ **separa A e B** se $\sup\langle A, \zeta \rangle \leq \alpha$ e $\inf\langle B, \zeta \rangle \geq \alpha$ e **separa, estritamente, A e B** se existe $\epsilon > 0$ para o qual $\sup\langle A, \zeta \rangle \leq \alpha - \epsilon$ e $\inf\langle B, \zeta \rangle \geq \alpha + \epsilon$. Ainda, sendo $A_1, \dots, A_m \subseteq X$ não-vazios, tais conjuntos são ditos **separáveis** se existe um hiperplano separando um deles da interseção dos demais.

Definição 156 (adaptado da página 6 de [9]). Sendo $C \in \zeta(X, \Phi)$ um aberto convexo contendo a origem, o funcional

$$x \mapsto \inf\left\{\alpha > 0 \mid \frac{x}{\alpha} \in C\right\}$$

é denominado **calibre de C** e denotado por p_C .

Lema 157 (adaptado das páginas 5 e 6 de [9]). Sendo $C \in \zeta(X, \Phi)$ um convexo contendo a origem, p_C é um funcional sublinear, e $C = p_C^{-1}([0, 1])$.

Demonstração. Supondo que $x \in C$, o fato de que tal conjunto é aberto garante a existência de $\epsilon(x) > 0$ para o qual $[1 + \epsilon(x)]x \in C$, o que mostra que $p_C(x) \leq \frac{1}{1 + \epsilon(x)} < 1$. Reciprocamente, se $p_C(x) < 1$, existe $\alpha \in]0, 1[$ para o qual $\frac{x}{\alpha} \in C$, então, a convexidade de tal conjunto e o fato de que este contém a origem mostram que $x = \alpha \frac{x}{\alpha} + (1 - \alpha)0 \in C$. Portanto, $C = p_C^{-1}([0, 1])$.

Agora, dado $t > 0$, segue que

$$p_C(tx) = \inf\left\{\alpha > 0 \mid \frac{tx}{\alpha} \in C\right\} = \inf\left\{t \frac{\alpha}{t} > 0 \mid \frac{x}{\frac{\alpha}{t}} \in C\right\} = t \inf\left\{\beta > 0 \mid \frac{x}{\beta} \in C\right\} = tp_C(x).$$

Ainda, fixados $x, y \in E$, seja $\epsilon > 0$. Sabendo que

$$p_C\left(\frac{x}{p_C(x) + \epsilon}\right) = \frac{p_C(x)}{p_C(x) + \epsilon} < 1, \quad p_C\left(\frac{y}{p_C(y) + \epsilon}\right) = \frac{p_C(y)}{p_C(y) + \epsilon} < 1,$$

segue que $\frac{x}{p_C(x)+\epsilon}, \frac{y}{p_C(y)+\epsilon} \in C$, logo, escolhendo $t = \frac{p_C(x)+\epsilon}{p_C(x)+p_C(y)+2\epsilon} \in [0, 1]$, a convexidade de C garante que

$$C \ni \frac{p_C(x)+\epsilon}{p_C(x)+p_C(y)+2\epsilon} \frac{x}{p_C(x)+\epsilon} + \frac{p_C(y)+\epsilon}{p_C(x)+p_C(y)+2\epsilon} \frac{y}{p_C(y)+\epsilon} = \frac{x+y}{p_C(x)+p_C(y)+2\epsilon},$$

o que mostra que $p_C(x+y) < p_C(x)+p_C(y)+2\epsilon$. Fazendo $\epsilon \downarrow 0$, segue que $p_C(x+y) \leq p_C(x)+p_C(y)$, portanto, p_C é sublinear. $\square$

Lema 158 (adaptado da página 6 de [9]). *Seja $C \in \mathcal{C}(X, \Phi)$ um convexo não-vazio e $x_0 \in X \setminus C$, existe $\zeta \in \Phi$ para o qual $\langle x_0, \zeta \rangle$ é cota superior de $\langle C, \zeta \rangle$.*

Demonstração. Supondo que $0 \in C$, seja p_C o calibre de C . Considerando o subespaço vetorial $[x_0]$ e o funcional linear definido por $tx_0 \mapsto t$: se $t > 0$, então, pelo **Lema 157**, $t \leq tp_C(x_0) = p_C(tx_0)$; senão, $t \leq 0 \leq p_C(tx_0)$. De qualquer modo, tal funcional é dominado, superiormente, por p_C em $[x_0]$, então, o **Teorema 143** garante a existência de $\zeta \in X'$ que o estende e é dominado, superiormente, por p_C .

Dos Lemas 157 e 152, é possível afirmar que ζ é contínuo, e o **Lema 157** garante, também, que $x \in C \Rightarrow \langle x, \zeta \rangle \leq p_C(x) < 1 = \langle x_0, \zeta \rangle$, portanto, $\langle x_0, \zeta \rangle$ é cota superior de $\langle C, \zeta \rangle$.

No caso em que $0 \notin C$, considerando $x_1 \in C$, $\tilde{x}_0 := x_0 - x_1$ e $\tilde{C} := C - x_1$, este é um aberto convexo contendo a origem, o que permite aplicar o raciocínio acima e obter $\zeta \in \Phi$ para o qual $\langle \tilde{x}_0, \zeta \rangle$ é cota superior de $\langle \tilde{C}, \zeta \rangle$, o que equivale, como tal funcional é linear, a $\langle x_0, \zeta \rangle$ ser cota superior de $\langle C, \zeta \rangle$. $\square$

Teorema 159 (página 50 de [11]). *Seja $A, B \subseteq X$ convexos não-vazios e disjuntos:*

- *Se $A \in \mathcal{C}(X, \Phi)$, então, existe um hiperplano fechado separando A e B ;*
- *Se A é fechado em $\mathcal{C}(X, \Phi)$ e B é compacto nesta, então, existe um hiperplano fechado separando, estritamente, os dois.*

Demonstração. • Definindo $C = A - B$, como A e B não são vazios, o mesmo vale para C . Ainda, dados $x, y \in C$, existem $x_A, y_A \in A$ e $x_B, y_B \in B$ tais que $x = x_A - x_B$ e $y = y_A - y_B$, assim, para todo $t \in [0, 1]$, as convexidades de A e B garantem que

$$\begin{aligned} (1-t)x + ty &= (1-t)x_A - (1-t)x_B + ty_A - ty_B \\ &= [(1-t)x_A + ty_A] - [(1-t)x_B + ty_B] \in C. \end{aligned}$$

Mais ainda, $C = \bigcup_{y \in B} (A - y) \in \mathcal{C}(X, \Phi)$ (**Teorema 148**), e a disjunção de A e B garante que $0 \notin C$. A partir disso e do **Lema 158**, existe $\zeta \in \Phi$ que é negativo em C , o que mostra que $\langle x, \zeta \rangle < \langle y, \zeta \rangle$ para todos $x \in A$ e $y \in B$, assim, chamando α a média aritmética entre $\sup \langle A, \zeta \rangle$ e $\inf \langle B, \zeta \rangle$, o hiperplano de equação $\zeta = \alpha$ separa A e B ;

- Do **Teorema 148**, cada ponto $y \in B$ admite um convexo $V_y \in \mathcal{S}(X, \Phi)$ disjunto de A , e a coleção $\{V_y \mid y \in B\}$ forma uma cobertura aberta de B . Da compacidade deste, existem $n \in \mathbb{N}$ e $y_1, \dots, y_n \in B$ cujos correspondentes convexos $V_1, \dots, V_n \in \mathcal{S}(X, \Phi)$ cobrem B , assim, o aberto

$$V := \bigcap_{i=1}^n (V_i - y_i) \in \mathcal{S}(X, \Phi)$$

é um convexo contendo a origem tal que $V + B$ é disjunto de A . Aplicando o primeiro item, encontramos um $\zeta \in \Phi$ de modo que $\sup\langle A, \zeta \rangle \leq \inf\langle V + B, \zeta \rangle$, assim, como $\langle B, \zeta \rangle$ é um subconjunto compacto do aberto $\langle V + B, \zeta \rangle$, vale $\sup\langle A, \zeta \rangle < \inf\langle B, \zeta \rangle$. Chamando α a média aritmética entre $\sup\langle A, \zeta \rangle$ e $\inf\langle B, \zeta \rangle$, o hiperplano de equação $\zeta = \alpha$ separa, estritamente, A e B . □

Corolário 160 (adaptado da página 8 de [9]). *Se $Y \subseteq X$ é um subespaço vetorial não-denso, então, existe $\zeta \in \Phi$ não-nulo que se anula, identicamente, em Y .*

Demonstração. Sendo $x_0 \in X \setminus \bar{Y}$, o conjunto $\bar{Y}$ é fechado, convexo e não-vazio, e $\{x_0\}$ é um compacto convexo não-vazio que é disjunto deste, assim, o **Teorema 159** garante a existência de um hiperplano fechado de equação $\zeta = \alpha$ separando, estritamente, os dois. Em particular, isso acarreta que $\langle kx, \zeta \rangle < \alpha$ para todos $x \in Y$ e $k \in \mathbb{Z}$, assim, qualquer $n \in \mathbb{N}$ satisfaz

$$-\frac{\alpha}{n} < \langle x, \zeta \rangle < \frac{\alpha}{n} \quad \forall x \in Y.$$

Fazendo $n \uparrow \infty$, concluímos que ζ se anula, identicamente, em Y . □

Observação 161 (página 8 de [9]). É mais frequente a utilização da recíproca desse corolário: “se $Y \subseteq X$ é um subespaço vetorial no qual o único funcional que se anula é o funcional nulo, então, Y é denso em X .”

A.3 Espaços normados

Os resultados, aqui, serão feitos supondo que X é espaço vetorial real.

Definição 162. Uma **norma** em X é um funcional $\|\cdot\|_X : X \rightarrow [0, \infty[$ 1-homogêneo e subaditivo que se anula, somente, no vetor nulo de X — funcionais com essa última propriedade são denominados **definidos positivos**.

Nesse caso, o par $(X, \|\cdot\|_X)$ é denominado **espaço vetorial real normado** — novamente, quando não houver confusão da norma, será dito que X é espaço vetorial real normado.

Proposição 163. *A aplicação $\text{dist}_X : X \times X \rightarrow [0, \infty[$ definida por $\text{dist}_X(x, y) = \|x - y\|_X$ é uma métrica em X , e tanto a soma quanto a multiplicação por escalar são contínuas na topologia determinada por essa métrica.*

A bola fechada, nesta métrica, de centro $x_0 \in X$ e raio $r > 0$ será denotada, em geral por $B_X[x_0, r]$, e a correspondente bola aberta será denotada por $B_X(x_0, r)$. Com essa métrica, X torna-se um espaço métrico sobre o qual é possível definir os conceitos de sequências convergentes e de Cauchy e de conjuntos abertos, fechados, limitados, compactos e, sequencialmente, compactos.

Definição 164. X é dito **espaço de Banach** se o espaço (X, dist_X) é completo.

Proposição 165. A aplicação $f \mapsto \sup\{|f(x)| \mid x \in B_X[0, 1]\}$ é uma norma em X' .

Demonstração. Claramente, o funcional em questão nunca assume valores negativos, e ele é nulo se, e somente se, o correspondente elemento de X' se anula em $B_X[0, 1]$, o que, pela linearidade de tal elemento, é equivalente ao seu anulamento em X , logo, tal funcional é positivo definido.

Ainda, dados $f, g \in X'$ e $\alpha \in \mathbb{R}$, vale que

$$\begin{aligned} \sup\{|f(x) + g(x)| \mid x \in B_X[0, 1]\} &\leq \sup\{|f(x)| + |g(x)| \mid x \in B_X[0, 1]\} \\ &\leq \sup\{|f(x)| \mid x \in B_X[0, 1]\} + \sup\{|g(x)| \mid x \in B_X[0, 1]\} \\ \sup\{|\alpha f(x)| \mid x \in B_X[0, 1]\} &= \sup\{|\alpha| |f(x)| \mid x \in B_X[0, 1]\} \\ &= |\alpha| \sup\{|f(x)| \mid x \in B_X[0, 1]\}, \end{aligned}$$

o que mostra que tal funcional é subaditivo e 1-homogêneo. Portanto, tal funcional é uma norma. $\square$

Definição 166. O espaço vetorial X' munido da norma da **Proposição 165** é denominado **dual topológico de X** e denotado por X^* .

Lema 167. Sendo $f \in X^*$,

- Qualquer $x \in X$ satisfaz $|\langle x, f \rangle| \leq \|f\|_{X^*} \|x\|_X$;
- Quaisquer $x \in X$ e $r > 0$ satisfazem $\sup |\langle B_X[x, r]; f \rangle| \geq \|f\|_{X^*} r$.

Demonstração. • A desigualdade é imediata para $x = 0$. Já para $x \neq 0$, temos que

$$|\langle x, f \rangle| = \|x\|_X \left| \left\langle \frac{x}{\|x\|_X}, f \right\rangle \right| \leq \|f\|_{X^*} \|x\|_X;$$

- (adaptada de [45]) Sabendo que $B_X[x, r] = x \pm rB_X[0, 1]$, qualquer $\xi \in B_X[0, 1]$ satisfaz

$$\sup\{|\langle x + r\xi, f \rangle|, |\langle x - r\xi, f \rangle|\} \geq \frac{|\langle x + r\xi, f \rangle| + |\langle x - r\xi, f \rangle|}{2} \geq r|\langle \xi, f \rangle|,$$

logo, basta tomar o supremo para ξ variando em $B_X[0, 1]$ e o resultado seguirá. $\square$

A partir disso, fica claro que todo funcional de X^* é contínuo em X .

Teorema 168. X^* é um espaço de Banach.

Demonstração. Sendo $\{f_i\} \subseteq X^*$ uma sequência de Cauchy, para todo $\epsilon > 0$, existe $\iota \in \mathbb{N}$ para o qual $i, j \geq \iota \Rightarrow \|f_i - f_j\|_{X^*} < \epsilon$, ou seja, $i, j \geq \iota \Rightarrow |\langle x, f_i - f_j \rangle| < \epsilon \|x\|_X$. Sabendo que, fixado $\epsilon_0 > 0$, é possível escolher, para cada $x \in X$, um $\epsilon_x > 0$ de modo que $\epsilon(x)\|x\|_X < \epsilon_0$, fica mostrado que $\{\langle x, f_i \rangle\}$ é uma sequência de Cauchy em $\mathbb{R}$. A partir disso, dado $x \in X$ existe um único $y \in \mathbb{R}$ tal que $\langle x, f_i \rangle \rightarrow y$, o que permite construir um funcional $f: X \rightarrow \mathbb{R}$ associando cada x ao correspondente y .

Afirmção. $f \in X^*$.

De fato, dados $x_1, x_2 \in X$ e $\alpha_1, \alpha_2 \in \mathbb{R}$, vale que

$$\begin{aligned} f(\alpha_1 x_1 + \alpha_2 x_2) &= \lim_{i \uparrow \infty} \langle \alpha_1 x_1 + \alpha_2 x_2, f_i \rangle = \alpha_1 \lim_{i \uparrow \infty} \langle x_1, f_i \rangle + \alpha_2 \lim_{i \uparrow \infty} \langle x_2, f_i \rangle \\ &= \alpha_1 f(x_1) + \alpha_2 f(x_2), \end{aligned}$$

o que mostra que f é linear. Ainda, dados $x \in X$ e $i, j \geq \iota$,

$$|\langle x, f_i - f \rangle| \leq |\langle x, f_i - f_j \rangle| + |\langle x, f_j - f \rangle| < \epsilon \|x\|_X + |\langle x, f_j - f \rangle|$$

assim, ao fazer $j \uparrow \infty$, segue que $|\langle x, f_i - f \rangle| \leq \epsilon \|x\|_X$ para todo $i \geq \iota$, o que permite afirmar que $f_i - f$ é lipschitziano em X para todo $i \geq \iota$, logo, $f = f_i - (f_i - f)$ é contínuo também.

Por fim, não é difícil ver que $i \geq \iota \Rightarrow \|f_i - f\|_{X^*} < \epsilon$, o que mostra que $f_i \rightarrow f$. Portanto, X^* é espaço de Banach. $\square$

Proposição 169. Todo subconjunto de X^* , pontualmente, limitado é limitado em norma.

Demonstração. Supondo que $K \subseteq X^*$ seja um conjunto ilimitado, em particular, para todo $i \in \mathbb{N}$, existe um $f_i \in K$ para o qual $\|f_i\|_{X^*} > \frac{3}{5} 8^{i+1}$. Ainda:

- Do **Lema 167**, $\sup \left| \left\langle B_X \left[0, \frac{1}{4} \right]; f_1 \right\rangle \right| \geq \frac{1}{4} \|f_1\|_{X^*}$, assim, da definição de supremo, existe um $x_1 \in B_X \left[0, \frac{1}{4} \right]$ satisfazendo $|\langle x_1, f_1 \rangle| \geq \sup \left| \left\langle B_X \left[0, \frac{1}{2} \right]; f_1 \right\rangle \right| - \frac{1}{16} \|f_1\|_{X^*} \geq \frac{3}{16} \|f_1\|_{X^*}$;
- Agora, do **Lema 167**, vale, também, que $\sup \left| \left\langle B_X \left[x_1, \frac{1}{16} \right]; f_2 \right\rangle \right| \geq \frac{1}{16} \|f_2\|_{X^*}$, assim, da definição de supremo, existe um $x_2 \in B_X \left[x_1, \frac{1}{16} \right]$ satisfazendo

$$|\langle x_2, f_2 \rangle| \geq \sup \left| \left\langle B_X \left[x_1, \frac{1}{16} \right]; f_2 \right\rangle \right| - \frac{1}{64} \|f_2\|_{X^*} \geq \frac{3}{64} \|f_2\|_{X^*};$$

- Continuando esse raciocínio, construímos uma sequência $\{x_i\} \subseteq X$ satisfazendo $\|x_i - x_{i-1}\| \leq \frac{1}{4^i}$ e $|\langle x_i, f_i \rangle| \geq \frac{3}{4^{i+1}} \|f_i\|_{X^*}$ para todo $i \in \mathbb{Z}_+$ — aqui, $x_0 = 0$.

Essa sequência é de Cauchy, já que

$$\begin{aligned}
j > i &\Rightarrow \|x_j - x_i\| \leq \sum_{\ell=1}^{j-i} \|x_{i+\ell} - x_{i+\ell-1}\| \leq \sum_{\ell=1}^{j-i} \frac{1}{4^{i+\ell}} \leq \frac{1}{4^{i+1}} \sum_{\ell=1}^{\infty} \frac{1}{4^{\ell-1}} = \frac{1}{3 \cdot 4^i}, \\
j = i &\Rightarrow \|x_j - x_i\| = 0, \\
j < i &\Rightarrow \|x_j - x_i\| \leq \frac{1}{3 \cdot 4^j},
\end{aligned}$$

o que mostra que ela tende a um certo ponto $x \in X$. A partir disso, segue que $\|x - x_i\| \leq 2^{-i}$ para todo $i \in \mathbb{Z}_+$, logo,

$$\frac{3}{4^{i+1}} \|f_i\|_{X^*} \leq |\langle x_i, f_i \rangle| \leq |\langle x_i - x, f_i \rangle| + |\langle x, f_i \rangle| \leq \frac{1}{3 \cdot 4^i} \|f_i\|_{X^*} + |\langle x, f_i \rangle|,$$

o que equivale a dizer que

$$|\langle x, f_i \rangle| \geq \left(3 - \frac{4}{3}\right) \frac{1}{4^{i+1}} \|f_i\|_{X^*} = \frac{5}{3} \frac{1}{4^{i+1}} \|f_i\|_{X^*} > 2^{i+1}.$$

Isso mostra que K não é, pontualmente, limitado. $\square$

Proposição 170 (página 3 de [9]). *Sendo $M \subseteq X$ um subespaço vetorial e $g : M \rightarrow \mathbb{R}$, existe $f \in X^*$ cuja norma coincide com a de g .*

Demonstração. Como $g \in M^*$, do **Lema 167**, segue que $\langle x, g \rangle \leq \|g\|_{X^*} \|x\|_X$ para todo $x \in M$, assim, do **Teorema 143**, existe $f \in X'$ estendendo g e satisfazendo $f(x) \leq \|g\|_{X^*} \|x\|_X$.

A linearidade de $f \in X^*$ permite, também, afirmar que $-f(x) \leq f(-x) \leq \|g\|_{X^*} \|x\|_X = \|g\|_{X^*} \|x\|_X$, então, $f \in X^*$ e $\|f\|_{X^*} \leq \|g\|_{X^*}$. Por outro lado,

$$\begin{aligned}
\|g\|_{X^*} &= \sup\{|g(x)| \mid x \in M \cap B_X[0, 1]\} = \sup\{|f(x)| \mid x \in M \cap B_X[0, 1]\} \\
&\leq \sup\{|f(x)| \mid x \in B_X[0, 1]\} = \|f\|,
\end{aligned}$$

portanto, $\|f\|_{X^*} = \|g\|_{X^*}$. $\square$

Corolário 171 (página 3 de [9]). *Para todo $x_0 \in X$, existe $f_0 \in X^*$ tal que $\|f_0\|_{X^*} = \|x_0\|_X$ e $\langle x_0, f_0 \rangle = \|x_0\|_X^2$.*

Demonstração. Considere o subespaço $[x_0] \subseteq X$ e o funcional $g : [x_0] \rightarrow \mathbb{R}$ dado por $g(tx_0) = t\|x_0\|_X^2$. Dados $\lambda_1, \lambda_2, t_1, t_2 \in \mathbb{R}$, vale que

$$\begin{aligned}
g(\lambda_1 t_1 x_0 + \lambda_2 t_2 x_0) &= g((\lambda_1 t_1 + \lambda_2 t_2)x_0) = (\lambda_1 t_1 + \lambda_2 t_2) \|x_0\|_X^2 = \lambda_1 t_1 \|x_0\|_X^2 + \lambda_2 t_2 \|x_0\|_X^2 \\
&= \lambda_1 g(t_1 x_0) + \lambda_2 g(t_2 x_0),
\end{aligned}$$

o que mostra que g é linear. Ainda, $|\langle tx_0, g \rangle| = \|x_0\|_X^2 |t| = \|x_0\|_X \|tx_0\|_X$, para todo $t \in \mathbb{R}$, o que mostra que $\|g\|_{X^*} = \|x_0\|_X$, logo, da **Proposição 170**, existe $f_0 \in X^*$ satisfazendo $\|f_0\|_{X^*} = \|x_0\|_X$ e $\langle tx_0, f_0 \rangle = t\|x_0\|_X^2$ para todo $t \in \mathbb{R}$. Em particular, para $t = 1$, vale $\langle x_0, f_0 \rangle = \|x_0\|_X^2$. $\square$

Lema 172 (página 14 de [21]). Se $f_1, f_2 \in X^*$, $Q_1 = \ker f_1$ e $Q_2 = \ker f_2$ são tais que $Q_1 \subseteq Q_2$, então, $f_2 = 0$ (consequentemente, $Q_2 = X$) ou existe $\lambda \neq 0$ tal que $f_1 = \lambda f_2$ (consequentemente, $Q_1 = Q_2$).

A.3.1 Topologias fracas em espaços normados

Topologia fraca

O corolário acima garante que X^* é separante, o que permite considerar a topologia fraca induzida em X por X^* . Por ser uma topologia de elevada importância, esta será chamada, simplesmente, de **topologia fraca de X** (com elementos denominados **abertos fracos**), enquanto que a topologia original de X (isto é, a induzida pela norma) será chamada **topologia forte de X** (com elementos denominados **abertos fortes**).

Como todos os elementos de X^* são contínuos na topologia forte de X , esta contém a topologia fraca de tal espaço.

Proposição 173 (página 51 de [11]). As topologias fraca e forte de X coincidem se, e somente se, X é dimensão finita.

Demonstração. Se X é de dimensão finita, então, existem $n \in \mathbb{N}$ e uma isometria linear $T: X \rightarrow \mathbb{R}^n$, o que mostra que, para $r > 0$, o conjunto $T(B_X(0, r))$ é uma vizinhança da origem em $\mathbb{R}^n$, o que significa que ela contém um conjunto da forma $\{u \in \mathbb{R}^n : |\langle u, e_i \rangle| < \epsilon \forall i \in \{1 : n\}\}$, em que $\epsilon > 0$ e os e_i são os vetores canônicos de $\mathbb{R}^n$. Isto implica que $B_X(0, r)$ contém um conjunto da forma $\{x \in X : |\langle T(x), e_i \rangle| < \epsilon \forall i \in \{1 : n\}\}$, que é um aberto fraco pelo fato de que $x \mapsto \langle T(x), e_i \rangle$ é um funcional linear contínuo para cada $i \in \{1 : n\}$.

Agora, seja X de dimensão infinita. Sendo W uma vizinhança fraca da origem, ele contém um conjunto da forma $\bigcap_{i=1}^n V(0, \varphi_i, \epsilon)$ com $\epsilon > 0$ e $\{\varphi_1, \dots, \varphi_n\} \subseteq X^*$. Ainda, a transformação linear $x \mapsto \sum_{i=1}^n \langle x, \varphi_i \rangle e_i$ não pode ser injetiva, já que, nesse caso, X seria isomorfo a um subespaço de dimensão finita de $\mathbb{R}^n$, o que mostra que existe $x_0 \neq 0$ para o qual $\langle x, \varphi_i \rangle = 0$ para todo $i \in \{1 : n\}$, consequentemente, $[x_0] \subseteq W$. Sabendo que $B_X(0, 1)$ não contém um subespaço não-trivial sequer, ela não pode ser aberto fraco de X . $\square$

Teorema 174 (página 52 de [11]). Sendo $C \subseteq X$ convexo, os fechos de C em relação às topologias fraca e forte coincidem.

Demonstração. Sendo C_w e C_s os fechos fraco e forte respectivamente, como cada aberto fraco é um aberto forte, cada fechado fraco é, também, um fechado forte, o que, pela definição do fecho como o menor fechado contendo o conjunto, mostra que $C_s \subseteq C_w$.

Agora, sendo $x \notin C_s$, do **Teorema 159**, existem $\zeta \in X^*$ e $\alpha \in \mathbb{R}$ satisfazendo $\langle x, \zeta \rangle < \alpha \leq \inf \langle C_s, \zeta \rangle$, assim, o conjunto $\zeta^{-1}([\alpha, \infty])$ é um fechado fraco contendo C que não contém x , o que mostra que $x \notin C_w$. $\square$

Corolário 175 (página 61 de [9]). *Sendo $x_0 \in X$, se $x_i \rightarrow x_0$, então, existe uma sequência $\{y_i\} \subseteq X$ tendendo a x_0 fortemente.*

Demonstração. Denotando, por C , a envoltória convexa de $\{x_i\}$, trata-se de um fechado fraco que é convexo, assim, do **Teorema 174**, tal conjunto é fechado forte, o que mostra que existe $\{y_i\}$ tendendo a x_0 fortemente. $\square$

Topologia fraca^{*}

Além de X^* , outro conjunto de funcionais lineares que é separante é o conjunto Φ das *avaliações de funcionais lineares*, ou seja, dado $\zeta \in X^*$, cada elemento de Φ é da forma $\zeta \mapsto \langle x, \zeta \rangle$ para $x \in X$. A correspondente topologia fraca induzida será chamada **topologia fraca^{*} de X^*** (com elementos denominados **abertos fracos^{*}**). Essa topologia é, em alguns casos, chamada *topologia da convergência pontual* pelo fato de que, do **Teorema 148**, uma sequência converge nessa topologia se, e somente se, ela converge pontualmente.

Apesar de ser possível mostrar que essa topologia é mais fraca que a topologia fraca de X^* , esse não será o foco do trabalho, de modo que a demonstração será omitida.

Proposição 176 (página 54 de [11]). *Supondo que $\Sigma \subseteq X^*$ seja um convexo não-vazio e fechado fraco^{*} e que $\zeta \in X^* \setminus \Sigma$, existem $x \in X$ e $\alpha \in \mathbb{R}$ tais que $\inf \langle x, \Sigma \rangle \leq \alpha < \langle x, \zeta \rangle$.*

Teorema 177 (página 55 de [11]). *Sendo V uma vizinhança forte da origem em X , o conjunto Γ dos funcionais lineares contínuos cuja imagem por V é subconjunto de $[-1, 1]$ é um compacto fraco^{*} em X^* .*

Demonstração. Chamando de P o conjunto dos funcionais de domínio em X , a topologia produto τ_p de P é, por definição, a topologia mais fraca que preserva a continuidade das aplicações $\zeta \mapsto \zeta(x)$, fica evidente que a topologia fraca^{*} de X^* coincide com a topologia induzida por τ_p em X^* . Ainda, como o limite fraco^{*} de funcionais lineares contínuos é um funcional linear contínuo e o conjunto $[-1, 1]$ é fechado, o conjunto Γ é fechado fraco^{*}. Mais ainda, fixando $\zeta \in \Gamma$, dado $x \in X$, existe $t_x > 0$ para o qual $t_x x \in V$, assim, $\langle x, \zeta \rangle \in \left[-\frac{1}{t_x}, \frac{1}{t_x}\right]$, o que mostra que

$$\Gamma \subseteq \prod_{x \in X} \left[-\frac{1}{t_x}, \frac{1}{t_x}\right].$$

O conjunto à direita é compacto pelo Teorema de Tihonov¹, portanto, Γ é compacto em τ_p , o que prova o teorema. $\square$

Corolário 178 (página 55 de [11]). *Todo conjunto fechado fraco^{*} e limitado é compacto fraco^{*}. Ainda, se X é um espaço de Banach, então, todo compacto fraco^{*} é fechado fraco^{*} e limitado.*

Demonstração. Supondo que K seja compacto fraco^{*}, como, pelo **Teorema 148**, tal espaço é de Hausdorff, K é um conjunto fechado. Ainda, a imagem de K por qualquer $\zeta \mapsto \langle x, \zeta \rangle$ é um conjunto compacto e, consequentemente, limitado, assim, pela **Proposição 169**, K é limitado.

¹Esse teorema e sua demonstração podem ser encontrados nas páginas 230 a 234 de [35].

Reciprocamente, supondo que K seja fechado fraco^{*} e limitado, existem $\zeta \in K$ e $r > 0$ tais que $K \subseteq B_{X^*}[\zeta, r]$, logo, como este é compacto fraco^{*} (**Teorema 177** combinado ao fato de que translações e homotetias são homeomorfismos), o conjunto K é compacto fraco^{*} também. $\square$

B Multifunções e limites

Este breve capítulo apresenta alguns conceitos e propriedades relativos a multifunções e seus limites que são necessários para compreender alguns resultados da dissertação. As referências, aqui, são [4] e [48], e os resultados serão formulados supondo que X é espaço de Banach real e Y é um espaço vetorial real normado.

Definição 179 (adaptada da página 64 de [48]). Uma aplicação $\Gamma : X \rightarrow \mathcal{G}(Y)$ (contradomínio é o conjunto das partes de Y) é denominada **multifunção** de X a Y e denotada por $\Gamma : X \rightsquigarrow Y$.

Definição 180 (página 34 de [4]). Seja $\Gamma : X \rightsquigarrow Y$:

- O **gráfico de Γ** é o conjunto dos pares $(x, y) \in X \times Y$ para os quais $y \in \Gamma(x)$. Ele será denotado por $\text{gr } \Gamma$;
- A **inversa de Γ** é a multifunção que associa cada $y_0 \in Y$ à seção y do gráfico de Γ relativa a y_0 . Ela será denotada por $\Gamma^{-1} : Y \rightsquigarrow X$;
- O **domínio efetivo de Γ** é o conjunto dos elementos de X cujas imagens por Γ não são vazias. Ele será denotado por $\text{def } \Gamma$;

Definição 181 (página 41 de [4]). • Dado $x_0 \in X$, uma multifunção $\Gamma : X \rightsquigarrow Y$ é dita **lipschitziana em torno de x_0** se existem $\ell > 0$ e uma vizinhança $U \subseteq \text{def } \Gamma$ de x_0 de modo que $x_1, x_2 \in U \Rightarrow \Gamma(x_1) \subseteq \Gamma(x_2) + \ell \|x_1 - x_2\| B_Y[0, 1]$.

- Dados $\Gamma : X \rightsquigarrow Y$ e $(x_0, y_0) \in \text{gr } \Gamma$, a multifunção Γ é dita ser **pseudo-lipschitziana em (x_0, y_0)** se existem $\ell > 0$ e vizinhanças $U \subseteq \text{def } \Gamma$ de x_0 e V de y_0 de modo que $x_1, x_2 \in U \Rightarrow \Gamma(x_1) \cap V \subseteq \Gamma(x_2) + \ell \|x_1 - x_2\| B_Y[0, 1]$.

Definição 182 (adaptada da página 63 de [48]). Sejam $\Gamma : X \rightsquigarrow Y$ e x_0 um ponto de acumulação de X .

- O **limite inferior de Γ para $x \rightarrow x_0$** é o conjunto dos pontos ξ para os quais, dada uma sequência $x_i \rightarrow x_0$, existe uma sequência $\xi_i \rightarrow \xi$ satisfazendo $\xi_i \in \Gamma(x_i)$ para todo $i \in \mathbb{N}$;
- O **limite superior de Γ para $x \rightarrow x_0$** é o conjunto dos pontos ξ para os quais existem sequências $x_i \rightarrow x_0$ e $\xi_i \rightarrow \xi$ satisfazendo $\xi_i \in \Gamma(x_i)$ para todo $i \in \mathbb{N}$.

Proposição 183 (adaptada da página 63 de [48]). Nas condições da *Definição 182*, vale

$$\liminf_{x \rightarrow x_0} \Gamma(x) = \left\{ \xi \in Y \mid \limsup_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0 \right\},$$

$$\limsup_{x \rightarrow x_0} \Gamma(x) = \left\{ \xi \in Y \mid \liminf_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0 \right\}.$$

Demonstração. Fixe $\xi \in \liminf_{x \rightarrow x_0} \Gamma(x)$. Dada $x_i \rightarrow x_0$, existe $\xi_i \rightarrow \xi$ satisfazendo $\xi_i \in \Gamma(x_i)$ para todo $i \in \mathbb{N}$, segue que $\text{dist}(\xi, \Gamma(x_i)) \leq \|\xi - \xi_i\| \rightarrow 0$, o que mostra que

$$\limsup_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0.$$

Reciprocamente, supondo que $\xi \in Y$ satisfaça $\limsup_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0$, defina $\alpha(\xi; \delta) = \sup\{\text{dist}(\xi, \Gamma(x)) \mid x \in B(x, \delta)\}$, tome $x_i \rightarrow x_0$ e escolha, para cada $i \in \mathbb{N}$, um $\xi_i \in \Gamma(x_i)$ de modo que

$$\text{dist}(\xi, \Gamma(x_i)) \leq \|\xi - \xi_i\|_Y < \frac{3}{2} \text{dist}(\xi, \Gamma(x_i))$$

se $\Gamma(x_i) \neq \emptyset$ ou $\xi_i = \xi$ caso contrário. Como $\alpha(\xi; \delta) \downarrow 0$ para $\delta \downarrow 0$ e $x_i \rightarrow x_0$, é possível afirmar que $\text{dist}(\xi, \Gamma(x_i)) \rightarrow 0$, e isso mostra que $\xi_i \rightarrow \xi$, portanto, $\xi \in \liminf_{x \rightarrow x_0} \Gamma(x)$.

Agora, tomando $\xi \in \limsup_{x \rightarrow x_0} \Gamma(x)$, existem seqüências $x_i \rightarrow x_0$ e $\xi_i \rightarrow \xi$ satisfazendo $\xi_i \in \Gamma(x_i)$ para todo $i \in \mathbb{N}$, assim, $\text{dist}(\xi, \Gamma(x_i)) \leq \|\xi - \xi_i\|_Y \rightarrow 0$, o que mostra que

$$\liminf_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0.$$

Reciprocamente, supondo que $\xi \in Y$ satisfaça $\liminf_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0$, escolha, para cada $i \in \mathbb{N}$, um $x_i \in Y$ cuja distância até x_0 seja menor que $\frac{1}{i}$ e de modo que $\text{dist}(\xi, \Gamma(x_i)) < \frac{2}{3i}$ e um $\xi_i \in \Gamma(x_i)$ satisfazendo

$$\text{dist}(\xi, \Gamma(x_i)) \leq \|\xi - \xi_i\| < \frac{3}{2} \text{dist}(\xi, \Gamma(x_i)) < \frac{1}{i}$$

se $\Gamma(x_i) \neq \emptyset$ ou $\xi_i = \xi$ caso contrário. A partir dessas escolhas, temos que $\xi_i \rightarrow \xi$, portanto, $\xi \in \limsup_{x \rightarrow x_0} \Gamma(x)$. $\square$

Proposição 184 (adaptada da página 63 de [48]). *Para todo $x_0 \in X$, os limites superior e inferior de Γ para $x \rightarrow x_0$ são fechados, e aquele inclui este.*

Demonstração. Para esta demonstração, serão usadas as caracterizações encontradas na **Proposição 183**.

- Tome uma seqüência $\{\xi_i\} \subseteq \liminf_{x \rightarrow x_0} \Gamma(x)$ tendendo a $\xi \in X$. Fixado $\epsilon > 0$, existe um $i_0 \in \mathbb{N}$ para o qual $\|\xi - \xi_{i_0}\| < \frac{\epsilon}{2}$, e existe um $\delta_0 > 0$ para o qual $\delta < \delta_0 \Rightarrow \alpha(\xi_{i_0}; \delta) < \frac{\epsilon}{2}$, logo,

$$\delta < \delta_0 \Rightarrow \alpha(\xi; \delta) \leq \|\xi_{i_0} - \xi\| + \alpha(\xi_{i_0}; \delta) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

o que mostra que $\xi \in \liminf_{x \rightarrow x_0} \Gamma(x)$ e, conseqüentemente, que tal conjunto é fechado;

- Tome uma seqüência $\{\xi_i\} \subseteq \limsup_{x \rightarrow x_0} \Gamma(x)$ tendendo a $\xi \in Y$ e defina

$$\beta(\xi; \delta) = \inf\{\text{dist}(\xi, \Gamma(x)) \mid x \in B(x, \delta)\}.$$

Fixado $\epsilon > 0$, existe um $i_0 \in \mathbb{N}$ para o qual $|\xi - \xi_{i_0}| < \frac{\epsilon}{2}$, e existe $\delta_0 > 0$ para o qual $\delta < \delta_0 \Rightarrow \beta(\xi_{i_0}; \delta) < \frac{\epsilon}{2}$, assim,

$$\delta < \delta_0 \Rightarrow \beta(\xi; \delta) \leq |\xi_{i_0} - \xi| + \beta(\xi_{i_0}; \delta) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

o que mostra que $\xi \in \limsup_{x \rightarrow x_0} \Gamma(x)$ e, consequentemente, que tal conjunto é fechado;

- Dado $\xi \in Y$, vale que

$$\begin{aligned} \xi \in \liminf_{x \rightarrow x_0} \Gamma(x) &\Rightarrow \limsup_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0 \Rightarrow \liminf_{x \rightarrow x_0} \text{dist}(\xi, \Gamma(x)) = 0 \\ &\Rightarrow \xi \in \limsup_{x \rightarrow x_0} \Gamma(x). \end{aligned}$$

□

C Envoltória afim e interior relativo

Este breve capítulo apresenta os conceitos de *envoltória afim* e *interior relativo* e algumas de suas propriedades encontrados em [42] e [51] e que são necessários para compreender o raciocínio usado na seção 2.4 e no Capítulo 4. Os resultados, aqui, serão feitos sob a suposição de que X é espaço vetorial topológico, localmente, convexo.

Definição 185 (página 1 de [51]). Um conjunto $A \subseteq X$ é dito **afim** se, dados $x, y \in A$ e $\lambda \in \mathbb{R}$, $\lambda x + (1 - \lambda)y \in A$.

Propriedade 186 (páginas 1 e 2 de [51]). • A afinidade de um conjunto $A \subseteq X$ é equivalente à existência de um subespaço $Y \subseteq X$ e de um ponto $a \in X$ para os quais $A = a + Y$;

• Se $\{A_i \mid i \in I\} \subseteq X$ é uma família de conjuntos afins, então, $A := \bigcap_{i \in I} A_i$ é, também, um conjunto afim.

Definição 187 (página 2 de [51]). Dado um conjunto $A \subseteq X$, o menor subconjunto afim de X contendo A é denominado **envoltória afim de A** .

Definição 188 (adaptada da página 2 de [51]). Dados $\{x_1, \dots, x_m\} \subseteq X$, uma **combinação afim de elementos** desse conjunto é um elemento da forma $\sum_{i=1}^m \lambda_i x_i$ com $\lambda_i \in \mathbb{R}$ para todo $i \in \{1 : m\}$ e $\sum_{i=1}^m \lambda_i = 1$.

Propriedade 189 (página 2 de [51]). A envoltória afim de A é o conjunto de todas as combinações afins de elementos de A .

Demonstração. Seja B o conjunto das referidas combinações.

Afirmção. B é um conjunto afim.

De fato, dados $x, y \in B$, existem $I, J \subseteq \mathbb{N}$ e $\{x_i \mid i \in I\}, \{y_j \mid j \in J\} \subseteq A$ e $\{\lambda_i \mid i \in I\}, \{\mu_j \mid j \in J\} \subseteq \mathbb{R}$ tais que as somatórias dos λ_i e dos μ_j é igual a um e que x é combinação afim dos x_i com correspondentes coeficientes λ_i e y é combinação afim dos y_j com correspondentes coeficientes μ_j , logo, dado $\lambda \in \mathbb{R}$,

$$(1 - \lambda)x + \lambda y = \sum_{i \in I} (1 - \lambda)\lambda_i x_i + \sum_{j \in J} \lambda \mu_j y_j,$$

$$\sum_{i \in I} (1 - \lambda)\lambda_i + \sum_{j \in J} \lambda \mu_j = (1 - \lambda) \sum_{i \in I} \lambda_i + \lambda \sum_{j \in J} \mu_j = (1 - \lambda) + \lambda = 1,$$

o que mostra que $(1 - \lambda)x + \lambda y \in B$.

Dito isso, como B contém A , segue que $B \supseteq \text{aff } A$. Para mostrar a inclusão contrária, usaremos indução no número de argumentos da combinação. Claramente, todo elemento de A é, também,

elemento de $\text{aff } A$, e, pelo fato de $\text{aff } A$ ser afim, qualquer combinação afim de dois elementos de A é elemento de $\text{aff } A$. Fixando $v \in \mathbb{N}$ e supondo que qualquer combinação afim de v elementos de A é elemento de $\text{aff } A$, sejam $x_1, \dots, x_{v+1} \in A$ e $\lambda_1, \dots, \lambda_{v+1} \in \mathbb{R}$ cuja somatória valha um. Se, ao menos, um dos λ_i é nulo, a combinação afim x dos x_i com coeficientes λ_i possui, na verdade, v ou menos elementos, o que já faz dela elemento de $\text{aff } A$; senão, ao menos um dos elementos é diferente de um (digamos, λ_{v+1}), logo,

$$x = \sum_{i=1}^{v+1} \lambda_i x_i = (1 - \lambda_{v+1}) \underbrace{\sum_{i=1}^v \frac{\lambda_i}{1 - \lambda_{v+1}} x_i}_{\in \text{aff } A} + \lambda_{v+1} \underbrace{x_{v+1}}_{\in \text{aff } A} \in \text{aff } A.$$

A partir disso, concluímos que $B \subseteq \text{aff } A$. □

Definição 190. Sendo $C \subseteq X$ um convexo, o **interior relativo de C** é o interior de tal conjunto em relação à topologia induzida na envoltória afim dele pela topologia do espaço. Ele será denotado por $\text{ri } C$.

Teorema 191 (página 45 de [42]). *Sendo $C \subseteq X$ um convexo, se $x \in \text{ri } C$ e $y \in \bar{C}$, então, $(1 - \lambda)x + \lambda y \in \text{ri } C$ para todo $\lambda \in [0, 1[$.*

Demonstração. Fixado $\lambda \in [0, 1[$, como $y \in \bar{C}$, toda bola centrada em y intercepta C , o que implica que $y \in C + B_X(0, \epsilon)$, logo,

$$\begin{aligned} B_X((1 - \lambda)x + \lambda y, \epsilon) &\subseteq (1 - \lambda)x + \lambda[C + B_X(0, \epsilon)] + B_X(0, \epsilon) \\ &= (1 - \lambda) \left[x + \frac{1 + \lambda}{1 - \lambda} B_X(0, \epsilon) \right] + \lambda C = (1 - \lambda) B_X \left(x, \frac{1 + \lambda}{1 - \lambda} \epsilon \right) + \lambda C. \end{aligned}$$

A interseção deste último conjunto com $\text{aff } C$ é subconjunto de $(1 - \lambda)C + \lambda C = C$ para $\epsilon > 0$ pequeno o suficiente, pois $x \in \text{ri } C$ por hipótese. □

Teorema 192 (página 46 de [42]). *Dado $C \subseteq X$ um convexo, $\overline{\text{ri } C} = \bar{C}$ e $\text{ri } \bar{C} = \text{ri } C$.*

Demonstração. Sabendo que $\text{ri } C \subseteq C$, vale que $\overline{\text{ri } C} \subseteq \bar{C}$. Por outro lado, dados $y \in \bar{C}$ e $x \in \text{ri } C$, o **Teorema 191** garante que o segmento ligando x a y está, exceto, possivelmente, pelo ponto y , contido em C , o que mostra que $y \in \overline{\text{ri } C}$. Portanto, $\overline{\text{ri } C} = \bar{C}$.

Agora, como as envoltórias afins de C e de $\bar{C}$ coincidem, é válido que $\text{ri } \bar{C} \supseteq \text{ri } C$. Reciprocamente, dado $z \in \text{ri } \bar{C}$, sendo $x \in \text{ri } C$, para valores $\mu > 1$ para os quais $\mu - 1$ é pequeno o suficiente, o ponto $y = (1 - \mu)x + \mu z = z - (\mu - 1)(x - z)$ pertence a $\text{ri } \bar{C} \subseteq \bar{C}$, e vale que $z = \left(1 - \frac{1}{\mu}\right)x + \frac{1}{\mu}y$, logo, o **Teorema 191** garante que $z \in \text{ri } C$. □

Corolário 193 (página 46 de [11]). *Sendo $C_1, C_2 \subseteq X$, $\bar{C}_1 = \bar{C}_2$ se, e somente se, $\text{ri } C_1 = \text{ri } C_2$.*

Proposição 194 (página 47 de [11]). *Sendo $C \subseteq X$ um convexo de interior relativo não-vazio, $z \in \text{ri } C$ se, e somente se, para cada $x \in C$, existe $\mu > 1$ tal que $(1 - \mu)x + \mu z \in C$ (em outras palavras, cada segmento contido em C tendo z como extremidade pode ser prolongado sem “escapar” de C).*

Demonstração. Claramente, se $z \in \text{ri } C$, então, a própria definição de interior relativo garante que cada segmento contido em C tendo z como extremidade pode ser prolongado sem “escapar” de C . Reciprocamente, sendo $z \in C$ tal que cada segmento contido em C tendo z como extremidade pode ser prolongado sem “escapar” de C e $x \in \text{ri } C$, tomando $\mu > 1$ e $y = (1 - \mu)x + \mu z$ como hipotetizado, segue que $z = \left(1 - \frac{1}{\mu}\right)x + \frac{1}{\mu}y$, logo, o **Teorema 191** garante que $z \in \text{ri } C$. $\square$

Teorema 195 (página 47 de [11]). *Sendo $\{C_i \mid i \in I\} \subseteq X$ convexos cujos interiores relativos possuem, ao menos, um ponto em comum,*

$$\overline{\bigcap_{i \in I} C_i} = \bigcap_{i \in I} \overline{C_i}.$$

Ainda, se I é finito, então,

$$\text{ri} \bigcap_{i \in I} C_i = \bigcap_{i \in I} \text{ri } C_i.$$

Demonstração. Fixado $x \in \bigcap_{i \in I} \text{ri } C_i$, para todo $y \in \bigcap_{i \in I} \overline{C_i}$, o vetor $(1 - \lambda)x + \lambda y \in \bigcap_{i \in I} \text{ri } C_i$ para todo $\lambda \in [0, 1[$ (**Teorema 191**), logo, como y é o limite desse vetor para $\lambda \uparrow 1$, vale que

$$\bigcap_{i \in I} \overline{C_i} \subseteq \overline{\bigcap_{i \in I} \text{ri } C_i} \subseteq \overline{\bigcap_{i \in I} C_i} = \bigcap_{i \in I} \overline{C_i},$$

o que mostra a igualdade relativa aos fechos e que $\bigcap_{i \in I} C_i$ e $\bigcap_{i \in I} \text{ri } C_i$ possuem o mesmo fecho. Do **Corolário 193**, os interiores relativos de tais conjuntos devem coincidir, o que mostra que

$$\text{ri} \bigcap_{i \in I} C_i \subseteq \bigcap_{i \in I} \text{ri } C_i.$$

Agora, sendo I finito e dado $z \in \bigcap_{i \in I} C_i$, seja ℓ um segmento de $\bigcap_{i \in I} C_i$ com extremidade em z . Fixado $i \in I$, ℓ pode ser prolongado sem “escapar” de cada um dos C_i (**Proposição 194**), assim, a finitude do conjunto I garante a existência de um prolongamento de ℓ que não “escapa” de $\bigcap_{i \in I} C_i$, o que mostra, pela **Proposição 194**, que $z \in \text{ri} \bigcap_{i \in I} C_i$. $\square$