



**COPPE/UFRJ**

UM ALGORITMO DE NEWTON DE PONTO INTERIOR E APLICAÇÕES NA  
FUNDIÇÃO ELETROMAGNÉTICA DE METAIS

Alfredo Canelas Botta

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Mecânica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Mecânica.

Orientadores: José Herskovits

Jean Rodolphe Roche

Rio de Janeiro

Abril de 2009

UM ALGORITMO DE NEWTON DE PONTO INTERIOR E APLICAÇÕES NA  
FUNDIÇÃO ELETROMAGNÉTICA DE METAIS

Alfredo Canelas Botta

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ  
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)  
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS  
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM  
CIÊNCIAS EM ENGENHARIA MECÂNICA.

Aprovada por:

---

Prof. José Herskovits, D.Ing.

---

Prof. Jean Rodolphe Roche, Ph.D.

---

Prof. Manuel Ernani de Carvalho Cruz, Ph.D.

---

Prof. José Claudio de Faria Telles, Ph.D.

---

Prof. Antonio André Novotny, D.Sc.

---

Prof. Berardi Sensale, D.Ing.

RIO DE JANEIRO, RJ – BRASIL

ABRIL DE 2009

Canelas Botta, Alfredo

Um Algoritmo de Newton de Ponto Interior e Aplicações na  
Fundição Eletromagnética de Metais/Alfredo Canelas Botta. –  
Rio de Janeiro: UFRJ/COPPE, 2009.

XXI, 189 p.: il.; 29,7cm.

Orientadores: José Herskovits

Jean Rodolphe Roche

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia  
Mecânica, 2009.

Referências Bibliográficas: p. 170 – 178.

1. Otimização de forma. 2. Algoritmos de ponto interior.
3. Fundição Electromagnética. I. Herskovits, José *et al.*. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Mecânica. III. Título.

# Agradecimentos

Agradeço aos professores José e Jean, pela orientação e pelo grande apoio para a realização deste trabalho.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro.

A Ana, e aos meus queridos pais, Carlos e Marisa, pelo carinho, incentivo e apoio constantes.

Aos amigos do Laboratório Optimize com os quais trabalhei junto, Sandro, Gabriel, Henry, Mario e Miguel.

Aos docentes do Programa de Engenharia Mecânica da UFRJ.

Aos docentes do “Instituto de Estructuras y Transporte” de Uruguai.

Ao pessoal administrativo do Programa de Engenharia Mecânica da UFRJ.

Às pessoas que colaboraram de forma direta ou indireta, por meio de incentivo, e troca de experiências.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

UM ALGORITMO DE NEWTON DE PONTO INTERIOR E APLICAÇÕES NA  
FUNDIÇÃO ELETROMAGNÉTICA DE METAIS

Alfredo Canelas Botta

Abril/2009

Orientadores: José Herskovits

Jean Rodolphe Roche

Programa: Engenharia Mecânica

Neste trabalho é apresentado um método para otimização em Engenharia Mecânica. O mesmo é baseado na utilização do Método dos Elementos de Contorno, na formulação SAND do problema de otimização e na utilização de um Algoritmo de Ponto Interior.

Um algoritmo de ponto interior baseado no algoritmo de direções viáveis FDIPA é proposto. O mesmo, aqui chamado FDIPA-H, permite utilizar a matriz hessiana nos sistemas lineares que definem a direção de busca. A demonstração de convergência global do algoritmo FDIPA-H é apresentada, assim como uma comparação com dois dos mais bem sucedidos algoritmos de otimização conhecidos para uma importante coleção de problemas teste.

O método proposto é utilizado na solução do problema de otimização de forma dos indutores utilizados na Fundição Eletromagnética de Metais. Duas formulações são apresentadas para este problema, assim como uma técnica geral para satisfazer as restrições geométricas de não superposição dos diferentes domínios.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

A NEWTON-BASED INTERIOR POINT ALGORITHM AND APLICATIONS TO  
ELECTROMAGNETIC CASTING OF MOLTEN METALS

Alfredo Canelas Botta

April/2009

Advisors: José Herskovits

Jean Rodolphe Roche

Department: Mechanical Engineering

This work presents a numerical method for optimization in Mechanical Engineering. It is based on the Boundary Element Method, the SAND formulation of the optimization problem and the application of an Interior Point Algorithm.

An interior point algorithm based on the feasible directions algorithm FDIPA is proposed. It was called FDIPA-H and allows the use of the hessian matrix in the linear systems that define the line search. A proof of global convergence of the algorithm is given, as well as a numerical comparison with two state of the art algorithms for a large collection of known benchmark problems.

The propose method is employed for the shape optimization of the inductors used in the Electromagnetic Casting. Two formulations are presented for this problem, as well as a general technique for taking into account the geometric constraints that avoid the overlapping between different domains.

# Sumário

<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xvii</b>
<b>Notação</b>	<b>xxi</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Preliminares</b>	<b>7</b>
2.1 Métodos de ponto interior para otimização não linear . . . . .	7
2.1.1 Problema de otimização não linear . . . . .	9
2.1.2 Convergência global de algoritmos . . . . .	11
2.1.3 Algoritmo de Newton . . . . .	13
2.1.4 Algoritmo FDIPA . . . . .	14
2.2 Método dos elementos de contorno . . . . .	17
2.2.1 Cálculo das matrizes elementares . . . . .	22
2.2.2 Equação para pontos no interior dos elementos . . . . .	23
2.2.3 Sistema de equações do MEC . . . . .	24
2.3 Formulação SAND . . . . .	25
2.3.1 Formulação SAND para otimização estrutural . . . . .	29
2.4 Diferenciação com relação à forma . . . . .	32
2.4.1 Movimentos e descrições material e espacial . . . . .	34
2.4.2 Derivadas com relação à forma de expressões integrais . . . . .	38
2.4.3 Análise de sensibilidade considerando o problema variacional . . . . .	39
2.4.4 Exemplo: Problema de Fundição Eletromagnética . . . . .	40

<b>3</b>	<b>Algoritmo de otimização FDIPA-H</b>	<b>43</b>
3.1	Convergência global do FDIPA . . . . .	44
3.1.1	Problema com restrições de igualdade . . . . .	51
3.2	Matrizes dos problemas de otimização . . . . .	56
3.2.1	Decomposição de matrizes simétricas indefinidas . . . . .	61
3.3	Algoritmo FDIPA-H . . . . .	63
3.3.1	Estudo da direção de Newton . . . . .	63
3.3.2	Modificação do sistema linear . . . . .	67
3.3.3	Definição do algoritmo FDIPA-H . . . . .	69
3.3.4	Convergência global do FDIPA-H . . . . .	72
3.4	Exemplos . . . . .	75
3.4.1	Alguns detalhes da implementação do algoritmo . . . . .	75
3.4.2	Resultados . . . . .	76
<b>4</b>	<b>Otimização usando o MEC e a formulação SAND</b>	<b>80</b>
4.1	Funções parcialmente separáveis . . . . .	82
4.2	Detalhes da implementação do MEC . . . . .	83
4.3	Minimização do volume com restrição na complacência . . . . .	85
4.4	Comparação com outros problemas de otimização . . . . .	99
<b>5</b>	<b>Problema de Fundição Eletromagnética</b>	<b>109</b>
5.1	O modelo matemático . . . . .	109
5.1.1	A formulação variacional . . . . .	111
5.2	O problema direto . . . . .	113
5.2.1	Algoritmo de resolução . . . . .	117
5.2.2	Exemplos . . . . .	118
5.3	O problema inverso . . . . .	124
5.3.1	Dois abordagens para o problema inverso . . . . .	126
5.3.2	Formulação SAND . . . . .	127
5.4	Otimização da posição dos indutores . . . . .	128
5.4.1	Problema discreto . . . . .	128
5.4.2	Exemplos . . . . .	130
5.5	Otimização da forma dos indutores . . . . .	142

5.5.1	Indutores . . . . .	143
5.5.2	Restrições geométricas . . . . .	144
5.5.3	Problema discreto . . . . .	146
5.5.4	Exemplos . . . . .	147
<b>6</b>	<b>Conclusões</b>	<b>167</b>
	<b>Referências bibliográficas</b>	<b>178</b>
<b>A</b>	<b>Resultados obtidos com o FDIPA-H</b>	<b>179</b>

# Lista de Figuras

2.1	Problema de elasticidade linear. . . . .	18
2.2	Comparação entre as técnicas de otimização clássica e SAND. . . . .	27
2.3	Definição do domínio transformado pelo campo vetorial $V$ . . . . .	33
3.1	Direção do algoritmo de Newton. . . . .	67
4.1	Tempo para o cálculo de $L_h$ e de suas derivadas. . . . .	90
4.2	Exemplo 1, configuração inicial da otimização de topologia. . . . .	92
4.3	Exemplo 1, <b>a)</b> Resultado da otimização de topologia para um 50% do volume final, <b>b)</b> Configuração inicial para otimização de forma. . . . .	93
4.4	Exemplo 1, resultados para a complacência igual a <b>a)</b> $2.22e-2$ kNm e <b>b)</b> $3.00e-2$ kNm. . . . .	93
4.5	Exemplo 1, evolução da função objetivo. . . . .	93
4.6	Exemplo 2, configuração inicial da otimização de topologia. . . . .	94
4.7	Exemplo 2, <b>a)</b> Resultado da otimização de topologia para um 40% do volume final, <b>b)</b> Configuração inicial para otimização de forma. . . . .	94
4.8	Exemplo 2, resultados para a complacência igual a <b>a)</b> $0.80e-2$ kNm e <b>b)</b> $1.00e-2$ kNm. . . . .	95
4.9	Exemplo 2, evolução da função objetivo. . . . .	95
4.10	Exemplo 3, configuração inicial da otimização de topologia. . . . .	96
4.11	Exemplo 3, <b>a)</b> Resultado da otimização de topologia para um 60% do volume final, <b>b)</b> Configuração inicial para otimização de forma. . . . .	96
4.12	Exemplo 3, resultados para a complacência igual a <b>a)</b> $2.52e-2$ kNm e <b>b)</b> $3.60e-2$ kNm. . . . .	96
4.13	Exemplo 3, evolução da função objetivo. . . . .	97
4.14	Exemplo 4, configuração inicial da otimização de topologia. . . . .	97

4.15	Exemplo 4, <b>a)</b> Resultado da otimização de topologia para um 40% do volume final, <b>b)</b> Configuração inicial para otimização de forma no exemplo 3. . . . .	98
4.16	Exemplo 4, resultado para a complacência igual a $3.42e-2$ kNm. . . . .	98
4.17	Exemplo 4, evolução da função objetivo. . . . .	99
4.18	Sistema de coordenadas locais em um ponto do contorno. . . . .	101
4.19	Exemplo 1, otimização da forma de um orifício. . . . .	102
4.20	Exemplo 1, tensões de von Mises - Problemas 1–3. . . . .	102
4.21	Exemplo 2, otimização do filete. . . . .	103
4.22	Exemplo 2, tensões de von Mises - Problema 1. . . . .	103
4.23	Exemplo 2, tensões de von Mises - Problema 2. . . . .	104
4.24	Exemplo 2, tensões de von Mises - Problema 3. . . . .	104
4.25	Exemplo 3, otimização de um pórtico plano. . . . .	105
4.26	Exemplo 3, tensões de von Mises - Problema 1. . . . .	105
4.27	Exemplo 3, tensões de von Mises - Problema 2. . . . .	106
4.28	Exemplo 3, tensões de von Mises - Problema 3. . . . .	106
5.1	Problema EMC. . . . .	110
5.2	Exemplo 1, <b>a)</b> forma inicial, <b>b)</b> forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	119
5.3	Exemplo 1, evolução de $\ G\ $ . . . . .	119
5.4	Exemplo 2, <b>a)</b> forma inicial, <b>b)</b> forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	120
5.5	Exemplo 2, evolução de $\ G\ $ . . . . .	120
5.6	Exemplo 3, <b>a)</b> forma inicial, <b>b)</b> forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	121
5.7	Exemplo 3, evolução de $\ G\ $ . . . . .	121
5.8	Exemplo 4, <b>a)</b> forma inicial, <b>b)</b> forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	122
5.9	Exemplo 4, evolução de $\ G\ $ . . . . .	122
5.10	Exemplo 5, <b>a)</b> forma inicial, <b>b)</b> forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	123
5.11	Exemplo 5, evolução de $\ G\ $ . . . . .	123

5.12	Exemplo 6, <b>a)</b> forma inicial, <b>b)</b> forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	124
5.13	Exemplo 6, evolução de $\ G\ $ . . . . .	124
5.14	Exemplo 1, posição dos indutores, <b>a)</b> distribuição inicial, <b>b)</b> distribuição final. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . .	131
5.15	Exemplo 1, evolução da função objetivo da primeira formulação. . . . .	131
5.16	Exemplo 1, evolução da função objetivo da segunda formulação. . . . .	132
5.17	Exemplo 2, configuração de indutores e forma final, <b>a)</b> primeira formulação, <b>b)</b> segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	133
5.18	Exemplo 2, evolução da função objetivo da primeira formulação. . . . .	133
5.19	Exemplo 2, evolução da função objetivo da segunda formulação. . . . .	134
5.20	Exemplo 3, posição dos indutores, <b>a)</b> distribuição inicial, <b>b)</b> distribuição final. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . .	134
5.21	Exemplo 3, evolução da função objetivo da primeira formulação. . . . .	135
5.22	Exemplo 3, evolução da função objetivo da segunda formulação. . . . .	135
5.23	Exemplo 4, configuração de indutores e forma final, <b>a)</b> primeira formulação, <b>b)</b> segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	136
5.24	Exemplo 4, evolução da função objetivo da primeira formulação. . . . .	136
5.25	Exemplo 4, evolução da função objetivo da segunda formulação. . . . .	137
5.26	Exemplo 5, forma objetivo e configuração inicial de indutores. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	137
5.27	Exemplo 5, configuração final de indutores e forma final, <b>a)</b> primeira formulação, <b>b)</b> segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	138

5.28	Exemplo 5, evolução da função objetivo da primeira formulação. . . . .	138
5.29	Exemplo 5, evolução da função objetivo da segunda formulação. . . . .	139
5.30	Exemplo 6, forma objetivo e configuração inicial de indutores. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	139
5.31	Exemplo 6, configuração final de indutores e forma final, <b>a)</b> primeira formulação, <b>b)</b> segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	140
5.32	Exemplo 6, evolução da função objetivo da primeira formulação. . . . .	140
5.33	Exemplo 6, evolução da função objetivo da segunda formulação. . . . .	141
5.34	Problema de otimização de forma dos indutores. . . . .	142
5.35	Forma paramétrica da seção transversal dos indutores. <b>1.</b> indutor retangular de quatro parâmetros; <b>2.</b> indutor de seis parâmetros com lados verticais parabólicos; <b>3.</b> indutor de seis parâmetros com lados horizontais parabólicos. . . . .	144
5.36	Restrições Geométricas. . . . .	145
5.37	Função $\psi$ para uma forma objetivo aproximadamente retangular. . . . .	146
5.38	Função $\psi$ para uma forma objetivo similar à letra “T”. . . . .	146
5.39	Exemplo 1, forma objetivo em equilíbrio com intensidades concentradas. Linha sólida: Forma em equilíbrio. Sinal de adição: Corrente positiva. Círculo: Corrente negativa. . . . .	149
5.40	Exemplo 1, configuração inicial e restrições geométricas dos exemplos Ex1a e Ex1b, <b>a)</b> exemplo Ex1a. <b>b)</b> exemplo Ex1b. Linha sólida: Curva $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa. . . . .	149
5.41	Solução do Exemplo 1, ambas as formulações, forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	150
5.42	Exemplo 1, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	150
5.43	Exemplo 1, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	151

5.44	Exemplo 2, configuração inicial e restrições geométricas dos exemplos Ex2a a Ex2d. Linha sólida: Curva $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa. . . . .	152
5.45	Solução do exemplo Ex2a, primeira formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	152
5.46	Solução do exemplo Ex2a, segunda formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	153
5.47	Solução do exemplo Ex2d, segunda formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	153
5.48	Exemplo 2, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	153
5.49	Exemplo 2, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	154
5.50	Exemplo 3, configuração inicial e restrições geométricas para os exemplos Ex3a e Ex3b. Linha sólida: Curva $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa. . . . .	155
5.51	Solução do exemplo Ex3a, primeira formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	155
5.52	Solução do exemplo Ex3b, segunda formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	156
5.53	Exemplo 3, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	156
5.54	Exemplo 3, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	157

5.55	Exemplo 4, configuração inicial e restrições geométricas para os exemplos Ex4a e Ex4b. Linha sólida: Curva $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa. . . . .	157
5.56	Solução do exemplo Ex4a, primeira formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	158
5.57	Solução do exemplo Ex4a, segunda formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	159
5.58	Exemplo 4, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	159
5.59	Exemplo 4, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	160
5.60	Exemplo 5, configuração inicial. Linha sólida: Curva $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa. . . . .	161
5.61	Solução do exemplo Ex5, primeira formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	161
5.62	Solução do exemplo Ex5, segunda formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	162
5.63	Exemplo 5, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	162
5.64	Exemplo 5, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	163
5.65	Exemplo 6, configuração inicial. Linha sólida: Curva $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa. . . . .	163

5.66	Solução do exemplo Ex6, primeira formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	164
5.67	Solução do exemplo Ex6, segunda formulação, <b>a)</b> solução e restrição geométrica, <b>b)</b> forma em equilíbrio e curvas de nível da função de fluxo magnético $\varphi$ . . . . .	164
5.68	Exemplo 6, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	165
5.69	Exemplo 6, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H. . . . .	165

# Lista de Tabelas

3.1	Resumo dos resultados obtidos para os problemas de Hock e Schittkowski.	78
3.2	Resumo dos resultados obtidos para os problemas do CUTE. . . . .	79
4.1	Parâmetros e valores obtidos para os exemplos apresentados. . . . .	107
5.1	Resumo dos resultados obtidos para os exemplos de otimização da posição dos indutores. . . . .	142
5.2	Resumo dos resultados obtidos para os exemplos de otimização de forma dos indutores. . . . .	166
A.1	Resultados obtidos para os problemas de Hock e Schittkowski. . . . .	179
A.2	Resultados obtidos para os problemas do CUTE, primeira parte. . . . .	182
A.3	Resultados obtidos para os problemas do CUTE, segunda parte. . . . .	186

# NOTAÇÃO

## Otimização não linear

$\mathbb{N}$	conjunto dos números naturais,
$\mathbb{R}$	conjunto dos números reais,
$\mathbb{R}^n$	conjunto dos vetores reais de dimensão $n$ ,
$\mathbb{R}^{m \times n}$	conjunto das matrizes reais de dimensões $m \times n$ ,
$\mathbf{x}$	vetor das variáveis primais ou variáveis de projeto, $\mathbf{x} \in \mathbb{R}^n$ .
$f$	função objetivo. $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,
$\mathbf{g}$	função das restrições de desigualdade. $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,
$\mathbf{h}$	função das restrições de igualdade. $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ ,
$\nabla f$	gradiente da função $f$ , $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,
$\nabla \mathbf{g}$	gradiente da função $\mathbf{g}$ , $\nabla \mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$ ,
$\nabla \mathbf{h}$	gradiente da função $\mathbf{h}$ , $\nabla \mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^{p \times n}$ ,
$\Omega$	região viável, conjunto dos pontos que satisfazem as restrições,
$\boldsymbol{\lambda}$	multiplicadores de Lagrange das restrições de desigualdade, $\boldsymbol{\lambda} \in \mathbb{R}^m$ ,
$\boldsymbol{\mu}$	multiplicadores de Lagrange das restrições de igualdade, $\boldsymbol{\mu} \in \mathbb{R}^p$ ,
$L$	função lagrangiana. $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ ,
$\mathbf{H}$	matriz hessiana. $\mathbf{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}^{n \times n}$ ,
$\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k$	valores das variáveis na iteração $k$ ,
$\mathbf{d}_x$	direção de busca. $\mathbf{d}_x \in \mathbb{R}^n$ ,
$t$	passo, $t \in \mathbb{R}$ .

## Método dos elementos de contorno

$\Omega$	corpo do espaço euclidiano tridimensional,
$\Gamma$	contorno de $\Omega$ ,
$x, \xi$	pontos do espaço euclidiano,
$\mathbf{u}$	vetor dos deslocamentos,
$\epsilon$	tensor de deformações,
$\sigma$	tensor de tensões,
$\mathbf{n}$	vetor unitário normal a $\Gamma$ ,
$\mathbf{b}$	forças de volume,
$\mathbf{p}$	forças de superfície,
$\mathbb{C}$	tensor elástico,
$G$	módulo de cisalhamento,
$\nu$	coeficiente de Poisson,
$\mathbf{u}_{ij}^*$	deslocamentos fundamentais,
$\mathbf{p}_{ij}^*$	forças de superfície fundamentais,
$\mathbf{h}_{ij}, \mathbf{g}_{ij}$	matrizes elementares correspondente ao nó $i$ e ao elemento $j$ ,
$\mathbf{H}, \mathbf{G}$	matrizes globais do MEC.

## Algoritmo FDIPA-H

$\sigma_k(A)$	$k$ -ésimo valor próprio da matriz $\mathbf{A}$ , ordenados em forma crescente,
$\ \mathbf{A}\ $	norma espectral da matriz $\mathbf{A}$ ,
$\mathcal{T}$	espaço tangente,
$\mathbf{I}_n$	matriz identidade de dimensão $n$ ,
$\phi_c$	função potencial para o parâmetro $c$ , $\phi_c : \mathbb{R}^n \rightarrow \mathbb{R}$ ,
$\mathbf{d}_x^\alpha$	direção de Newton do algoritmo FDIPA-H,
$\mathbf{d}_x^\beta$	direção de desvio,

$\mathbf{d}_x^1$	direção de otimalidade,
$\mathbf{d}_x^2$	direção de viabilidade,
$\gamma, \epsilon_1, \epsilon_2$	parâmetros para modificação da matriz hessiana,
$\xi, \varphi$	parâmetros do desvio,
$\eta, \nu$	parâmetros da busca linear.

### Fundição Eletromagnética

$\mathbf{B}$	campo magnético,
$\mathbf{j}_0$	densidade de corrente elétrica,
$\mu_0$	permeabilidade magnética do vácuo,
$\omega$	seção transversal do metal líquido,
$\Omega$	domínio exterior ao metal líquido, $\Omega = \mathbb{R}^2 \setminus \omega$
$\Gamma$	contorno do metal líquido,
$\mathbf{n}$	vetor unitário normal a $\Gamma$ ,
$\mathcal{C}$	curvatura de $\Gamma$ ,
$\sigma$	tensão superficial do metal líquido,
$\bar{p}$	diferença entre as pressões internas e externas,
$S_0$	área do metal líquido,
$\varphi$	função de fluxo magnético,
$E(\omega)$	energia total de $\omega$ ,
$P(\omega)$	perímetro de $\omega$ ,
$m$	número de indutores,
$\Theta_p$	seção transversal do indutor $p$ ,
$I$	valor de referência da intensidade ou densidade da intensidade da corrente elétrica,
$\alpha_p$	valores adimensionados da densidade de corrente elétrica,

$W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2)$	conjunto das funções uniformemente limitadas de $\mathbb{R}^2$ em $\mathbb{R}^2$ com gradiente uniformemente limitado,
$V, Z$	funções em $W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2)$ ,
$Id$	função identidade em $W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2)$ ,
$\omega_V, (\Gamma_V)$	domínio perturbado $(Id + V)(\omega)$ ,
$\omega^*, (\Gamma^*)$	forma objetivo,
$\omega_Z, (\Gamma_Z)$	domínio perturbado $(Id + Z)(\omega^*)$ ,
$\Gamma^h$	aproximação linear por partes de $\Gamma^*$ ,
$\mathbf{u}$	vetor dos parâmetros de forma do contorno do metal líquido,
$\Gamma_{\mathbf{u}}$	contorno do metal líquido correspondente a $\mathbf{u}$ ,
$\mathbf{u}_p$	vetor dos parâmetros de forma dos indutores,
$\mathbf{q}$	vetor das variáveis nodais de estado,
$\mathbf{p}$	vetor das pressões nodais fictícias,
$\psi$	função que define as restrições geométricas.

# Capítulo 1

## Introdução

Neste trabalho é apresentado um método numérico para resolver problemas de otimização em Engenharia Mecânica. O método é baseado na formulação de *Análise e Otimização Simultâneas (SAND, do inglês Simultaneous Analysis and Design)* do problema de otimização e na utilização do *Método dos Elementos de Contorno (MEC)* para a análise estrutural. Para a resolução dos problemas de otimização, utiliza-se um *Algoritmo de ponto interior*.

Antes de passar a descrever o conteúdo dos capítulos seguintes, é conveniente relatar, brevemente, as motivações que conduziram ao estudo do método proposto. No início optou-se por utilizar o MEC em problemas de otimização de forma em elasticidade linear. O MEC é um método numérico amplamente utilizado como ferramenta de análise para diversos problemas da engenharia. Ele é baseado em uma fórmula integral, equivalente às equações de campo do problema, definida no contorno do domínio, que em problemas de elasticidade é conhecida como identidade de Somigliana. Portanto, a característica mais importante do método é que, na sua implementação numérica, somente necessita da discretização do contorno do domínio do problema.

Esta característica sugeria que o MEC devia ser uma ferramenta ideal para otimização de forma, já que a malha do contorno é suficiente, ao mesmo tempo, tanto para representar aproximadamente a geometria da estrutura que se pretende otimizar, quanto para realizar a análise estrutural. Porém, o MEC pode ser aproveitado somente nos casos onde qualquer das considerações sobre a representação geométrica do problema esteja relacionada com o contorno, e nos quais as restrições geométricas e mecânicas possam ser expressadas com relação às funções definidas no contorno, ou com relação à restrição ao

contorno das funções que estejam definidas em todo o domínio. Somente nestes casos é possível prescindir completamente de cálculos destinados a obter valores no interior. Porém, estes casos são usuais. De fato, tanto as funções definidas no contorno quanto diversos funcionais definidos como integrais no domínio ou no contorno podem ser calculados utilizando os resultados fornecidos pelo MEC. Isto também é válido tratando-se da análise de sensibilidade.

Assim, utilizando o MEC evita-se a tarefa, não de menor importância em termos de custo computacional, de discretizar o domínio do problema. Mais importante ainda, evita-se tratar com o usualmente enorme número de variáveis nodais, comum nos métodos de domínio como o *Método dos Elementos Finitos* (MEF) ou o *Método das Diferenças Finitas* (MDF), correspondente aos nós interiores ao domínio.

De fato, após trabalhar um tempo com o MEC, tem-se a sensação de que as variáveis correspondentes ao interior do domínio, ou inclusive a discretização do domínio pertencentes a outros métodos de análise, são puramente auxiliares. Esta sensação é obviamente incorreta e somente se explica pelo sucesso do MEC no cálculo aproximado das funções no contorno. Mais correto é dizer que o MEC produz uma redução significativa do número de variáveis de estado necessárias para este propósito específico de análise para otimização de forma, quando comparado com os métodos de domínio, que podem ser muito mais convenientes para outros propósitos e em outras circunstâncias.

Por outro lado, o MEC como ferramenta de análise é também muito indicado para problemas de domínio infinito, sendo ótimo no caso freqüente em que o contorno de interesse desse domínio é limitado. Por esta causa tem sido muito utilizado em mecânica da fratura, acústica, análise de túneis, interação solo-estrutura e em problemas de dinâmica em sismologia. Também é conhecido que os resultados obtidos com o método são, na maioria dos casos, consideravelmente mais precisos que os obtidos utilizando outras ferramentas de análise, como o MEF ou o MDF, isto faz do MEC uma boa alternativa por exemplo nos problemas de concentração de tensões em elasticidade.

Por todas estas razões, já existiam no começo do trabalho numerosos aportes sobre a aplicação do MEC na otimização de forma, assim como trabalhos sobre análise de sensibilidade utilizando o MEC, motivados pela sua aplicação a estes problemas.

Porém, existia uma diferença fundamental entre as pesquisas realizadas até o momento e as que eram iniciadas neste trabalho. Essa diferença estava relacionada com os

então recentes resultados obtidos na área da otimização não linear, que indicavam que os algoritmos de ponto interior que utilizavam a informação relativa às derivadas segundas das funções que definem o problema de otimização eram de longe os mais convenientes para resolver problemas de otimização de *grande porte*, isto é, problemas com grande número de variáveis e/ou restrições. De fato, diversos estudos os apontavam como os mais robustos e eficientes algoritmos de otimização de propósito geral. Por outro lado, nestes problemas de grande porte os métodos de ponto interior apresentam-se como os de melhor escalabilidade.

Iniciaram-se então as pesquisas com a convicção de que estes resultados deviam ser considerados, e que um método computacional para otimização de forma que utilize o MEC devia utilizar também estes algoritmos de otimização.

Cabe questionar-se se a utilização dos algoritmos de otimização mais convenientes para solução de problemas de grande porte é necessária, visto que poderão ser considerados unicamente os problemas que podem ser analisados utilizando o MEC. Portanto, tanto a função objetivo quanto as restrições geométricas e mecânicas do problema de otimização serão dependentes somente das variáveis de projeto e das variáveis de estado correspondentes ao contorno. Visto que o MEC produz uma significativa redução do número de variáveis, o problema de otimização será consideravelmente menor que no caso de utilizar um método de análise com discretização de domínio. De fato, muitos dos problemas resolvidos e apresentados nos capítulos seguintes podem ser considerados de pequeno tamanho, enquanto que, se um método de análise com discretização de domínio tivesse sido utilizado, o mesmo problema seria de grande porte e, certamente, de difícil solução em um computador de escritório usual.

Embora essas considerações não sejam incorretas, persiste o interesse de utilizar estes algoritmos de otimização por duas razões: A primeira é a evidência “empírica” que diz que uma vez que uma ferramenta computacional está disponível, esta sempre é utilizada até o limite da capacidade do computador empregado. Na verdade, sempre existem ou aparecem razões legítimas para resolver problemas cada vez maiores, por exemplo, a necessidade de representação de geometrias complexas, requerimentos de precisão, etc. Visto de outro ponto de vista, o tamanho dos problemas que são resolvidos tem sido, desde sempre, limitado pela capacidade dos computadores existentes. Portanto, já que esta capacidade está sempre em contínuo crescimento, o estudo dos algoritmos que pos-

sibilitarem a resolução dos problemas de maior porte que possam ser considerados deve estar sempre na atenção dos pesquisadores. A segunda razão é que, até onde chega o conhecimento do autor, a utilização do MEC em conjunto com um algoritmo de otimização de ponto interior para problemas de otimização de forma nunca tinha sido realizada. Portanto existia o desejo de conhecer as características que teria um método que integrasse as duas técnicas. Embora nunca se tenha pensado em substituir nenhuma das técnicas clássicas de otimização de forma, tinha-se a esperança de obter uma ferramenta final muito adequada para alguns problemas específicos.

Desde o início começou-se utilizando a formulação SAND, visto que esta reduz significativamente o custo computacional do cálculo das derivadas segundas, necessárias para o algoritmo de otimização. A formulação SAND considera as variáveis de estado como incógnitas do problema de otimização e inclui as equações de estado como restrições de igualdade. A vantagem imediata desta técnica aparece em problemas não lineares, nos quais a análise deve ser feita por um processo iterativo. Como o algoritmo de otimização é também iterativo e em cada passo utiliza os resultados da análise estrutural e da análise de sensibilidade, a resolução de um problema de otimização não linear pelo método clássico resulta em dois processos iterativos aninhados, o de otimização e o de análise, que podem resultar em um custo computacional muito elevado. A formulação SAND, por incluir as variáveis de estado no problema de otimização, evita o processo iterativo de análise em cada iteração do algoritmo de otimização.

Outras vantagens da formulação SAND têm sido reconhecidas, por exemplo, a maior simplicidade da análise de sensibilidade, característica pelo qual foi inicialmente considerado e, no caso da utilização do MEF, a natureza esparsa das matrizes de derivadas das funções que definem o problema de otimização. Esta última vantagem tem permitido obter um ganho computacional significativo inclusive na aplicação da formulação SAND em problemas de análise linear.

Por causa destas observações, considerou-se viável e relevante na área de otimização estrutural a elaboração de uma técnica numérica que, baseada na formulação SAND e o MEC, utilize algoritmos de otimização de ponto interior para resolver eficientemente problemas de otimização de forma não lineares.

É importante notar que, assim como o uso de algoritmos de ponto interior que necessitam das derivadas segundas torna conveniente a utilização da formulação SAND,

reciprocamente, esta última torna obrigatório o uso de um algoritmo de otimização apropriado para problemas de grande porte, visto que eleva consideravelmente o número de variáveis pela adição das variáveis de estado ao problema de otimização, e eleva consideravelmente o número de restrições pela adição das equações de estado. Por outro lado, observa-se que o uso da formulação SAND conduz a problemas de otimização mal condicionados, para os quais os algoritmos de otimização que utilizam derivadas segundas são mais apropriados.

Cabe mencionar que o sucesso da formulação SAND foi possível em razão das características específicas observadas nas funções que definem o problema de otimização. Quando o MEF foi utilizado para a análise, a resolução numérica do problema de otimização foi possível por causa da natureza esparsa das matrizes de derivadas das funções do problema. A esparsidade das matrizes de derivadas contribui enormemente na redução do custo computacional associado ao cálculo e armazenamento de matrizes de derivadas e do custo de solução dos sistemas lineares do algoritmo de otimização. Desafortunadamente, as matrizes de derivadas não são esparsas quando o MEC é utilizado. Esta desvantagem do MEC fez necessário estudar as seguintes questões:

1. Custo computacional associado ao cálculo das matrizes de derivadas das funções do problema de otimização.
2. Possibilidade do cálculo das derivadas segundas.
3. Custo computacional associado à solução dos sistemas lineares do algoritmo de otimização.

Os itens 1 e 2 determinam a viabilidade da utilização do MEC, em quanto que o item 3 define a sua conveniência com relação a outros métodos de análise. A redução do número de variáveis do MEC com relação aos métodos de domínio é uma vantagem inicial. Os casos onde esta redução seja significativa ou, mais precisamente, os casos onde o MEC seja a ferramenta de análise mais conveniente, provavelmente serão os mais indicados para aplicar esta técnica de otimização.

A necessidade de utilizar um algoritmo de otimização eficiente para problemas de otimização de grande porte motivou um estudo sobre um algoritmo de otimização. A partir do algoritmo quasi-Newton de ponto interior e direções viáveis *FDIPA*, desenvolveu-se uma extensão do mesmo, aqui chamada *FDIPA-H*, que permite utilizar a matriz hessiana

do problema de otimização no lugar da matriz quasi-Newton BFGS da implementação original. A razão da escolha do algoritmo FDIPA, cujo nome provem do inglês *Feasible Directions Interior Point Algorithm*, é que este tinha sido utilizado com muito sucesso em alguns dos problemas aqui tratados e outros da Engenharia Mecânica.

Nos seguintes capítulos são descritos, de forma mais detalhada, os conceitos aqui expostos brevemente. O Capítulo 2 é uma breve revisão dos métodos de ponto interior para otimização não linear, focalizada no algoritmo FDIPA, o MEC para problemas de elasticidade linear e a formulação SAND para otimização estrutural. No Capítulo 3 é descrito o algoritmo de otimização FDIPA-H que utiliza as derivadas segundas das funções que definem o problema. Nesse capítulo é mostrada a convergência global da sequência gerada pelo algoritmo FDIPA-H para pontos de Karush-Kuhn-Tucker do problema de otimização, considerando hipóteses similares às consideradas no caso do algoritmo FDIPA. No Capítulo 4 são discutidas as características das funções que definem o problema de otimização, no caso em que a formulação SAND é utilizada e o MEC é empregado para a análise estrutural. Nesse capítulo, os resultados apresentados são particularizados para o caso de otimização de forma em elasticidade linear. No Capítulo 5 é apresentada uma aplicação do método proposto ao problema de *Fundição Eletromagnética* ou *EMC*, do inglês *Electromagnetic Casting*. O método proposto é especialmente indicado para este problema por causa da não linearidade das equações de equilíbrio e a necessidade de considerar um domínio infinito. Por último, no Capítulo 6 são descritas as conclusões deste trabalho e são sugeridos assuntos para futuras pesquisas.

# Capítulo 2

## Preliminares

### 2.1 Métodos de ponto interior para otimização não linear

Em diversas áreas da engenharia aparecem problemas de otimização complexos e de grande porte. Isto acontece, por exemplo, quando as variáveis e as funções que definem o problema de otimização derivam da discretização de corpos materiais contínuos. Com o objetivo de resolver estes problemas de otimização foram estudadas diversas técnicas baseadas nos métodos de ponto interior utilizados na otimização não linear.

Os métodos de ponto interior têm sido intensamente estudados a partir da apresentação do algoritmo de Karmarkar [1], o primeiro algoritmo de complexidade polinomial e eficiente na prática capaz de ser um forte competidor do algoritmo SIMPLEX [2] na resolução do problema de programação linear. A partir da apresentação do algoritmo de Karmarkar, muitos pesquisadores têm se dedicado ao estudo dos métodos de ponto interior, produzindo um grande número de trabalhos. Os métodos de ponto interior têm produzido um grande avanço nas áreas de otimização linear e não linear. Em alguns casos têm se apresentado como as melhores ferramentas para resolver importantes classes de problemas, e em outros casos têm produzido uma forte competência que derivou em avanços importantes em outras classes de métodos [3].

Os algoritmos de ponto interior são considerados hoje os mais robustos e eficientes entre os métodos de otimização de propósito geral. Diversos estudos mostram que estes métodos constituem poderosas ferramentas para a resolução de problemas de otimização não linear de grande escala [4–6]. Nestes problemas os métodos de ponto interior apresentam-se como os métodos de melhor escalabilidade [4, 6].

De acordo com Tapia [7], o papel fundamental da metodologia de ponto interior é produzir, de forma significativa e efetiva, um sistema de equações não lineares que represente suficientemente bem o problema de otimização com restrições de desigualdade de forma tal que a aplicação do método de Newton nesse sistema seja efetiva e bem-sucedida.

Na atualidade os métodos de ponto interior de tipo Primal-dual (métodos onde a teoria de dualidade desempenha um importante papel e é dado um tratamento igual às variáveis primais e duais) são considerados os mais promissores, de modo que grande parte da pesquisa em métodos de ponto interior está concentrada no estudo destes métodos.

Em geral, a única forma de resolver o *problema de otimização não linear* apresentado na Seção 2.1.1 é por meio de um método iterativo. Os métodos de ponto interior compreendem um conjunto de técnicas iterativas para achar um mínimo local do problema. Os mesmos geram uma seqüência de pontos convergente para um mínimo local e que satisfazem em forma estrita as restrições de desigualdade do problema.

Algumas das vantagens dos métodos de ponto interior são:

- Permitem resolver problemas esparsos ou com outro tipo de estrutura de forma muito eficiente. É simples adaptá-los para tomar o máximo proveito das características destes problemas [2, 3, 8].
- São muito convenientes para executar em computadores de arquitetura paralela, pois podem ser implementados de forma a aproveitar ao máximo o potencial destes computadores [2, 3, 8].

Neste capítulo, na Seção 2.1.1 é apresentada brevemente o problema de otimização não linear, as definições e teoremas principais. A teoria de convergência global de algoritmos de Zangwill é descrita brevemente na Seção 2.1.2, a mesma será de fundamental importância na demonstrações do Capítulo 3.3.3. Na Seção 2.1.3 é apresentado o método de Newton, o método iterativo para solução de sistemas de equações não lineares no qual estão baseados a maioria dos métodos de ponto interior. Por último, na Seção 2.1.4 é descrito o método de ponto interior e direções viáveis FDIPA.

### 2.1.1 Problema de otimização não linear

Um problema de otimização consiste na minimização ou maximização de uma determinada função, chamada de *função objetivo*, sujeito a restrições nas suas variáveis. O problema geral de otimização não linear pode ser formulado precisamente como:

$$\begin{aligned} & \text{minimize} && f(\mathbf{x}), \\ & \text{sujeito a:} && \mathbf{g}(\mathbf{x}) \geq 0, \\ & && \mathbf{h}(\mathbf{x}) = 0. \end{aligned} \tag{2.1}$$

Na Eq. (2.1)  $\mathbf{x} \in \mathbb{R}^n$  é o vetor das variáveis do problema, ou variáveis de projeto,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é a função objetivo,  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é a função que define as restrições de desigualdade e  $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$  é a função que define as restrições de igualdade. As funções  $f$ ,  $\mathbf{g}$  e  $\mathbf{h}$  devem ser funções contínuas com derivadas primeiras e segundas contínuas.

**Definição 2.1** *A região viável  $\Omega$  do Problema (2.1) é o conjunto de pontos que satisfazem as restrições:*

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0, \mathbf{h}(\mathbf{x}) = 0\}. \tag{2.2}$$

**Definição 2.2** *O ponto  $\mathbf{x}^*$  é um mínimo local do Problema (2.1) se existe uma vizinhança  $\mathcal{N}$ , com  $\mathbf{x}^* \in \mathcal{N}$ , tal que a desigualdade  $f(\mathbf{x}) \geq f(\mathbf{x}^*)$  é satisfeita para todo ponto  $\mathbf{x}$  pertencente à região  $\Omega \cap \mathcal{N}$ .*

**Definição 2.3** *Dado um ponto  $\mathbf{x} \in \Omega$ , o conjunto das restrições de desigualdade ativas é:*

$$A(\mathbf{x}) = \{i \mid 1 \leq i \leq m, \mathbf{g}_i(\mathbf{x}) = 0\}. \tag{2.3}$$

**Definição 2.4** *Dado um ponto  $\mathbf{x} \in \Omega$ , o Requisito de Independência Linear das Restrições (RILR) é satisfeito em  $\mathbf{x}$  se os gradientes das restrições de igualdade e de desigualdade ativas são linearmente independentes, ou seja, o conjunto  $\{\nabla \mathbf{g}_i(\mathbf{x}) \mid i \in A(\mathbf{x}), \nabla \mathbf{h}_i(\mathbf{x}) \mid 1 \leq i \leq p\}$  é linearmente independente.*

Neste trabalho será sempre assumido que o requisito RILR, definido acima é válido. Outras condições mais fracas, como por exemplo o Requisito de Mangasarian-Fromovitz [9], não serão consideradas aqui.

**Definição 2.5** Seja  $\mathbf{x} \in \mathbb{R}^n$ . As condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker (KKT) são satisfeitas no ponto  $\mathbf{x}$  se existem vetores  $\boldsymbol{\lambda} \in \mathbb{R}^m$  e  $\boldsymbol{\mu} \in \mathbb{R}^p$  que satisfazem:

$$\nabla f(\mathbf{x}) - \sum_{i=1}^m \lambda_i \nabla \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^p \mu_i \nabla \mathbf{h}_i(\mathbf{x}) = 0, \quad (2.4)$$

$$\mathbf{g}_i(\mathbf{x}) \lambda_i = 0, \quad 1 \leq i \leq m, \quad (2.5)$$

$$\mathbf{h}(\mathbf{x}) = 0, \quad (2.6)$$

$$\mathbf{g}(\mathbf{x}) \geq 0, \quad (2.7)$$

$$\boldsymbol{\lambda} \geq 0. \quad (2.8)$$

Os vetores  $\boldsymbol{\lambda}$  e  $\boldsymbol{\mu}$ , quando existem, são chamados multiplicadores de Lagrange.

**Teorema 2.6** (Condições necessárias de otimalidade de primeira ordem) Seja  $\mathbf{x}$  um mínimo local do Problema (2.1) onde é satisfeito o requisito RILR. Então, as condições de otimalidade de KKT são satisfeitas no ponto  $\mathbf{x}$  e, além disso, os vetores  $\boldsymbol{\lambda}$  e  $\boldsymbol{\mu}$  são únicos.

Prova: Ver na referência [8, Teorema 12.1 e Seção 12.3] e também em [9].  $\square$

**Definição 2.7** A função lagrangiana para o Problema (2.1) é a função  $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$  definida como:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i=1}^m \lambda_i \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^p \mu_i \mathbf{h}_i(\mathbf{x}). \quad (2.9)$$

Com esta definição a Eq. (2.4) pode ser escrita de forma mais compacta como:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0. \quad (2.10)$$

**Definição 2.8** A hessiana da função lagrangiana para o Problema (2.1) é a função  $\mathbf{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}^{n \times n}$  definida como:

$$\mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla_{\mathbf{xx}}^2 L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla^2 f(\mathbf{x}) - \sum_{i=1}^m \lambda_i \nabla^2 \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^p \mu_i \nabla^2 \mathbf{h}_i(\mathbf{x}). \quad (2.11)$$

**Teorema 2.9** (Condições necessárias de otimalidade de segunda ordem) Seja  $\mathbf{x}$  um mínimo local do Problema (2.1) onde é satisfeito o requisito RILR. Sejam  $\boldsymbol{\lambda}$  e  $\boldsymbol{\mu}$  os multiplicadores de Lagrange que satisfazem as condições de otimalidade de KKT. Então:

$$\mathbf{w}^T \mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{w} \geq 0, \quad \forall \mathbf{w} \in \mathbb{R}^n \quad \text{que satisfaz:} \quad (2.12)$$

$$\begin{cases} \nabla \mathbf{h}(\mathbf{x}) \mathbf{w} = 0, \\ \nabla \mathbf{g}_i(\mathbf{x}) \mathbf{w} = 0, \quad \forall i \in A(\mathbf{x}) \text{ com } \lambda_i > 0, \\ \nabla \mathbf{g}_i(\mathbf{x}) \mathbf{w} \geq 0, \quad \forall i \in A(\mathbf{x}) \text{ com } \lambda_i = 0. \end{cases} \quad (2.13)$$

Prova: Ver na referência [8, Teorema 12.5].  $\square$

**Teorema 2.10** (*Condições suficientes de otimalidade de segunda ordem*) *Seja  $\mathbf{x}$  um ponto de KKT e sejam  $\lambda$  e  $\mu$  os multiplicadores de Lagrange correspondentes. Seja  $\mathbf{H}(\mathbf{x}, \lambda, \mu)$  tal que satisfaz:*

$$\mathbf{w}^T \mathbf{H}(\mathbf{x}, \lambda, \mu) \mathbf{w} > 0, \quad \forall \mathbf{w} \neq 0 \text{ que satisfaz a Eq. (2.13)}. \quad (2.14)$$

Então  $\mathbf{x}$  é um mínimo local do Problema (2.1).

Prova: Ver na referência [8, Teorema 12.6].  $\square$

## 2.1.2 Convergência global de algoritmos

A seguinte definição foi introduzida por Zangwill [10, 11]:

**Definição 2.11** *Um algoritmo  $\mathcal{A}$  definido em um espaço  $X$  é um mapeamento ponto-conjunto, isto é, a cada ponto  $\mathbf{x} \in X$  este associa um subconjunto de  $X$ .*

Assim, um algoritmo permite gerar uma seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  da seguinte forma: Dado  $\mathbf{x}_k \in X$  o ponto  $\mathbf{x}_{k+1}$  é escolhido no conjunto  $\mathcal{A}(\mathbf{x}_k)$ , ou seja,  $\mathbf{x}_{k+1} \in \mathcal{A}(\mathbf{x}_k)$ .

**Definição 2.12** *Seja  $E \subset X$  um conjunto solução e seja  $\mathcal{A}$  um algoritmo definido em  $X$ . Uma função contínua  $F : X \rightarrow \mathbb{R}$  é chamada função de descida para  $E$  e  $\mathcal{A}$  se satisfaz:*

- a) *Se  $\mathbf{x} \notin E$  e  $\mathbf{y} \in \mathcal{A}(\mathbf{x})$ , então  $F(\mathbf{y}) < F(\mathbf{x})$ ,*
- b) *Se  $\mathbf{x} \in E$  e  $\mathbf{y} \in \mathcal{A}(\mathbf{x})$ , então  $F(\mathbf{y}) \leq F(\mathbf{x})$ .*

**Definição 2.13** *Um mapeamento ponto-conjunto  $\mathcal{A}$  de  $X$  em  $Y$  é dito fechado no ponto  $\mathbf{x} \in X$  se as hipóteses*

- a)  $\mathbf{x}_k \rightarrow \mathbf{x}$ ,  $\mathbf{x}_k \in X$ ,
- b)  $\mathbf{y}_k \rightarrow \mathbf{y}$ ,  $\mathbf{y}_k \in \mathcal{A}(\mathbf{x}_k)$ .

*implicam:*

- c)  $\mathbf{y} \in \mathcal{A}(\mathbf{x})$ .

**Teorema 2.14** (Zangwill) *Seja  $\mathcal{A}$  um algoritmo em  $X$  e suponha que dado  $\mathbf{x}_0 \in X$  a seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  é definida satisfazendo  $\mathbf{x}_{k+1} \in \mathcal{A}(\mathbf{x}_k)$ . Seja  $E$  um conjunto solução dado e suponha:*

- a) *Toda a seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  é contida em um conjunto compacto  $S \subset X$ ,*
- b) *Existe uma função de descida  $F$  para  $E$  e  $\mathcal{A}$ ,*
- c) *O mapeamento  $\mathcal{A}$  é fechado nos pontos no exterior de  $E$ .*

*Então, o limite de qualquer subseqüência convergente da seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  pertence ao conjunto solução  $E$ .*

Prova: Ver [11, pág. 187]. □

**Corolário 2.15** *Suponha que além das condições do Teorema 2.14 tem-se que o conjunto solução consiste somente do ponto  $\bar{\mathbf{x}} \in X$ . Então a seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  converge para o ponto  $\bar{\mathbf{x}}$ .*

Prova: Ver [11, pág. 188]. □

**Definição 2.16** *Seja  $\mathcal{A}$  um mapeamento de  $X$  em  $Y$  e  $\mathcal{B}$  um mapeamento de  $Y$  em  $Z$ . O mapeamento composto  $\mathcal{C} = \mathcal{B}\mathcal{A}$  é definido como:*

$$\mathcal{C}(\mathbf{x}) = \bigcup_{\mathbf{y} \in \mathcal{A}(\mathbf{x})} \mathcal{B}(\mathbf{y}).$$

**Teorema 2.17** *Seja  $\mathcal{A}$  um mapeamento ponto-conjunto de  $X$  em  $Y$  e  $\mathcal{B}$  um mapeamento de  $Y$  em  $Z$ . Suponha que  $\mathcal{A}$  é fechado em  $\mathbf{x}$  e  $\mathcal{B}$  é fechado em  $\mathcal{A}(\mathbf{x})$ . Além disso, suponha que se  $\mathbf{x}_k \rightarrow \mathbf{x}$  e  $\mathbf{y}_k \in \mathcal{A}(\mathbf{x}_k)$ , existe um ponto  $\mathbf{y}$  tal que para alguma subseqüência  $\{\mathbf{y}_{k_i}\}_{i \in \mathbb{N}}$ ,  $\mathbf{y}_{k_i} \rightarrow \mathbf{y}$ . Então o mapeamento composto  $\mathcal{C} = \mathcal{B}\mathcal{A}$  é fechado em  $\mathbf{x}$ .*

Prova: Ver [11, pág. 186]. □

**Corolário 2.18** *Seja  $\mathcal{A}$  um mapeamento ponto-conjunto de  $X$  em  $Y$  e  $\mathcal{B}$  um mapeamento de  $Y$  em  $Z$ . Suponha que  $\mathcal{A}$  é fechado em  $\mathbf{x}$  e  $\mathcal{B}$  é fechado em  $\mathcal{A}(\mathbf{x})$ . Além disso, suponha que  $Y$  é compacto. Então o mapeamento composto  $\mathcal{C} = \mathcal{B}\mathcal{A}$  é fechado em  $\mathbf{x}$ .* □

**Teorema 2.19** *Seja  $\mathcal{A}$  um algoritmo em  $X$  fechado fora do conjunto solução  $E$ . Seja  $F$  uma função de descida para  $E$  e  $\mathcal{A}$ . Suponha que a seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  é gerada satisfazendo  $\mathbf{x}_{k+1} \in \mathcal{A}(\mathbf{x}_k)$  para  $k$  no conjunto de índices infinito  $\mathbb{K} \subset \mathbb{N}$  e que  $F(\mathbf{x}_{k+1}) \leq$*

$F(\mathbf{x}_k)$  para todo  $k \in \mathbb{N}$ . Além disso, suponha que o conjunto  $S = \{\mathbf{x} \mid F(\mathbf{x}) \leq F(\mathbf{x}_0)\}$  é compacto. Então, o limite de qualquer subsequência convergente de  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  é uma solução.

Prova: Ver o “Spacer Step Theorem” [11, pág. 231]. □

### 2.1.3 Algoritmo de Newton

O método de Newton-Raphson é um método iterativo que resolve um sistema de  $n$  equações não lineares com  $n$  incógnitas resolvendo uma seqüência de sistemas de equações lineares. Seja o seguinte sistema de equações não lineares:

$$\mathbf{F}(\mathbf{x}) = 0, \quad (2.15)$$

onde a função  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  é uma função contínua com derivadas contínuas. Seja  $\mathbf{x}_0$  um ponto tal que  $\mathbf{F}(\mathbf{x}_0) \approx 0$ . Uma melhor aproximação pode ser obtida linearizando a função  $\mathbf{F}$  na vizinhança do ponto  $\mathbf{x}_0$ . Tal aproximação é:

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0) + \nabla \mathbf{F}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

O ponto  $\mathbf{x}$  que anula a aproximação linear da função  $\mathbf{F}$  é o ponto que verifica:

$$\mathbf{F}(\mathbf{x}_0) + \nabla \mathbf{F}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = 0.$$

Então, chamando  $\mathbf{d}_x = \mathbf{x} - \mathbf{x}_0$ , o ponto  $\mathbf{x}$  pode ser calculado na forma:

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{d}_x, \quad (2.16)$$

onde o vetor  $\mathbf{d}_x$  pode ser obtido resolvendo o seguinte sistema linear:

$$\nabla \mathbf{F}(\mathbf{x}_0) \mathbf{d}_x = -\mathbf{F}(\mathbf{x}_0). \quad (2.17)$$

Como foi utilizada uma aproximação linear da função  $\mathbf{F}$ , o ponto  $\mathbf{x}$  definido pela Eq. (2.16) não é, em geral, solução da Eq. (2.15). Neste caso deve ser feita uma nova aproximação linear, agora no ponto  $\mathbf{x}$ , e continuar com o mesmo procedimento até satisfazer algum critério de convergência.

### Algoritmo de Newton-Raphson

Dados:  $\mathbf{x}_0$  tal que  $F(\mathbf{x}_0) \approx 0$ .

**Passo 1:** Teste de convergência

**Passo 2:** Cálculo da direção de Newton solução do sistema:

$$\nabla F(\mathbf{x}_k) \mathbf{d}_x = -F(\mathbf{x}_k).$$

**Passo 3:** Atualização

Defina  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_x$ .

Retorne ao Passo 1.

Considere uma solução  $\mathbf{x}^*$  da Eq. (2.15). Então, é válido do seguinte teorema:

**Teorema 2.20** *Se  $\nabla F$  é contínua e satisfaz a condição de Lipschitz em um entorno de  $\mathbf{x}^*$ , e  $\nabla F(\mathbf{x}^*)$  não é singular, então para  $\mathbf{x}_0$  suficientemente próximo a  $\mathbf{x}^*$  a iteração de Newton está bem definida ( $\nabla F(\mathbf{x}_k)$  não é singular para todo  $k \geq 0$ ) e converge ao ponto  $\mathbf{x}^*$ . Além disso, a velocidade de convergência da seqüência é quadrática, ou seja, existe  $K$  tal que:*

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq K \|\mathbf{x}_k - \mathbf{x}^*\|^2 \quad \forall k \text{ suficientemente grande.}$$

Prova: Ver nas referências [12, 13].

□

### 2.1.4 Algoritmo FDIPA

O algoritmo FDIPA [14] é um método iterativo que serve para resolver o Problema (2.1). Este algoritmo requer um ponto inicial  $\mathbf{x}_0$  pertencente ao interior da região  $\Delta$  definida como:

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0, \mathbf{h}(\mathbf{x}) \geq 0\}.$$

A partir do ponto inicial o algoritmo FDIPA gera uma seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  totalmente contida no interior de  $\Delta$  e que converge a um ponto de KKT do Problema (2.1). A seqüência gerada pelo algoritmo reduz monotonamente o valor da *função potencial*:

$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^p c_i |\mathbf{h}_i(\mathbf{x})|,$$

onde o vetor  $\mathbf{c} \in \mathbb{R}^p$  é definido pelo algoritmo. Pode ser mostrado que se as componentes  $c_i$  do vetor  $\mathbf{c}$  são suficientemente grandes, então a função  $\phi_{\mathbf{c}}$  é uma função de

penalidade exata para o Problema (2.1), ver [14]. Em outras palavras, existe um vetor  $\mathbf{c}$  tal que os mínimos da função  $\phi_{\mathbf{c}}$  sujeitos às restrições de desigualdade são soluções do Problema (2.1).

A iteração do algoritmo FDIPA é dada pela regra  $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{d}_x$  onde  $\mathbf{x}_k$  é o ponto da iteração atual,  $\mathbf{d}_x$  é chamada de *direção de busca* e  $t_k$  é o *passo* da iteração  $k$  definido por um procedimento chamado *busca linear*. A definição da direção de busca do algoritmo FDIPA é inspirada na aplicação do método de Newton no sistema não linear de equações dado pelas Eqs. (2.4)-(2.6). A iteração de Newton para este sistema de equações é  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_x$  onde  $\mathbf{d}_x$  é obtida resolvendo o sistema linear:

$$\begin{pmatrix} \mathbf{H}(\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k) & -\nabla \mathbf{g}(\mathbf{x}_k)^T & -\nabla \mathbf{h}(\mathbf{x}_k)^T \\ \boldsymbol{\Lambda}_k \nabla \mathbf{g}(\mathbf{x}_k) & \mathbf{G}(\mathbf{x}_k) & 0 \\ \nabla \mathbf{h}(\mathbf{x}_k) & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x \\ \mathbf{d}_\lambda \\ \mathbf{d}_\mu \end{pmatrix} = \begin{pmatrix} -\nabla_x L(\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k)^T \\ -\mathbf{G}(\mathbf{x}_k) \boldsymbol{\lambda}_k \\ -\mathbf{h}(\mathbf{x}_k) \end{pmatrix},$$

onde  $\mathbf{H}$  é a matriz hessiana definida na Eq. (2.11),  $L$  é definida na Eq. (2.9),  $\boldsymbol{\Lambda}_k = \text{diag}(\boldsymbol{\lambda}_k)$  e  $\mathbf{G}(\mathbf{x}_k) = \text{diag}(\mathbf{g}(\mathbf{x}_k))$ . Definindo as variáveis auxiliares  $\bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}_k + \mathbf{d}_\lambda$  e  $\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_k + \mathbf{d}_\mu$ , o sistema anterior pode ser escrito como:

$$\begin{pmatrix} \mathbf{H}(\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k) & -\nabla \mathbf{g}(\mathbf{x}_k)^T & -\nabla \mathbf{h}(\mathbf{x}_k)^T \\ \boldsymbol{\Lambda}_k \nabla \mathbf{g}(\mathbf{x}_k) & \mathbf{G}(\mathbf{x}_k) & 0 \\ \nabla \mathbf{h}(\mathbf{x}_k) & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x \\ \bar{\boldsymbol{\lambda}} \\ \bar{\boldsymbol{\mu}} \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x}_k)^T \\ 0 \\ \mathbf{h}(\mathbf{x}_k) \end{pmatrix}. \quad (2.18)$$

Na referência [14] é mostrado que a direção de Newton pode não ser uniformemente viável para a região  $\Delta$  nem de descida para a função potencial  $\phi_{\mathbf{c}}$ . Por esta razão a direção de busca do algoritmo FDIPA é diferente da dada pela Eq. (2.18). O pseudo-código do algoritmo FDIPA é o seguinte:

### Algoritmo FDIPA

Parâmetros:  $\xi \in (0, 1)$ ,  $\eta \in (0, 1)$ ,  $\varphi > 0$  e  $\nu \in (0, 1)$ .

Dados:  $\mathbf{x}_0 \in \Delta^0$ ,  $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$  positivo,  $\mathbf{B}_0 \in \mathbb{R}^{n \times n}$  positiva definida,  $\boldsymbol{\omega}^D \in \mathbb{R}^m$  positivo,  $\boldsymbol{\omega}^E \in \mathbb{R}^p$  positivo e  $\mathbf{c}_0 \in \mathbb{R}^p$  positivo.

**Passo 1:** Teste de convergência

**Passo 2:** Cálculo da direção de busca

2.1 Calcule  $(\mathbf{d}_x^\alpha, \boldsymbol{\lambda}^\alpha, \boldsymbol{\mu}^\alpha)$  e  $(\mathbf{d}_x^\beta, \boldsymbol{\lambda}^\beta, \boldsymbol{\mu}^\beta)$  como a solução do sistema:

$$\begin{pmatrix} \mathbf{B}_k & -\nabla \mathbf{g}(\mathbf{x}_k)^T & -\nabla \mathbf{h}(\mathbf{x}_k)^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}(\mathbf{x}_k) & -\mathbf{G}(\mathbf{x}_k) & 0 \\ -\nabla \mathbf{h}(\mathbf{x}_k) & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^\alpha & \mathbf{d}_x^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x}_k)^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k \boldsymbol{\omega}^D \\ \mathbf{h}(\mathbf{x}_k) & -\boldsymbol{\omega}^E \end{pmatrix}.$$

2.2 Se  $\mathbf{d}_x^\alpha = 0$  pare.

2.3 Se  $(\mathbf{c}_k)_i < -1.2\boldsymbol{\mu}_i^\alpha$  faça  $(\mathbf{c}_k)_i = -2\boldsymbol{\mu}_i^\alpha$ ,  $1 \leq i \leq p$ .

2.4 Se  $\nabla \phi_{\mathbf{c}_k} \mathbf{d}_x^\beta > 0$ , faça:

$$\rho = \min(\varphi \|\mathbf{d}_x^\alpha\|^2, (\xi - 1) \nabla \phi_{\mathbf{c}_k} \mathbf{d}_x^\alpha / \nabla \phi_{\mathbf{c}_k} \mathbf{d}_x^\beta).$$

Se não, faça:

$$\rho = \varphi \|\mathbf{d}_x^\alpha\|^2.$$

2.5 Calcule a direção de busca  $\mathbf{d}_x$  e o vetor  $\bar{\boldsymbol{\lambda}}$ :

$$\mathbf{d}_x = \mathbf{d}_x^\alpha + \rho \mathbf{d}_x^\beta, \quad \bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^\alpha + \rho \boldsymbol{\lambda}^\beta. \quad (2.19)$$

### Passo 3: Busca linear

Calcule  $t_k$  como o primeiro número da seqüência  $\{1, \nu, \nu^2, \nu^3 \dots\}$  que satisfaz:

$$\begin{aligned} \phi_{\mathbf{c}_k}(\mathbf{x}_k + t_k \mathbf{d}_x) &\leq \phi_{\mathbf{c}_k}(\mathbf{x}_k) + t_k \eta \nabla \phi_{\mathbf{c}_k} \mathbf{d}_x, \\ \mathbf{h}(\mathbf{x}_k + t_k \mathbf{d}_x) &\geq 0, \\ \mathbf{g}_i(\mathbf{x}_k + t_k \mathbf{d}_x) &> 0, \quad \text{se } \bar{\boldsymbol{\lambda}}_i \geq 0 \text{ ou} \\ \mathbf{g}_i(\mathbf{x}_k + t_k \mathbf{d}_x) &\geq \mathbf{g}_i(\mathbf{x}_k), \quad \text{caso contrário.} \end{aligned}$$

### Passo 4: Atualização

4.1 Faça:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{d}_x,$$

$$\mathbf{c}_{k+1} = \mathbf{c}_k,$$

e defina novos valores:  $\boldsymbol{\omega}^D > 0$ ,  $\boldsymbol{\omega}^E > 0$ ,  $\boldsymbol{\lambda}_{k+1} > 0$ , e  $\mathbf{B}_{k+1}$  positiva definida.

4.2 Retorne ao Passo 1.

O algoritmo FDIPA tem convergência global a pontos de KKT se algumas condições de regularidade forem satisfeitas pelas funções que definem o problema de otimização. Ver na Seção 3.1.

## 2.2 Método dos elementos de contorno

O Método dos Elementos de Contorno (MEC) é um método numérico amplamente utilizado como ferramenta de análise para diversos problemas da engenharia [15–18]. Ele é baseado em uma fórmula integral, equivalente às equações de equilíbrio do problema, que em elasticidade é conhecida como identidade de Somigliana. Uma característica importante do método é que somente requer de uma discretização do contorno do domínio. Esta característica faz do MEC uma ferramenta ideal em problemas de domínio infinito como por exemplo em mecânica da fratura, acústica, análise de túneis, interação solo-estrutura e em problemas de dinâmica em sismologia. Por outro lado, os resultados obtidos com o método são mais precisos que os obtidos utilizando outras ferramentas de análise, como o Método dos Elementos Finitos (MEF) ou o Método das Diferenças Finitas (MDF), isto faz do MEC uma boa alternativa por exemplo nos problemas de concentração de tensões em elasticidade [15, 17].

As características principais do MEC são [15–18]:

- Somente é necessária a discretização do contorno do problema. Isto conduz a problemas com um número de variáveis muito menor quando comparado com o MEF ou o MDF. Esta característica é acentuada em aplicações tridimensionais ou com domínio infinito. Por outro lado, a menor dificuldade de discretização o faz ideal para aplicações como otimização de forma, onde a análise de muitas geometrias diferentes pode ser necessária antes de encontrar a forma ótima.
- As matrizes dos sistemas de equações obtidos utilizando o MEC são cheias e não simétricas. Estes sistemas lineares são, em geral, de mais difícil solução.
- O tratamento das integrais singulares que aparecem no método o tornam mais difícil de implementar quando comparado com o MEF. No entanto, desde o ponto de vista do usuário, talvez o MEC seja mais fácil de utilizar já que não é necessária a discretização do interior do domínio.
- Alguns problemas não possuem uma solução fundamental que permita a aplicação do MEC. Problemas com não homogeneidade do domínio ou não lineares são mais difíceis de resolver utilizando o MEC.

- No caso de otimização estrutural, o MEC permite obter diretamente o valor das variáveis de estado no contorno do problema, o que é, em geral, suficiente tanto para calcular as funções de performance do problema de otimização quanto para realizar a análise de sensibilidade.

Considere o problema de elasticidade linear. O mesmo consiste em determinar as funções  $\mathbf{u}$ ,  $\boldsymbol{\epsilon}$  e  $\boldsymbol{\sigma}$  que definem o deslocamento, as deformações e as tensões em um corpo elástico que ocupa um domínio  $\Omega$  do espaço euclidiano bidimensional, no caso de estado plano de tensões ou deformações, ou tridimensional, no caso geral [15–18]. O domínio  $\Omega$  é um aberto limitado por uma superfície suficientemente regular para que seja aplicável o teorema da divergência de Gauss [19, 20]. As condições de contorno são dadas pelas funções  $\bar{\mathbf{u}}$  e  $\bar{\mathbf{p}}$  definidas em regiões complementares do contorno de  $\Omega$  (ver Figura 2.1).

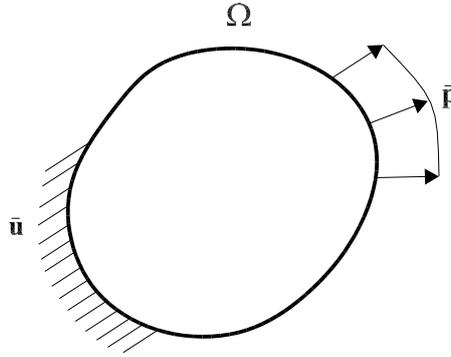


Figura 2.1: Problema de elasticidade linear.

O tensor de deformações  $\boldsymbol{\epsilon}$  pode ser expresso em termos das derivadas das componentes do deslocamento na forma:

$$\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \quad (2.20)$$

A equação de equilíbrio no interior do domínio é:

$$\frac{\partial \sigma_{ij}}{\partial x_i} + \mathbf{b}_i = 0, \quad (2.21)$$

onde as componentes do tensor de tensões são representadas por  $\sigma_{ij}$  e as forças de volume por  $\mathbf{b}_i$ .

A condição de equilíbrio no contorno  $\Gamma$  do corpo é dada por:

$$\mathbf{p}_i - \sigma_{ij} \mathbf{n}_j = 0, \quad (2.22)$$

a qual expressa que o vetor força de superfície, de componentes  $\mathbf{p}_i$ , é igual ao obtido pela aplicação do tensor de tensões ao vetor normal exterior a  $\Gamma$ , de componentes  $\mathbf{n}_j$ .

Para um material elástico, linear e isótropo, onde não existem mudanças de temperatura, a lei de Hooke que relaciona tensões e deformações é:

$$\sigma_{ij} = \mathbb{C}_{ijkl}\epsilon_{kl}, \quad (2.23)$$

onde o tensor  $\mathbb{C}_{ijkl}$  pode ser escrito como:

$$\mathbb{C}_{ijkl} = \frac{2G\nu}{1-2\nu}\delta_{ij}\delta_{kl} + G(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), \quad (2.24)$$

onde  $G$  é o módulo de cisalhamento do material,  $\nu$  é o coeficiente de Poisson e  $\delta_{ij}$  é o símbolo delta de Kronecker.

A Eq. (2.24) é válida para o caso tridimensional e para problemas de estado plano de deformações. Para o estado plano de tensões a mesma expressão pode ser empregada substituindo  $\nu$  por  $\bar{\nu} = \nu/(1+\nu)$ .

A substituição da Eq. (2.20) na Eq. (2.23) fornece o tensor de tensões em termos das derivadas dos deslocamentos. Esta nova expressão pode ser substituída na Eq. (2.21) para a obtenção da equação de Navier, que junto com as condições de contorno define o problema de elasticidade linear:

$$\begin{cases} G \frac{\partial^2 \mathbf{u}_j}{\partial x_k \partial x_k} + \frac{G}{1-2\nu} \frac{\partial^2 \mathbf{u}_k}{\partial x_k \partial x_j} + \mathbf{b}_j = 0, \\ \mathbf{u}_i = \bar{\mathbf{u}}_i \quad \text{em } \Gamma_{\mathbf{u}}, \\ \mathbf{p}_i = \bar{\mathbf{p}}_i \quad \text{em } \Gamma_{\mathbf{p}}, \end{cases} \quad (2.25)$$

onde  $\bar{\mathbf{u}}_i$  e  $\bar{\mathbf{p}}_i$  são os valores prescritos e o contorno total do corpo é dado por  $\Gamma = \Gamma_{\mathbf{u}} \cup \Gamma_{\mathbf{p}}$ .

Uma equação integral, para o problema da Eqs. (2.25), pode ser agora obtida através da formulação de resíduos ponderados ou pelo princípio da reciprocidade baseado na simetria do tensor elástico  $\mathbb{C}_{ijkl}$  [15–18]. Esta equação é conhecida como *identidade de Somigliana* [15–18] para os deslocamentos e pode ser escrita na forma:

$$\mathbf{u}_i(\xi) = \int_{\Gamma} \mathbf{u}_{ij}^*(\xi, x) \mathbf{p}_j(x) d\Gamma - \int_{\Gamma} \mathbf{p}_{ij}^*(\xi, x) \mathbf{u}_j(x) d\Gamma, \quad (2.26)$$

onde  $\mathbf{b}_j = 0$  foi assumido e as funções estrela representam a *solução fundamental* para a Eq. (2.25) na qual as forças de volume correspondem a uma carga unitária pontual aplicada em  $\xi$  na direção “ $i$ ”. Dito de outra forma, a componente  $\mathbf{b}_i^*$  da carga associada à solução fundamental é igual a função delta de Dirac. As funções  $\mathbf{u}_{ij}^*$  e  $\mathbf{p}_{ij}^*$  representam,

respectivamente, as componentes na direção “ $j$ ” dos deslocamentos e forças de superfície da solução do problema fundamental.

Para estado plano de deformações, a solução fundamental corresponde a uma carga unitária pontual aplicada dentro de um plano infinito (solução de Kelvin). As componentes de deslocamento e forças de superfície fundamentais são dados por:

$$\mathbf{u}_{ij}^*(\xi, x) = \frac{-1}{8\pi(1-\nu)G} \{ (3-4\nu) \log(r) \delta_{ij} - r_{,i} r_{,j} \}, \quad (2.27)$$

$$\mathbf{p}_{ij}^*(\xi, x) = \frac{-1}{4\pi(1-\nu)r} \left\{ [(1-\nu) \delta_{ij} + 2r_{,i} r_{,j}] \frac{\partial r}{\partial \mathbf{n}} - (1-\nu)(r_{,i} \mathbf{n}_j - r_{,j} \mathbf{n}_i) \right\}, \quad (2.28)$$

onde  $r = r(\xi, x)$  representa a distância entre o ponto fonte  $\xi$  e o ponto campo  $x$ . As derivadas são sempre calculadas em relação a  $x$ , isto é:

$$r = \|x - \xi\|, \quad (2.29)$$

$$r_{,i} = \frac{\partial r}{\partial x_i} = \frac{x_i - \xi_i}{r}. \quad (2.30)$$

Assim, a identidade de Somigliana fornece uma representação contínua dos deslocamentos que pode ser avaliada em qualquer ponto  $\xi$  no interior de  $\Omega$ .

As componentes das tensões em pontos internos podem ser determinadas derivando a Eq. (2.26) com relação às coordenadas de  $\xi$  e substituindo nas Eqs. (2.20) e (2.23). O resultado é dado por:

$$\sigma_{ij}(\xi) = \int_{\Gamma} \mathbf{u}_{ijk}^*(\xi, x) \mathbf{p}_k(x) d\Gamma - \int_{\Gamma} \mathbf{p}_{ijk}^*(\xi, x) \mathbf{u}_k(x) d\Gamma, \quad (2.31)$$

na qual os novos termos de tensão fundamental são:

$$\mathbf{u}_{ijk}^*(\xi, x) = \frac{1}{4\pi(1-\nu)r} \{ (1-2\nu)(r_{,i} \delta_{jk} + r_{,j} \delta_{ik} - r_{,k} \delta_{ij}) \}, \quad (2.32)$$

$$\begin{aligned} \mathbf{p}_{ijk}^*(\xi, x) = \frac{G}{2\pi(1-\nu)r^2} \left\{ 2 \frac{\partial r}{\partial \mathbf{n}} [(1-2\nu) \delta_{ij} r_{,k} + \nu(r_{,i} \delta_{jk} + r_{,j} \delta_{ik}) - 4r_{,i} r_{,j} r_{,k}] \right. \\ \left. + 2\nu(\mathbf{n}_i r_{,j} r_{,k} + \mathbf{n}_j r_{,i} r_{,k}) + (1-2\nu)(2\mathbf{n}_k r_{,i} r_{,j} + \mathbf{n}_j \delta_{ik} + \mathbf{n}_i \delta_{jk}) \right. \\ \left. - (1-4\nu) \mathbf{n}_k \delta_{ij} \right\}. \end{aligned} \quad (2.33)$$

Para obter uma equação integral envolvendo variáveis apenas do contorno do corpo, deve-se efetuar o limite da Eq. (2.26) com o ponto  $\xi$  tendendo para um ponto em  $\Gamma$ . Esse limite deve ser calculado com cuidado visto que a integral de contorno torna-se singular em  $\xi$ . Ver, por exemplo, nas referências [15–18]. A equação resultante é:

$$\mathbf{c}_{ij}(\xi) \mathbf{u}_j(\xi) = \int_{\Gamma} \mathbf{u}_{ij}^*(\xi, x) \mathbf{p}_j(x) d\Gamma - \int_{\Gamma} \mathbf{p}_{ij}^*(\xi, x) \mathbf{u}_j(x) d\Gamma, \quad (2.34)$$

onde o coeficiente  $c_{ij}$  depende da geometria de  $\Gamma$  no ponto  $\xi$  e a segunda integral deve ser calculada no sentido do valor principal de Cauchy. Vale a pena mencionar que este coeficiente, junto com o valor principal associado, pode ser calculado indiretamente se for aplicada a Eq. (2.34) para representar duas translações independentes de corpo rígido.

Em forma matricial, a Eq. (2.34) fica:

$$\mathbf{c}(\xi)\mathbf{u}(\xi) = \int_{\Gamma} \mathbf{u}^*(\xi, x)\mathbf{p}(x) d\Gamma - \int_{\Gamma} \mathbf{p}^*(\xi, x)\mathbf{u}(x) d\Gamma. \quad (2.35)$$

A discretização do MEC consiste em aproximar o contorno  $\Gamma$  por  $n_e$  elementos  $\Gamma^{(j)}$  de forma tal que  $\Gamma \approx \bigcup_{i=1}^{n_e} \Gamma^{(j)}$ . No que segue  $\xi^i$ ,  $\mathbf{u}^i$  e  $\mathbf{p}^i$  denotarão a posição, deslocamento e força de superfície do nó  $i$ .  $\xi^{(j)}$ ,  $\mathbf{u}^{(j)}$  e  $\mathbf{p}^{(j)}$  denotarão as posições, deslocamentos e forças de superfície nodais do elemento  $\Gamma^{(j)}$ . Para elementos lineares no caso bidimensional, a representação paramétrica de  $\Gamma^{(j)}$  é:

$$\xi(\eta) = \Phi(\eta)\xi^{(j)}, \quad (2.36)$$

onde a matriz  $\Phi(\eta)$  é:

$$\Phi(\eta) = \begin{pmatrix} \phi_1(\eta) & 0 & \phi_2(\eta) & 0 \\ 0 & \phi_1(\eta) & 0 & \phi_2(\eta) \end{pmatrix}, \quad (2.37)$$

com as funções de interpolação  $\phi_1$  e  $\phi_2$  definidas como:

$$\phi_1(\eta) = (1/2)(1 - \eta), \quad \phi_2(\eta) = (1/2)(1 + \eta). \quad (2.38)$$

Utilizando a mesma interpolação para deslocamentos e forças de superfície tem-se:

$$\mathbf{u}(\eta) = \Phi(\eta)\mathbf{u}^{(i)}, \quad \mathbf{p}(\eta) = \Phi(\eta)\mathbf{p}^{(i)}. \quad (2.39)$$

O método de Colocação [15–18] impõe que a Eq. (2.35) seja satisfeita nos nós do problema. Assim, para o nó  $\xi^i$ , tem-se:

$$\mathbf{c}(\xi^i)\mathbf{u}^i = \sum_{j=1}^{n_e} \left\{ \int_{\Gamma^{(j)}} \mathbf{u}^*(\xi^i, x)\Phi\mathbf{p}^{(j)} d\Gamma \right\} - \sum_{j=1}^{n_e} \left\{ \int_{\Gamma^{(j)}} \mathbf{p}^*(\xi^i, x)\Phi\mathbf{u}^{(j)} d\Gamma \right\}. \quad (2.40)$$

Definem-se as matrizes:

$$\hat{\mathbf{h}}_{ij} = \int_{\Gamma^{(j)}} \mathbf{p}^*(\xi^i, x)\Phi d\Gamma, \quad \mathbf{g}_{ij} = \int_{\Gamma^{(j)}} \mathbf{u}^*(\xi^i, x)\Phi d\Gamma. \quad (2.41)$$

Com estas definições, a Eq. (2.40) pode ser escrita como:

$$\mathbf{c}(\xi^i)\mathbf{u}^i + \sum_{j=1}^{n_e} \hat{\mathbf{h}}_{ij}\mathbf{u}^{(j)} - \sum_{j=1}^{n_e} \mathbf{g}_{ij}\mathbf{p}^{(j)} = 0. \quad (2.42)$$

O primeiro termo pode ser adicionado no primeiro somatório, neste caso fica:

$$\sum_{j=1}^{n_e} \mathbf{h}_{ij} \mathbf{u}^{(j)} - \sum_{j=1}^{n_e} \mathbf{g}_{ij} \mathbf{p}^{(j)} = 0, \quad (2.43)$$

onde as matrizes  $\mathbf{h}_{ij}$  foram definidas de forma implícita pela Eq. (2.43).

## 2.2.1 Cálculo das matrizes elementares

As integrais não singulares da Eq. (2.41) podem ser calculadas aproximadamente utilizando o método de integração numérica por pontos de Gauss. Definem-se então as matrizes:

$$\mathbf{h}_{ij} = \sum_{k=1}^{n_g} \omega_k J \mathbf{p}^*(\xi^i, x(\eta_k)) \Phi(\eta_k), \quad \mathbf{g}_{ij} = \sum_{k=1}^{n_g} \omega_k J \mathbf{u}^*(\xi^i, x(\eta_k)) \Phi(\eta_k), \quad (2.44)$$

onde  $n_g$  é o número de pontos de Gauss,  $\eta_k$  e  $\omega_k$  são as abscissas e pesos de Gauss respectivamente, e  $J$  é o jacobiano da transformação de variável (constante para elementos lineares). Para o cálculo analítico destas integrais ver nas referências [21, 22].

As integrais singulares que existem no sentido de Riemann são aqui calculadas de forma analítica. Considerando que os nós do elemento  $\Gamma^{(j)}$  são  $\xi^{(j,1)}$  e  $\xi^{(j,2)}$ , então são válidas as seguintes fórmulas:

1) Caso  $\xi^i = \xi^{(j,1)}$

$$\mathbf{h}_{ij} = C_4 \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad (2.45)$$

$$\mathbf{g}_{ij} = C_2 L \begin{pmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \end{pmatrix} - C_3 L^{-1} \begin{pmatrix} l_1^2 & l_1 l_2 & l_1^2 & l_1 l_2 \\ l_1 l_2 & l_2^2 & l_1 l_2 & l_2^2 \end{pmatrix}, \quad (2.46)$$

2) Caso  $\xi^i = \xi^{(j,2)}$

$$\mathbf{h}_{ij} = C_4 \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (2.47)$$

$$\mathbf{g}_{ij} = C_2 L \begin{pmatrix} a_2 & 0 & a_1 & 0 \\ 0 & a_2 & 0 & a_1 \end{pmatrix} - C_3 L^{-1} \begin{pmatrix} l_1^2 & l_1 l_2 & l_1^2 & l_1 l_2 \\ l_1 l_2 & l_2^2 & l_1 l_2 & l_2^2 \end{pmatrix}, \quad (2.48)$$

onde, nas equações acima  $C_1, C_2, C_3, C_4, l_1, l_2, L, a_1$  e  $a_2$  são:

$$\begin{aligned} C_1 &= \frac{-1}{4(1-\nu)\pi}, & C_3 &= C_1/G, \\ C_2 &= (3 - 4\nu)C_1/(2G), & C_4 &= (1 - 2\nu)C_1, \\ l_1 &= \xi_1^{(j,2)} - \xi_1^{(j,1)}, & a_1 &= \log(L) - 1.5, \\ l_2 &= \xi_2^{(j,2)} - \xi_2^{(j,1)}, & a_2 &= \log(L) - 0.5, \\ L &= \sqrt{l_1^2 + l_2^2}. \end{aligned}$$

As integrais que devem ser calculadas no sentido do valor principal de Cauchy podem ser determinadas de forma indireta utilizando a Eq. (2.49). Visto que os deslocamentos rígidos verificam exatamente a equação de Navier, eles devem ser solução da Eq. (2.42). Uma equação que permite calcular as matrizes  $\mathbf{c}(\xi^i)$  e as integrais de Cauchy da Eq. (2.40) pode ser obtida considerando deslocamentos rígidos horizontais e verticais. Neste caso  $\mathbf{u}^i$  é constante independente de  $i$  e  $\mathbf{p}^i = 0$  obtendo:

$$\sum_{j=1}^{n_e} (\mathbf{h}_{ij}^1 + \mathbf{h}_{ij}^2) = 0, \quad (2.49)$$

onde  $\mathbf{h}_{ij}^1$  e  $\mathbf{h}_{ij}^2$  referem-se às duas primeiras e duas últimas colunas de  $\hat{\mathbf{h}}_{ij}$ . Utilizando esta equação, os coeficientes onde aparecem as matrizes  $\mathbf{c}$  e os definidos por integrais de valor principal de Cauchy podem ser calculados a partir dos obtidos pelas Eqs. (2.44)-(2.48).

## 2.2.2 Equação para pontos no interior dos elementos

Em alguns casos torna-se necessária a aplicação da Eq. (2.35) no interior de algum dos elementos, por exemplo, quando a estratégia de nós duplos é utilizada e as condições de contorno fazem necessária a introdução de uma equação adicional. A estratégia de nós duplos consiste em definir nós diferentes, portanto associados a variáveis diferentes, mas que estão localizados em um mesmo ponto do espaço euclidiano. Esta estratégia é útil para modelar casos onde existe uma descontinuidade da normal ou da função que define a condição de contorno. Ver, por exemplo, em [15].

Se a Eq. (2.35) é aplicada em um ponto do elemento  $\Gamma^{(i)}$ , por exemplo, no ponto  $z^{(i)} = \Phi(\eta_0)\xi^{(i)}$ , a equação integral fica:

$$\frac{1}{2}\Phi(\eta_0)\mathbf{u}^{(i)} + \sum_{j=1}^{n_e} \hat{\mathbf{h}}_{ij}\mathbf{u}^{(j)} - \sum_{j=1}^{n_e} \mathbf{g}_{ij}\mathbf{p}^{(j)} = 0, \quad (2.50)$$

onde, neste caso, as matrizes  $\hat{\mathbf{h}}_{ij}$  e  $\mathbf{g}_{ij}$  são definidas como:

$$\hat{\mathbf{h}}_{ij} = \int_{\Gamma^{(j)}} \mathbf{p}^*(z^{(i)}, x) \Phi \, d\Gamma, \quad \hat{\mathbf{g}}_{ij} = \int_{\Gamma^{(j)}} \mathbf{u}^*(z^{(i)}, x) \Phi \, d\Gamma. \quad (2.51)$$

Neste caso também é possível adicionar o primeiro termo da Eq. (2.43) no primeiro somatório. Assim:

$$\sum_{j=1}^{n_e} \mathbf{h}_{ij} \mathbf{u}^{(j)} - \sum_{j=1}^{n_e} \mathbf{g}_{ij} \mathbf{p}^{(j)} = 0, \quad (2.52)$$

onde, novamente, a matriz  $\mathbf{h}_{ij}$  foi definida de forma implícita pela Eq. (2.52).

As integrais não singulares da Eq. (2.51) podem ser calculadas utilizando o método de Gauss (Ver [21, 22] para o cálculo analítico destas integrais). As singulares podem ser calculadas utilizando as seguintes expressões:

$$\mathbf{h}_{ij} = C_4 \begin{pmatrix} 0 & -c_1 & 0 & c_2 \\ c_1 & 0 & -c_2 & 0 \end{pmatrix} + \begin{pmatrix} c_3 & 0 & c_4 & 0 \\ 0 & c_3 & 0 & c_4 \end{pmatrix}, \quad (2.53)$$

$$\mathbf{g}_{ij} = C_2 L \begin{pmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \end{pmatrix} - C_3 L^{-1} \begin{pmatrix} l_1^2 & l_1 l_2 & l_1^2 & l_1 l_2 \\ l_1 l_2 & l_2^2 & l_1 l_2 & l_2^2 \end{pmatrix}, \quad (2.54)$$

onde:

$$\begin{aligned} c_1 &= 1 + b \log(a/b), & c_3 &= 0.5b, \\ c_2 &= 1 + a \log(b/a), & c_4 &= 0.5a, \\ a_1 &= \log(L) - 0.5 - b - b^2 \log(a/b) + \log(a), & a &= 1 + \eta_0, \\ a_2 &= \log(L) - 0.5 - a - a^2 \log(b/a) + \log(b), & b &= 1 - \eta_0. \end{aligned}$$

A definição da Eq. (2.53) inclui o termo  $(1/2)\Phi(\eta_0)\mathbf{u}^{(i)}$  da Eq. (2.43).

### 2.2.3 Sistema de equações do MEC

A partir da Eq. (2.43) aplicada em todos os nós da discretização e, ocasionalmente, da Eq. (2.52) para pontos no interior dos elementos, obtém-se a equação:

$$\mathbf{H}\mathbf{u} - \mathbf{G}\mathbf{p} = 0, \quad (2.55)$$

onde, neste caso os vetores  $\mathbf{u}$  e  $\mathbf{p}$  representam os valores nodais em todos os nós da discretização. Cada equação (que corresponde a uma linha de  $\mathbf{H}$  e  $\mathbf{G}$ ) corresponde a uma equação integral aplicada em um ponto particular.

Ao aplicar as condições de contorno e separar as variáveis conhecidas das desconhecidas tem-se o sistema:

$$\mathbf{Ax} = \mathbf{b}, \quad (2.56)$$

onde as colunas da matriz  $\mathbf{A}$  são as colunas das matrizes  $\mathbf{H}$  e  $\mathbf{G}$  que ficavam multiplicadas por deslocamentos ou forças de superfície desconhecidas, agora agrupados no vetor  $\mathbf{x}$ . O vetor  $\mathbf{b}$  resulta de somar as outras colunas das matrizes  $\mathbf{H}$  e  $\mathbf{G}$  multiplicadas pelo valor do deslocamento ou da força de superfície dados pela condição de contorno.

O MEC resolve este sistema linear, por exemplo utilizando o método de eliminação de Gauss-Jordan ou a decomposição LU [23, 24], para obter todos os deslocamentos e forças de superfície no contorno. Valores do deslocamento ou do tensor de tensões no interior do domínio podem ser calculados a partir dos valores no contorno utilizando as Eqs. (2.26) e (2.31).

## 2.3 Formulação SAND

O método clássico de otimização estrutural começa por estabelecer uma parametrização, ou seja, uma relação entre o conjunto das estruturas admissíveis e um conjunto de parâmetros reais, chamados *variáveis de projeto* [25–27]. Estas variáveis reais definem propriedades geométricas específicas da estrutura em questão, ou outras propriedades mecânicas como por exemplo as propriedades dos materiais. Áreas de barras, espessuras de placas, coordenadas de pontos ou parâmetros de interpolação que definem curvas e superfícies são usualmente empregados como variáveis de projeto em otimização estrutural.

Para um conjunto específico de valores reais das variáveis de projeto tem-se uma estrutura específica que pode ser analisada utilizando métodos numéricos de análise estrutural como o MEF ou o MEC. A análise estrutural permite definir uma ou várias funções de performance que servem como critério de comparação entre as diferentes estruturas. A massa da estrutura, a complacência, tensões, deformações, deslocamentos, frequências naturais, entre outras são utilizadas frequentemente como funções de performance [25, 27, 28].

Os métodos de projeto estrutural que utilizam algoritmos de otimização definem uma destas funções de performance como *função objetivo* e utilizam outras para definir restrições nas variáveis de projeto. A função do algoritmo de otimização consiste então

em encontrar os valores para as variáveis de projeto que minimizem a função objetivo, atendendo o conjunto das restrições definidas sobre as variáveis de projeto.

Chamando  $\mathbf{x} \in \mathbb{R}^n$  ao vetor de variáveis de projeto,  $\phi(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$  à função objetivo e  $\gamma(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  à função que define as restrições, o problema de otimização estrutural pode ser formulado como:

$$\begin{aligned} & \text{minimize } \phi(\mathbf{x}), \\ & \text{sujeito a: } \gamma(\mathbf{x}) \geq 0. \end{aligned} \tag{2.57}$$

Algumas das funções de performance presentes no problema dado pela Eq. (2.57) podem ser avaliadas diretamente a partir da variável de projeto,  $\mathbf{x}$ , como por exemplo o volume da estrutura. Outras devem ser avaliadas indiretamente a partir dos resultados da análise estrutural empregada, por exemplo, tensões e deformações. Outras funções de performance dependem das variáveis de projeto tanto diretamente como indiretamente, como acontece por exemplo com as funções que definem a complacência da estrutura e as frequências naturais.

Na forma mais geral, as funções do problema da Eq. (2.57) são definidas na forma:

$$\phi(\mathbf{x}) = f(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})), \tag{2.58}$$

$$\gamma(\mathbf{x}) = \mathbf{g}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})), \tag{2.59}$$

onde  $f : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}$ ,  $\mathbf{g} : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^m$ , e a função  $\tilde{\mathbf{u}} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ , é a função que define as chamadas *variáveis de estado* que são calculadas pelo método de análise estrutural, como por exemplo os deslocamentos nodais no MEF, ou os deslocamentos e forças de superfície nodais definidos sobre o contorno no MEC.

A função  $\tilde{\mathbf{u}}$  é usualmente não linear e fica definida de forma implícita como a solução de uma equação chamada de equação de equilíbrio ou equação de estado do problema. Na forma mais geral, a função  $\tilde{\mathbf{u}}$  fica definida como solução da seguinte equação:

$$\mathbf{h}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) = 0, \tag{2.60}$$

onde a função  $\mathbf{h} : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ .

A análise é dita linear quando a função  $\mathbf{h}$  da Eq. (2.60) tem a seguinte forma:

$$\mathbf{h}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) = \mathbf{A}(\mathbf{x})\tilde{\mathbf{u}}(\mathbf{x}) - \mathbf{b}(\mathbf{x}), \tag{2.61}$$

onde a matriz  $\mathbf{A}(\mathbf{x}) \in \mathbb{R}^{p \times p}$  e o vetor  $\mathbf{b}(\mathbf{x}) \in \mathbb{R}^p$ . Quando a função  $\mathbf{h}$  não é da forma dada pela Eq. (2.61), a análise é dita não linear.

Quando a análise é linear, a avaliação da função  $\tilde{u}$  se reduz à solução de um sistema linear de equações com matriz de coeficientes igual à matriz  $A(x)$ . Quando a análise é não linear, a avaliação da função  $\tilde{u}$  requer a utilização de um método numérico iterativo para resolver o sistema não linear dado pela Eq. (2.60). Por isto último, a solução de problemas de otimização estrutural onde a análise é não linear implica, geralmente, em um custo computacional elevado, já que cada iteração do algoritmo de otimização requer uma ou várias avaliações da função  $\tilde{u}$  e de suas derivadas.

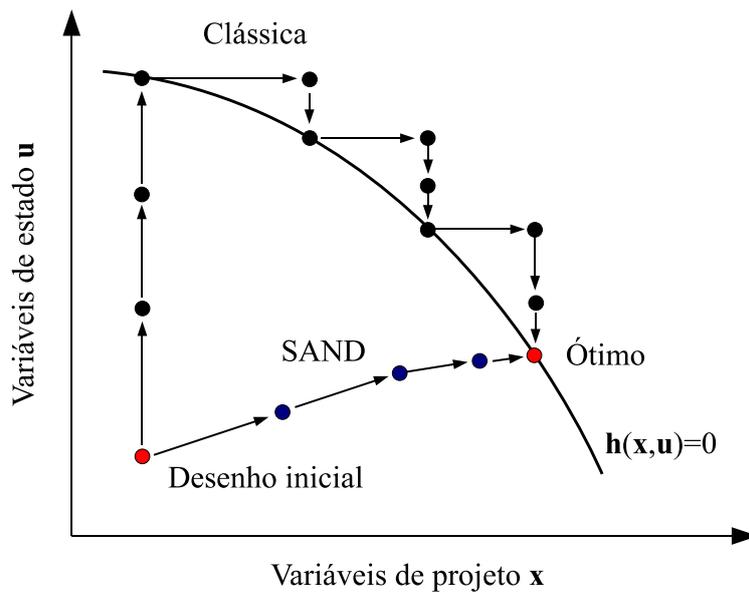


Figura 2.2: Comparação entre as técnicas de otimização clássica e SAND.

Com o objetivo de diminuir o custo computacional relativo à solução de problemas de otimização estrutural de análise não linear, diversos pesquisadores começaram a utilizar a técnica de *Análise e Otimização Simultâneas*, conhecida pela sigla *SAND* da frase em inglês *Simultaneous Analysis and Design* [29–31]. Esta técnica considera as variáveis de estado como incógnitas do problema de otimização e inclui as equações de estado como restrições de igualdade. O problema de otimização estrutural utilizando a técnica SAND é formulado como segue:

$$\begin{aligned} & \text{minimize} && f(\mathbf{x}, \mathbf{u}), \\ & \text{sujeito a:} && \mathbf{g}(\mathbf{x}, \mathbf{u}) \geq 0, \\ & && \mathbf{h}(\mathbf{x}, \mathbf{u}) = 0. \end{aligned} \tag{2.62}$$

A vantagem principal da formulação SAND é que em cada iteração do algoritmo de otimização já não é mais necessária a solução do sistema não linear de equações que

define as variáveis de estado. O principal inconveniente das formulações SAND é o grande incremento do número de variáveis do problema. Afortunadamente, existem também outras vantagens da formulação SAND, que em alguns casos tornam conveniente a sua aplicação inclusive em problemas de análise linear. Dentre as vantagens que já foram citadas estão [29–31]:

1. Torna desnecessários os métodos iterativos para avaliação das variáveis de estado, evitando desta forma os processos iterativos aninhados.
2. A análise de sensibilidade das funções do problema de otimização SAND é, em geral, muito mais simples, por causa da dependência explícita com relação às variáveis do problema.
3. No caso de utilizar o MEF ou outros métodos com discretização do domínio, as funções do problema de otimização SAND tem matrizes de derivadas esparsas. Isto em alguns casos reduz significativamente o custo de armazenamento e de solução dos sistemas lineares dos algoritmos de otimização.
4. É freqüente que muitas das funções que definem o problema e que na formulação clássica são altamente não lineares, na formulação SAND se tornem lineares. Isto, em geral, é aproveitado pelos algoritmos de otimização para reduzir significativamente o custo computacional.
5. Por definir problemas de otimização mais simples torna viável, em alguns casos, a utilização de algoritmos de otimização muito eficientes que existem para problemas de otimização específicos. Por exemplo, algumas vezes é possível a utilização de algoritmos para programação quadrática e inclusive para programação linear.
6. As funções do problema de otimização SAND têm freqüentemente a propriedade de separabilidade parcial (ver Seção 4.1), isto reduz significativamente o custo computacional associado à avaliação das derivadas primeiras e segundas. Isto permite a utilização de algoritmos de tipo Newton que usualmente são muito mais efetivos em problemas de grande escala e que na formulação clássica são de muito difícil aplicação.

O item 1 é uma das vantagens principais quando a análise é não linear. Os itens 2, 3 e 4 aparecem freqüentemente quando é utilizado o MEF. Estes itens foram destacados

no trabalho [32] onde é considerado como exemplo o problema de otimização de treliças, para o qual é mostrado que a utilização da formulação SAND permite um ganho computacional significativo com relação à formulação clássica. Este resultado é surpreendente considerando que o problema é de análise linear e que é utilizado o mesmo algoritmo de otimização para as duas formulações, o algoritmo de tipo quasi-Newton SNOPT [33]. A utilização de um algoritmo de tipo Newton poderia ter sido utilizado para resolver o problema no caso da formulação SAND, como por exemplo foi feito no trabalho [34], onde o número de iterações requerido para resolver esses problemas e outros ainda maiores foi consideravelmente menor. O item 5 é freqüente nos problemas de otimização topológica [28]. Algoritmos de programação linear e semidefinida já foram utilizados com grande sucesso em problemas deste tipo [28, 35]. O item 6 foi destacado no trabalho [36]. A esparsidade citada no item 3 é na verdade um caso particular da propriedade mais geral de separabilidade parcial. É esta última propriedade que permite uma redução significativa do custo computacional associado à avaliação das derivadas primeiras e segundas. Nos trabalhos [36] e [37] é mostrado como esta propriedade pode ser utilizada para calcular eficientemente estas derivadas no caso de utilizar o MEC como método de análise.

### 2.3.1 Formulação SAND para otimização estrutural

Considere primeiro a formulação clássica do problema de otimização estrutural:

$$\begin{aligned} \text{minimize } & \phi(\mathbf{x}) = f(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})), \\ \text{sujeito a: } & \gamma(\mathbf{x}) = \mathbf{g}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) \geq 0, \end{aligned} \quad (2.63)$$

onde a função  $\tilde{\mathbf{u}}$  é definida de forma implícita pela seguinte equação:

$$\mathbf{h}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) = 0. \quad (2.64)$$

As condições de otimalidade de KKT do problema da Eq. (2.63) são:

$$\nabla \phi(\mathbf{x}) - \boldsymbol{\lambda}^T \nabla \gamma(\mathbf{x}) = 0, \quad (2.65)$$

$$\gamma_i(\mathbf{x}) \lambda_i = 0, \quad 1 \leq i \leq m, \quad (2.66)$$

$$\gamma(\mathbf{x}) \geq 0, \quad (2.67)$$

$$\boldsymbol{\lambda} \geq 0. \quad (2.68)$$

As condições acima podem ser expressadas com relação às funções  $f$  e  $g$ . Chamando  $\tilde{\mathbf{z}} = (\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x}))$ , tem-se:

$$\begin{aligned} \nabla_{\mathbf{x}}f(\tilde{\mathbf{z}}) + \nabla_{\mathbf{u}}f(\tilde{\mathbf{z}})\nabla_{\mathbf{x}}\tilde{\mathbf{u}}(\mathbf{x}) \\ - \boldsymbol{\lambda}^T(\nabla_{\mathbf{x}}\mathbf{g}(\tilde{\mathbf{z}}) + \nabla_{\mathbf{u}}\mathbf{g}(\tilde{\mathbf{z}})\nabla_{\mathbf{x}}\tilde{\mathbf{u}}(\mathbf{x})) = 0, \end{aligned} \quad (2.69)$$

$$\mathbf{g}_i(\tilde{\mathbf{z}})\boldsymbol{\lambda}_i = 0, \quad 1 \leq i \leq m, \quad (2.70)$$

$$\mathbf{g}(\tilde{\mathbf{z}}) \geq 0, \quad (2.71)$$

$$\boldsymbol{\lambda} \geq 0. \quad (2.72)$$

As derivadas da função  $\tilde{\mathbf{u}}$  podem ser obtidas a partir da Eq. (2.64) derivando com relação a  $\mathbf{x}$ , assim:

$$\nabla_{\mathbf{x}}\mathbf{h}(\tilde{\mathbf{z}}) + \nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})\nabla_{\mathbf{x}}\tilde{\mathbf{u}}(\mathbf{x}) = 0.$$

A equação acima permite expressar explicitamente o gradiente da função  $\tilde{\mathbf{u}}$  como:

$$\nabla_{\mathbf{x}}\tilde{\mathbf{u}}(\mathbf{x}) = -\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}\nabla_{\mathbf{x}}\mathbf{h}(\tilde{\mathbf{z}}).$$

Substituindo esta expressão na Eq. (2.69) tem-se:

$$\begin{aligned} \nabla_{\mathbf{x}}f(\tilde{\mathbf{z}}) - \nabla_{\mathbf{u}}f(\tilde{\mathbf{z}})\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}\nabla_{\mathbf{x}}\mathbf{h}(\tilde{\mathbf{z}}) \\ - \boldsymbol{\lambda}^T(\nabla_{\mathbf{x}}\mathbf{g}(\tilde{\mathbf{z}}) - \nabla_{\mathbf{u}}\mathbf{g}(\tilde{\mathbf{z}})\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}\nabla_{\mathbf{x}}\mathbf{h}(\tilde{\mathbf{z}})) = 0. \end{aligned}$$

A equação anterior pode ser reescrita em forma compacta como:

$$\nabla_{\mathbf{x}}f(\tilde{\mathbf{z}}) - \boldsymbol{\lambda}^T\nabla_{\mathbf{x}}\mathbf{g}(\tilde{\mathbf{z}}) - (\nabla_{\mathbf{u}}f(\tilde{\mathbf{z}}) - \boldsymbol{\lambda}^T\nabla_{\mathbf{u}}\mathbf{g}(\tilde{\mathbf{z}}))\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}\nabla_{\mathbf{x}}\mathbf{h}(\tilde{\mathbf{z}}) = 0.$$

Utilizando esta equação, as condições de KKT Eq. (2.65)-(2.68) podem ser escritas como:

$$\begin{aligned} \nabla_{\mathbf{x}}f(\tilde{\mathbf{z}}) - \boldsymbol{\lambda}^T\nabla_{\mathbf{x}}\mathbf{g}(\tilde{\mathbf{z}}) \\ - (\nabla_{\mathbf{u}}f(\tilde{\mathbf{z}}) - \boldsymbol{\lambda}^T\nabla_{\mathbf{u}}\mathbf{g}(\tilde{\mathbf{z}}))\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}\nabla_{\mathbf{x}}\mathbf{h}(\tilde{\mathbf{z}}) = 0, \end{aligned} \quad (2.73)$$

$$\mathbf{g}_i(\tilde{\mathbf{z}})\boldsymbol{\lambda}_i = 0, \quad 1 \leq i \leq m, \quad (2.74)$$

$$\mathbf{g}(\tilde{\mathbf{z}}) \geq 0, \quad (2.75)$$

$$\boldsymbol{\lambda} \geq 0. \quad (2.76)$$

A presença da matriz  $\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}$  na Eq. (2.73) torna muito difícil, se não impossível, resolver o problema da Eq. (2.63) utilizando um algoritmo de otimização baseado na

iteração de Newton no sistema dado pelas Eqs. (2.73)-(2.74). A matriz  $\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}$  pode ser evitada definindo o vetor auxiliar  $\boldsymbol{\mu}$  que verifica:

$$\nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^T \boldsymbol{\mu} = \nabla_{\mathbf{u}}f(\tilde{\mathbf{z}})^T - \nabla_{\mathbf{u}}\mathbf{g}(\tilde{\mathbf{z}})^T \boldsymbol{\lambda}. \quad (2.77)$$

A partir desta definição é válida a seguinte igualdade:

$$\boldsymbol{\mu}^T = (\nabla_{\mathbf{u}}f(\tilde{\mathbf{z}}) - \boldsymbol{\lambda}^T \nabla_{\mathbf{u}}\mathbf{g}(\tilde{\mathbf{z}})) \nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}})^{-1}. \quad (2.78)$$

Substituindo esta última expressão nas Eqs. (2.73)-(2.76) e adicionando a Eq. (2.77) tem-se:

$$\nabla_{\mathbf{x}}f(\tilde{\mathbf{z}}) - \boldsymbol{\lambda}^T \nabla_{\mathbf{x}}\mathbf{g}(\tilde{\mathbf{z}}) - \boldsymbol{\mu}^T \nabla_{\mathbf{x}}\mathbf{h}(\tilde{\mathbf{z}}) = 0, \quad (2.79)$$

$$\nabla_{\mathbf{u}}f(\tilde{\mathbf{z}}) - \boldsymbol{\lambda}^T \nabla_{\mathbf{u}}\mathbf{g}(\tilde{\mathbf{z}}) - \boldsymbol{\mu}^T \nabla_{\mathbf{u}}\mathbf{h}(\tilde{\mathbf{z}}) = 0, \quad (2.80)$$

$$\mathbf{g}_i(\tilde{\mathbf{z}})\boldsymbol{\lambda}_i = 0, \quad 1 \leq i \leq m, \quad (2.81)$$

$$\mathbf{g}(\tilde{\mathbf{z}}) \geq 0, \quad (2.82)$$

$$\boldsymbol{\lambda} \geq 0. \quad (2.83)$$

O sistema dado pelas Eqs. (2.79)-(2.83) é bem mais conveniente que o dado pelas Eqs. (2.69)-(2.72). Por outro lado, é fácil ver que estes sistemas são equivalentes. Considere uma solução  $(\mathbf{x}, \boldsymbol{\lambda})$  do sistema das Eqs. (2.69)-(2.72), definindo  $\boldsymbol{\mu}$  pela Eq. (2.78) tem-se que  $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$  verifica todas as Eqs. (2.79)-(2.83). Da mesma forma, partindo de uma solução  $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$  das Eqs. (2.79)-(2.83), é imediato verificar que  $(\mathbf{x}, \boldsymbol{\lambda})$  é solução das Eqs. (2.69)-(2.72).

O maior inconveniente do sistema das Eqs. (2.79)-(2.83) é que as funções são avaliadas no ponto  $(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x}))$ , mas a função  $\tilde{\mathbf{u}}$  é em geral difícil de avaliar. Visto que ela fica definida de forma implícita pela Eq. (2.64), a avaliação da função  $\tilde{\mathbf{u}}$  somente pode ser feita por um método iterativo. Tem-se então uma iteração que é executada dentro da iteração principal do algoritmo, também iterativo, empregado para resolver o sistema das Eqs. (2.79)-(2.83). Como em cada iteração do algoritmo principal podem ser requeridas várias avaliações da função  $\tilde{\mathbf{u}}$ , o custo computacional pode ser elevado.

Uma alternativa a este procedimento é empregar a variável auxiliar  $\mathbf{u}$  em substituição da função  $\tilde{\mathbf{u}}$  e impor a condição  $\tilde{\mathbf{u}}(\mathbf{x}) = \mathbf{u}$  adicionando a equação  $\mathbf{h}(\mathbf{x}, \mathbf{u}) = 0$ . Desta

forma tem-se o seguinte sistema:

$$\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{u}) - \boldsymbol{\lambda}^T \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}, \mathbf{u}) - \boldsymbol{\mu}^T \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}, \mathbf{u}) = 0, \quad (2.84)$$

$$\nabla_{\mathbf{u}} f(\mathbf{x}, \mathbf{u}) - \boldsymbol{\lambda}^T \nabla_{\mathbf{u}} \mathbf{g}(\mathbf{x}, \mathbf{u}) - \boldsymbol{\mu}^T \nabla_{\mathbf{u}} \mathbf{h}(\mathbf{x}, \mathbf{u}) = 0, \quad (2.85)$$

$$\mathbf{g}_i(\mathbf{x}, \mathbf{u}) \boldsymbol{\lambda}_i = 0, \quad 1 \leq i \leq m, \quad (2.86)$$

$$\mathbf{h}(\mathbf{x}, \mathbf{u}) = 0, \quad (2.87)$$

$$\mathbf{g}(\mathbf{x}, \mathbf{u}) \geq 0, \quad (2.88)$$

$$\boldsymbol{\lambda} \geq 0, \quad (2.89)$$

onde, neste caso, as funções devem ser avaliadas no ponto  $(\mathbf{x}, \mathbf{u})$ .

É fácil verificar que as Eqs. (2.84)-(2.89) são exatamente as condições de otimalidade de KKT do problema de otimização na formulação SAND, dado pela Eq. (2.90):

$$\begin{aligned} & \text{minimize} && f(\mathbf{x}, \mathbf{u}), \\ & \text{sujeito a:} && \mathbf{g}(\mathbf{x}, \mathbf{u}) \geq 0, \\ & && \mathbf{h}(\mathbf{x}, \mathbf{u}) = 0. \end{aligned} \quad (2.90)$$

Tem-se então que todo ponto crítico do problema de otimização estrutural clássico é também ponto crítico do problema de otimização obtido utilizando a formulação SAND. A utilização de um método de otimização baseado no método de Newton aplicado nas Eqs. (2.84)-(2.87) requer o cálculo das derivadas segundas das funções  $f$ ,  $\mathbf{h}$  e  $\mathbf{g}$ , porém, estas derivadas são bem mais fáceis de calcular que as derivadas que aparecem na aplicação do método de Newton às Eqs. (2.73)-(2.74) da formulação clássica.

## 2.4 Diferenciação com relação à forma

Nesta seção são introduzidas algumas definições necessárias para esclarecer o conceito de *diferenciação com relação à forma*. Este conceito é útil tanto para caracterizar as formas ótimas quanto para definir métodos numéricos para o cálculo dos gradientes requeridos pelos algoritmos de otimização. Para isso será adotado o ponto de vista de Murat e Simon [38–40], ver também [41].

Considere um domínio de referência  $\Omega_0$ , suposto aberto, limitado e de contorno regular em  $\mathbb{R}^n$ . Para este domínio é definida uma classe de formas admissíveis  $\mathcal{O}(\Omega_0)$  composta pelos abertos  $\Omega = T(\Omega_0)$ , com a transformação  $T$  dada por:

$$T = Id + V, \quad \text{com} \quad V \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n), \quad \|V\|_{W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)} < 1, \quad (2.91)$$

onde  $Id$  é a transformação identidade, isto é  $Id(x) = x, \forall x \in \mathbb{R}^n$ , e  $W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$  é o conjunto das funções lipschitzianas  $V : \mathbb{R}^n \rightarrow \mathbb{R}^n$  tais que  $V$  e  $\nabla V$  são uniformemente limitadas em  $\mathbb{R}^n$ . O conjunto  $W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$  é um espaço Banach quando munido da norma  $\|\cdot\|_{W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)}$  definida como:

$$\|V\|_{W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)} = \sup_{x \in \mathbb{R}^n} (\|V(x)\|_{\mathbb{R}^n} + \|\nabla V(x)\|_{\mathbb{R}^{n \times n}}), \quad (2.92)$$

onde  $\|\cdot\|_{\mathbb{R}^n}$  é a norma euclidiana em  $\mathbb{R}^n$  e  $\|\cdot\|_{\mathbb{R}^{n \times n}}$  é a norma matricial associada.

Para o conjunto  $\mathcal{O}(\Omega_0)$  é válido o seguinte lema:

**Lema 2.21** *Para todo  $V \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$  que verifique  $\|V\|_{W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)} < 1$ , a aplicação  $T = Id + V$  é um difeomorfismo, ou seja, uma bijeção diferenciável de inversa diferenciável. Além disso,  $T$  está no conjunto:*

$$\mathcal{T} = \{T \mid (T - Id) \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n), (T^{-1} - Id) \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)\}. \quad (2.93)$$

Prova: Ver [38, Lema 6.13] □

Assim, o domínio transformado  $\Omega = (Id + V)(\Omega_0)$  fica definido como:

$$\Omega = \{x + V(x) \mid x \in \Omega_0\}, \quad (2.94)$$

ver na Figura 2.3.

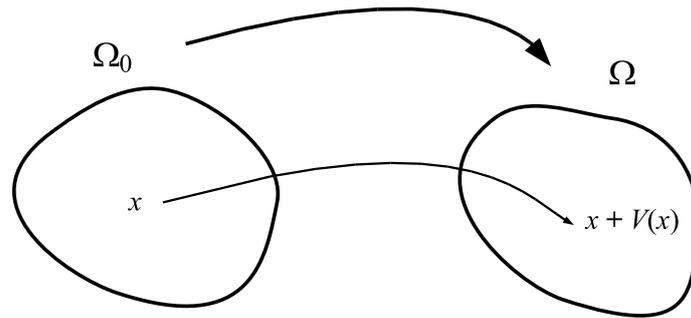


Figura 2.3: Definição do domínio transformado pelo campo vetorial  $V$ .

**Definição 2.22** *Seja  $F(\Omega)$  uma aplicação escalar definida sobre o conjunto de formas admissíveis  $\mathcal{O}(\Omega)$ . Diz-se que  $F$  é diferenciável com relação à forma em  $\Omega_0$  se a função*

$$V \rightarrow F((Id + V)(\Omega_0)) \quad (2.95)$$

*é diferenciável em  $V = 0$  no espaço de Banach  $W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$ .*

A Definição 2.22 pode ser considerada no sentido de Gâteaux ou Fréchet. Por exemplo, no sentido de Fréchet tem-se:

$$F((Id + V)(\Omega_0)) = F(\Omega_0) + F'(\Omega_0)(V) + o(V), \quad \text{com} \quad \lim_{V \rightarrow 0} \frac{o(V)}{\|V\|_{W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)}} = 0,$$

onde  $F'(\Omega_0)$  deve ser linear e contínua sobre  $W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$ .

### 2.4.1 Movimentos e descrições material e espacial

Seja  $\mathcal{X} : \Omega_0 \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$  um *movimento* do domínio  $\Omega_0$ , isto é, a função  $\mathcal{X}$  é inversível e de classe  $C^1(\Omega_0 \times \mathbb{R}^+)$ . O domínio  $\Omega_\tau = \mathcal{X}(\Omega_0, \tau)$  é dado por:

$$\Omega_\tau = \{x_\tau \mid x_\tau = \mathcal{X}(x, \tau), \quad \forall x \in \Omega_0\}.$$

Adotando a nomenclatura da mecânica do contínuo, a *trajetória* do domínio  $\Omega_\tau$  é definida como sendo o conjunto:

$$\mathcal{M} = \{(x_\tau, \tau) \mid x_\tau = \mathcal{X}(x, \tau), \quad \forall x \in \Omega_0 \text{ e } \forall \tau \in \mathbb{R}^+\}.$$

Por ser  $\mathcal{X}$  inversível, existe a função inversa  $\mathcal{X}^{-1} : \mathcal{M} \rightarrow \Omega_0$  que verifica:

$$\mathcal{X}^{-1}(\mathcal{X}(x, \tau), \tau) = x \quad \text{e} \quad \mathcal{X}(\mathcal{X}^{-1}(x_\tau, \tau), \tau) = x_\tau.$$

**Definição 2.23** A representação espacial *de uma função*  $\phi_m : \Omega_0 \times \mathbb{R}^+ \rightarrow \mathbb{R}$  é a função  $\phi_e = [\phi_m]_e : \mathcal{M} \rightarrow \mathbb{R}$  que satisfaz:

$$\phi_e(x_\tau, \tau) = \phi_m(\mathcal{X}^{-1}(x_\tau, \tau), \tau). \quad (2.96)$$

Em quanto que a representação material *de uma função*  $\varphi_e : \mathcal{M} \rightarrow \mathbb{R}$  é a função  $\varphi_m = [\varphi_e]_m : \Omega_0 \times \mathbb{R}^+ \rightarrow \mathbb{R}$  que satisfaz:

$$\varphi_m(x, \tau) = \varphi_e(\mathcal{X}(x, \tau), \tau). \quad (2.97)$$

Claramente, da Definição 2.23 tem-se:

$$\phi_m = \{[\phi_m]_e\}_m \quad \text{e} \quad \varphi_e = \{[\varphi_e]_m\}_e.$$

No seguinte, considere a  $\varphi_m$  e  $\varphi_e$  como as representações material e espacial de um campo escalar, e a  $\varphi_m$  e  $\varphi_e$  como as representações material e espacial de um campo vetorial.

**Definição 2.24** *Os gradientes material e espacial são definidos na forma:*

$$\nabla\varphi_m(x, \tau) = \frac{\partial}{\partial x}\varphi_m(x, \tau) \quad e \quad \nabla_\tau\varphi_e(x_\tau, \tau) = \frac{\partial}{\partial x}\varphi_e(x_\tau, \tau).$$

Assim, chamando  $\mathbf{J} = \nabla\mathcal{X}$ , da Eq. (2.97) tem-se:

$$\nabla\varphi_m = \mathbf{J}^T[\nabla_\tau\varphi_e]_m, \quad (2.98)$$

De forma análoga, para uma função vetorial tem-se:

$$\nabla\varphi_m = [\nabla_\tau\varphi_e]_m\mathbf{J}. \quad (2.99)$$

**Definição 2.25** *As derivadas material e espacial com relação a  $\tau$  são definidas na forma:*

$$\begin{aligned} \dot{\varphi}_m(x, \tau) &= \frac{\partial}{\partial \tau}\varphi_m(x, \tau), \\ \varphi'_e(x_\tau, \tau) &= \frac{\partial}{\partial \tau}\varphi_e(x_\tau, \tau). \end{aligned}$$

Para a derivada material também será utilizada a notação:

$$\frac{D}{D\tau}\varphi_m(x, \tau) = \dot{\varphi}_m(x, \tau).$$

e para a descrição espacial também é definida a derivada material na forma:

$$\frac{D}{D\tau}\varphi_e(x_\tau, \tau) = \dot{\varphi}_e(x_\tau, \tau) = [\dot{\varphi}_m]_e(x_\tau, \tau)$$

Por exemplo, as descrições material e espacial da velocidade são definidas como:

$$\begin{aligned} \mathbf{v}_m(x, \tau) &= \dot{\mathcal{X}}(x, \tau), \\ \mathbf{v}_e(x_\tau, \tau) &= [\mathbf{v}_m]_e(x_\tau, \tau), \end{aligned}$$

Assim, chamando  $\mathbf{L}_\tau = \nabla_\tau\mathbf{v}_e$ , da Eq. (2.99) tem-se:

$$\begin{aligned} \dot{\mathbf{J}} &= \nabla\mathbf{v}_m = [\nabla_\tau\mathbf{v}_e]_m\mathbf{J}, \\ &= [\mathbf{L}_\tau]_m\mathbf{J}. \end{aligned}$$

Por outro lado, da Eq. (2.97) tem-se para  $\varphi_e$  escalar:

$$\dot{\varphi}_e = \varphi'_e + \nabla_\tau\varphi_e \cdot \mathbf{v}_e. \quad (2.100)$$

De forma análoga, para um campo vetorial  $\varphi_e$  tem-se:

$$\dot{\varphi}_e = \varphi'_e + (\nabla_\tau\varphi_e)\mathbf{v}_e. \quad (2.101)$$

Finalmente, a derivada material de  $\nabla_\tau \varphi_e$  pode ser calculada como:

$$\begin{aligned}
(\nabla_\tau \varphi_e)' &= (\nabla_\tau \varphi_e)' + (\nabla_\tau (\nabla_\tau \varphi_e)) \mathbf{v}_e, \\
&= \nabla_\tau (\dot{\varphi}_e - \nabla_\tau \varphi_e \cdot \mathbf{v}_e) + (\nabla_\tau (\nabla_\tau \varphi_e)) \mathbf{v}_e, \\
&= \nabla_\tau \dot{\varphi}_e - (\nabla_\tau \mathbf{v}_e)^T \nabla_\tau \varphi_e.
\end{aligned} \tag{2.102}$$

**Definição 2.26** *O divergente da descrição espacial de um campo vetorial é:*

$$\operatorname{div}_\tau \boldsymbol{\varphi}_e = \operatorname{tr}(\nabla_\tau \boldsymbol{\varphi}_e) = \mathbf{I} \cdot \nabla_\tau \boldsymbol{\varphi}_e.$$

Antes de passar a calcular as derivadas materiais de expressões integrais é necessário o seguinte lema sobre a transformação de domínio dada pela função  $\mathcal{X}(\cdot, \tau)$ :

**Lema 2.27** *Sejam  $\varphi_m$  e  $\varphi_e$  as descrições material e espacial de um campo escalar. Considerando  $\Gamma_\tau = \partial\Omega_\tau$  são válidas as seguintes equações:*

$$\begin{aligned}
\int_{\Omega_\tau} \varphi_e \, d\Omega_\tau &= \int_{\Omega_0} \varphi_m \det(\mathbf{J}) \, d\Omega_0, \\
\int_{\Gamma_\tau} \varphi_e \, d\Gamma_\tau &= \int_{\Gamma_0} \varphi_m \|\mathbf{m}\| \det(\mathbf{J}) \, d\Gamma_0,
\end{aligned}$$

onde  $\mathbf{m} = \mathbf{J}^{-T} \mathbf{n}_0$  define a direção normal a  $\Gamma_\tau$ , ou seja,  $[\mathbf{n}_\tau]_m = \|\mathbf{m}\|^{-1} \mathbf{m}$ . Além disso, são válidas as fórmulas:

$$\begin{aligned}
\frac{D}{D\tau} \det(\mathbf{J}) &= [\operatorname{div}_\tau \mathbf{v}_e]_m \det(\mathbf{J}), \\
\frac{D}{D\tau} (\|\mathbf{m}\| \det(\mathbf{J})) &= [\operatorname{div}_{\Gamma_\tau} \mathbf{v}_e]_m \|\mathbf{m}\| \det(\mathbf{J}),
\end{aligned}$$

onde  $\operatorname{div}_{\Gamma_\tau} \mathbf{v}_e$  é o divergente superficial de  $\mathbf{v}_e$ , que pode ser escrito como:

$$\operatorname{div}_{\Gamma_\tau} \mathbf{v}_e = \operatorname{div}_\tau \mathbf{v}_e - \mathbf{n}_\tau \cdot \mathbf{L}_\tau \mathbf{n}_\tau.$$

Prova: Ver nas referências [19, 20, 42]. □

Considere agora a derivada material da integral no domínio  $\Omega_\tau$ :

$$\frac{D}{D\tau} \int_{\Omega_\tau} \varphi_e \, d\Omega_\tau.$$

**Lema 2.28** *Seja  $\varphi_e$  a descrição espacial de um campo escalar. Considerando o contorno  $\Gamma_\tau = \partial\Omega_\tau$ , tem-se:*

$$\begin{aligned}
\frac{D}{D\tau} \int_{\Omega_\tau} \varphi_e \, d\Omega_\tau &= \int_{\Omega_\tau} \dot{\varphi}_e + \varphi_e \operatorname{div}_\tau \mathbf{v}_e \, d\Omega_\tau, \\
&= \int_{\Omega_\tau} \varphi_e' \, d\Omega_\tau + \int_{\Gamma_\tau} \varphi_e (\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau
\end{aligned}$$

Prova: Pelo Lema 2.27 tem-se:

$$\begin{aligned}
\frac{D}{D\tau} \int_{\Omega_\tau} \varphi_e \, d\Omega_\tau &= \frac{D}{D\tau} \int_{\Omega_0} \varphi_m \det(\mathbf{J}) \, d\Omega_0, \\
&= \int_{\Omega_0} \dot{\varphi}_m \det(\mathbf{J}) + \varphi_m \frac{D}{D\tau} \det(\mathbf{J}) \, d\Omega_0, \\
&= \int_{\Omega_0} (\dot{\varphi}_m + \varphi_m [\operatorname{div}_\tau \mathbf{v}_e]_m) \det(\mathbf{J}) \, d\Omega_0, \\
&= \int_{\Omega_\tau} \dot{\varphi}_e + \varphi_e \operatorname{div}_\tau \mathbf{v}_e \, d\Omega_\tau.
\end{aligned}$$

Utilizando a Eq. (2.100) tem-se:

$$\begin{aligned}
\frac{D}{D\tau} \int_{\Omega_\tau} \varphi_e \, d\Omega_\tau &= \int_{\Omega_\tau} \varphi'_e + \nabla_\tau \varphi_e \mathbf{v}_e + \varphi_e \operatorname{div}_\tau \mathbf{v}_e \, d\Omega_\tau, \\
&= \int_{\Omega_\tau} \varphi'_e \, d\Omega_\tau + \int_{\Omega_\tau} \operatorname{div}_\tau (\varphi_e \mathbf{v}_e) \, d\Omega_\tau, \\
&= \int_{\Omega_\tau} \varphi'_e \, d\Omega_\tau + \int_{\Gamma_\tau} \varphi_e (\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau.
\end{aligned}$$

O que completa a tese do lema. □

Considere agora a derivada material da seguinte integral no contorno  $\Gamma_\tau$  de  $\Omega_\tau$ :

$$\frac{D}{D\tau} \int_{\Gamma_\tau} \varphi_e \, d\Gamma_\tau.$$

**Lema 2.29** *Seja  $\varphi_e$  a descrição espacial de um campo escalar. Considerando o contorno  $\Gamma_\tau = \partial\Omega_\tau$ , tem-se:*

$$\begin{aligned}
\frac{D}{D\tau} \int_{\Gamma_\tau} \varphi_e \, d\Gamma_\tau &= \int_{\Gamma_\tau} \dot{\varphi}_e + \varphi_e \operatorname{div}_{\Gamma_\tau} \mathbf{v}_e \, d\Gamma_\tau, \\
&= \int_{\Gamma_\tau} \varphi'_e \, d\Gamma_\tau + \int_{\Gamma_\tau} \left( \frac{\partial \varphi_e}{\partial \mathbf{n}_\tau} + \varphi_e \mathcal{C} \right) (\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau
\end{aligned}$$

onde  $\mathcal{C}$  é a curvatura média de  $\Gamma_\tau$ .

Prova: Pelo Lema (2.27) tem-se:

$$\begin{aligned}
\frac{D}{D\tau} \int_{\Gamma_\tau} \varphi_e \, d\Gamma_\tau &= \frac{D}{D\tau} \int_{\Gamma_0} \varphi_m \|\mathbf{m}\| \det(\mathbf{J}) \, d\Gamma_\tau, \\
&= \int_{\Gamma_0} \dot{\varphi}_m \|\mathbf{m}\| \det(\mathbf{J}) + \varphi_m \frac{D}{D\tau} (\|\mathbf{m}\| \det(\mathbf{J})) \, d\Gamma_\tau, \\
&= \int_{\Gamma_0} (\dot{\varphi}_m + \varphi_m [\operatorname{div}_{\Gamma_\tau} \mathbf{v}_e]_m) \|\mathbf{m}\| \det(\mathbf{J}) \, d\Gamma_0, \\
&= \int_{\Gamma_\tau} \dot{\varphi}_e + \varphi_e \operatorname{div}_{\Gamma_\tau} \mathbf{v}_e \, d\Gamma_\tau.
\end{aligned}$$

Utilizando a Eq. (2.100) e a definição de  $\text{div}_{\Gamma_\tau}$  tem-se:

$$\frac{D}{D\tau} \int_{\Gamma_\tau} \varphi_e \, d\Gamma_\tau = \int_{\Gamma_\tau} \varphi'_e + \nabla_\tau \varphi_e \cdot \mathbf{v}_e + \varphi_e (\text{div}_\tau \mathbf{v}_e - \mathbf{n}_\tau \cdot \mathbf{L}_\tau \mathbf{n}_\tau) \, d\Gamma_\tau.$$

Considerando o resultado clássico (ver [38, Lemma 6.25]):

$$\int_{\Gamma_\tau} \nabla_\tau \varphi_e \cdot \mathbf{v}_e + \varphi_e (\text{div}_\tau \mathbf{v}_e - \mathbf{n}_\tau \cdot \mathbf{L}_\tau \mathbf{n}_\tau) \, d\Gamma_\tau = \int_{\Gamma_\tau} \left( \frac{\partial \varphi_e}{\partial \mathbf{n}_\tau} + \varphi_e \mathcal{C} \right) (\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau,$$

tem-se a tese do lema. □

## 2.4.2 Derivadas com relação à forma de expressões integrais

Seja  $F : \mathcal{O}(\Omega_0) \rightarrow \mathbb{R}$  definida pela expressão:

$$F(\Omega) = \int_{\Omega} f \, d\Omega,$$

com  $f \in W^{1,1}(\mathbb{R}^n)$ . Então, a derivada de Gâteaux com relação à forma de  $F$  é:

$$F'(\Omega_0)(V) = \left. \frac{d}{d\tau} F((Id + \tau V)(\Omega_0)) \right|_{\tau=0}.$$

Considerando o movimento  $\mathcal{X}(x, \tau) = x + \tau V$ , tem-se:

$$F'(\Omega_0)(V) = \left. \frac{D}{D\tau} \int_{\Omega_\tau} f \, d\Omega_\tau \right|_{\tau=0}.$$

Como  $\mathbf{v}_e = V$ , e  $f' = 0$ , pelo Lema 2.28 tem-se:

$$\begin{aligned} F'(\Omega_0)(V) &= \left. \int_{\Gamma_\tau} f(V \cdot \mathbf{n}_\tau) \, d\Gamma_\tau \right|_{\tau=0}, \\ &= \int_{\Gamma_0} f(V \cdot \mathbf{n}_0) \, d\Gamma_0, \end{aligned}$$

Portanto, como  $f \in W^{1,1}(\mathbb{R}^n)$  e  $V \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$ ,  $F$  é diferenciável no sentido de Fréchet.

Analogamente, para  $\Gamma = \partial\Omega$  e uma função  $F$  definida na forma:

$$F(\Omega) = \int_{\Gamma} f \, d\Gamma,$$

Utilizando o movimento  $\mathcal{X}(x, \tau) = x + \tau V$ , do Lema 2.29 tem-se:

$$F'(\Omega_0)(V) = \int_{\Gamma_0} \left( \frac{\partial f}{\partial \mathbf{n}_\tau} + f \mathcal{C} \right) (V \cdot \mathbf{n}_0) \, d\Gamma_0.$$

Portanto, para  $f \in W^{2,1}(\mathbb{R}^n)$ ,  $F$  é diferenciável no sentido de Fréchet.

### 2.4.3 Análise de sensibilidade considerando o problema variacional

Considere uma função  $F(\Omega)$  definida como  $F(\Omega) = \mathcal{F}(\Omega, u_\Omega)$ . Assim,  $F$  depende explicitamente de  $\Omega$  e implicitamente através de  $u_\Omega$ , definida como a solução do problema variacional que consiste em achar a função em  $\mathcal{U}(\Omega)$  que verifica:

$$a_\Omega(u_\Omega, \eta) = l_\Omega(\eta) \quad \forall \eta \in \mathcal{V}(\Omega).$$

onde  $\mathcal{U}(\Omega)$  e  $\mathcal{V}(\Omega)$  são, respectivamente, os conjuntos das funções e variações admissíveis definidas no domínio  $\Omega$ ,  $a_\Omega$  é um funcional bilinear e  $l_\Omega$  um funcional linear.

Com o intuito de calcular  $F'(\Omega_0)$  considera-se um movimento  $\mathcal{X} : \Omega_0 \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$  e a derivada material:

$$\frac{D}{D\tau} F(\Omega_\tau) = \frac{D}{D\tau} \mathcal{F}(\Omega_\tau, u_\tau), \quad (2.103)$$

onde  $u_\tau \in \mathcal{U}(\Omega_\tau)$  verifica:

$$a_\tau(u_\tau, \eta) = l_\tau(\eta) \quad \forall \eta \in \mathcal{V}(\Omega_\tau).$$

Considere a função lagrangiana:

$$\mathcal{L}_\tau(u, \lambda) = \mathcal{F}(\Omega_\tau, u) + a_\tau(u, \lambda) - l_\tau(\lambda) \quad \forall \lambda \in \mathcal{V}(\Omega_\tau), \forall u \in \mathcal{U}(\Omega_\tau). \quad (2.104)$$

Para o caso  $u = u_\tau$  tem-se:

$$\mathcal{L}_\tau(u_\tau, \lambda) = \mathcal{F}(\Omega_\tau, u_\tau) \quad \forall \lambda \in \mathcal{V}(\Omega_\tau),$$

portanto:

$$\frac{D}{D\tau} \mathcal{L}_\tau(u_\tau, \lambda) = \frac{D}{D\tau} \mathcal{F}(\Omega_\tau, u_\tau) \quad \forall \lambda \in \mathcal{V}(\Omega_\tau), \quad (2.105)$$

O primeiro termo da Eq. (2.105) pode ser calculado como:

$$\frac{D}{D\tau} \mathcal{L}_\tau(u_\tau, \lambda) = \frac{D}{D\tau} \mathcal{L}_\tau(u, \lambda) \Big|_{u=u_\tau} + \left\langle \frac{\partial}{\partial u} \mathcal{L}_\tau(u_\tau, \lambda), \dot{u}_\tau \right\rangle \quad \forall \lambda \in \mathcal{V}(\Omega_\tau), \quad (2.106)$$

onde  $\dot{u}_\tau \in \mathcal{V}(\Omega_\tau)$ . Da definição da função lagrangiana, Eq. (2.104), tem-se:

$$\left\langle \frac{\partial}{\partial u} \mathcal{L}_\tau(u_\tau, \lambda), \dot{u}_\tau \right\rangle = \left\langle \frac{\partial}{\partial u} \mathcal{F}(\Omega_\tau, u_\tau), \dot{u}_\tau \right\rangle + a_\tau(\dot{u}_\tau, \lambda) \quad \forall \lambda \in \mathcal{V}(\Omega_\tau).$$

Assim, é possível anular esse termo escolhendo  $\lambda = \lambda_\tau \in \mathcal{V}(\Omega_\tau)$  como a solução da equação variacional adjunta:

$$\left\langle \frac{\partial}{\partial u} \mathcal{F}(\Omega_\tau, u_\tau), \eta \right\rangle + a_\tau(\eta, \lambda_\tau) = 0 \quad \forall \eta \in \mathcal{V}(\Omega_\tau). \quad (2.107)$$

Portanto, para  $\lambda = \lambda_\tau$ , a Eq. (2.106) fica:

$$\left. \frac{D}{D\tau} \mathcal{L}_\tau(u_\tau, \lambda) \right|_{\lambda=\lambda_\tau} = \left. \frac{D}{D\tau} \mathcal{L}_\tau(u, \lambda) \right|_{\substack{u=u_\tau \\ \lambda=\lambda_\tau}}.$$

Finalmente, considerando as Eqs. (2.103), (2.105) e (2.106) tem-se:

$$\begin{aligned} \frac{D}{D\tau} F(\Omega_\tau) &= \left. \frac{D}{D\tau} \mathcal{L}_\tau(u, \lambda) \right|_{\substack{u=u_\tau \\ \lambda=\lambda_\tau}}, \\ &= \left. \frac{D}{D\tau} \mathcal{F}(\Omega_\tau, u) \right|_{u=u_\tau} + \left. \frac{D}{D\tau} a_\tau(u, \lambda) \right|_{\substack{u=u_\tau \\ \lambda=\lambda_\tau}} - \left. \frac{D}{D\tau} l_\tau(\lambda) \right|_{\lambda=\lambda_\tau}. \end{aligned} \quad (2.108)$$

#### 2.4.4 Exemplo: Problema de Fundição Eletromagnética

Considere o caso em que se deseja obter a derivada material da função  $F(\Omega_\tau) = \mathcal{F}(\Omega_\tau, u_\tau)$ , onde  $u_\tau$  é solução do problema:

$$\begin{aligned} -\Delta_\tau u_\tau &= \mu_0 j_0 \quad \text{em } \Omega_\tau, \\ u_\tau &= 0 \quad \text{sobre } \Gamma_\tau, \\ u_\tau(x) &= O(1) \quad \text{quando } x \rightarrow \infty, \end{aligned} \quad (2.109)$$

onde  $\Gamma_\tau = \partial\Omega_\tau$ , e  $\Omega_\tau$  é o exterior de um domínio compacto  $\omega_\tau$ .

O problema variacional associado ao dado pela Eq. (2.109) consiste em achar a função  $u_\tau \in \mathcal{U}(\Omega_\tau)$  tal que:

$$a_\tau(u_\tau, \eta) = l_\tau(\eta) \quad \forall \eta \in \mathcal{V}(\Omega_\tau),$$

onde, neste caso:

$$\mathcal{U}(\Omega) = \{u \in H^1(\Omega_\tau) \mid u|_{\Gamma_\tau} = 0\} \quad \text{e} \quad \mathcal{V}(\Omega) = \mathcal{U}(\Omega),$$

e as expressões para  $a(u, \eta)$  e  $l(\eta)$  são:

$$a(u, \eta) = \frac{1}{\mu_0} \int_{\Omega_\tau} \nabla_\tau u \cdot \nabla_\tau \eta \, d\Omega_\tau, \quad (2.110)$$

$$l(\eta) = \int_{\Omega_\tau} j_0 \eta \, d\Omega_\tau. \quad (2.111)$$

Considere o funcional:

$$F(\Omega_\tau) = \mathcal{F}(\Omega_\tau, u_\tau) = -\frac{1}{2} a_\tau(u_\tau, u_\tau) - \int_{\omega_\tau} \bar{p} \, d\Omega_\tau + \int_{\Gamma_\tau} \sigma \, d\Gamma_\tau. \quad (2.112)$$

Para esta função o problema adjunto da Eq. (2.107) consiste em encontrar  $\lambda_\tau \in \mathcal{V}(\Omega_\tau)$  que verifica:

$$a_\tau(\eta, \lambda_\tau) = - \left\langle \frac{\partial}{\partial u} \mathcal{F}(\Omega_\tau, u_\tau), \eta \right\rangle = a_\tau(u_\tau, \eta) \quad \forall \eta \in \mathcal{V}(\Omega_\tau).$$

Pela simetria do funcional bilinear, a solução é  $\lambda_\tau = u_\tau$ .

Então, da Eq. (2.108):

$$\begin{aligned} \frac{D}{D\tau} F(\Omega_\tau) &= \frac{D}{D\tau} \mathcal{F}(\Omega_\tau, u) \Big|_{u=u_\tau} + \frac{D}{D\tau} a_\tau(u, \lambda) \Big|_{\substack{u=u_\tau \\ \lambda=\lambda_\tau}} - \frac{D}{D\tau} l_\tau(\lambda) \Big|_{\lambda=\lambda_\tau}, \\ &= -\frac{1}{2} \frac{D}{D\tau} a_\tau(u, u) \Big|_{u=u_\tau} - \frac{D}{D\tau} \int_{\omega_\tau} \bar{p} \, d\Omega_\tau + \frac{D}{D\tau} \int_{\Gamma_\tau} \sigma \, d\Gamma_\tau \\ &\quad + \frac{D}{D\tau} a_\tau(u, \lambda) \Big|_{\substack{u=u_\tau \\ \lambda=\lambda_\tau}} - \frac{D}{D\tau} l_\tau(\lambda) \Big|_{\lambda=\lambda_\tau}, \\ &= \frac{1}{2} \frac{D}{D\tau} a_\tau(u, u) \Big|_{u=u_\tau} - \frac{D}{D\tau} \int_{\omega_\tau} \bar{p} \, d\Omega_\tau + \frac{D}{D\tau} \int_{\Gamma_\tau} \sigma \, d\Gamma_\tau \\ &\quad - \frac{D}{D\tau} l_\tau(u) \Big|_{u=u_\tau}. \end{aligned} \tag{2.113}$$

Para o termo com  $a_\tau(u, u)$ , utilizando a Eq. (2.102) tem-se:

$$\begin{aligned} \frac{D}{D\tau} a_\tau(u, u) &= \frac{1}{\mu_0} \frac{D}{D\tau} \int_{\Omega_\tau} \nabla_\tau u \cdot \nabla_\tau u \, d\Omega_\tau, \\ &= \frac{1}{\mu_0} \int_{\Omega_\tau} (\nabla_\tau u \cdot \nabla_\tau u)' + (\nabla_\tau u \cdot \nabla_\tau u) \operatorname{div}_\tau \mathbf{v}_e \, d\Omega_\tau, \\ &= \frac{1}{\mu_0} \int_{\Omega_\tau} -2(\mathbf{L}_\tau^T \nabla_\tau u) \cdot \nabla_\tau u + (\nabla_\tau u \cdot \nabla_\tau u) \mathbf{I} \cdot \mathbf{L}_\tau \, d\Omega_\tau, \\ &= \frac{1}{\mu_0} \int_{\Omega_\tau} [(\nabla_\tau u \cdot \nabla_\tau u) \mathbf{I} - 2\nabla_\tau u \otimes \nabla_\tau u] \cdot \mathbf{L}_\tau \, d\Omega_\tau, \end{aligned}$$

Para o termo com  $\bar{p}$ :

$$\frac{D}{D\tau} \int_{\omega_\tau} \bar{p} \, d\Omega_\tau = \int_{\Gamma_\tau} -\bar{p} (\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau$$

Para o termo com  $\sigma$ :

$$\frac{D}{D\tau} \int_{\Gamma_\tau} \sigma \, d\Gamma_\tau = \int_{\Gamma_\tau} \sigma \mathcal{C}(\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau$$

Finalmente, para o termo com  $l_\tau(u)$  tem-se:

$$\begin{aligned} \frac{D}{D\tau} l_\tau(u) &= \frac{D}{D\tau} \int_{\Omega_\tau} j_0 u \, d\Omega_\tau, \\ &= \int_{\Omega_\tau} (j_0 u)' \, d\Omega_\tau + \int_{\Gamma_\tau} (j_0 u) \mathbf{v}_e \cdot \mathbf{n}_\tau \, d\Gamma_\tau. \end{aligned}$$

Assim, considerando que  $j'_0 = 0$  e  $\dot{u} = 0$ :

$$\frac{D}{D\tau} l_\tau(u) = \int_{\Omega_\tau} -j_0 \nabla_\tau u \cdot \mathbf{v}_e \, d\Omega_\tau + \int_{\Gamma_\tau} (j_0 u) \mathbf{v}_e \cdot \mathbf{n}_\tau \, d\Gamma_\tau.$$

Substituindo todos estes resultados na Eq. (2.113), considerando que  $u_\tau|_{\Gamma_\tau} = 0$ , tem-se:

$$\frac{D}{D\tau} F(\Omega_\tau) = \int_{\Omega_\tau} \boldsymbol{\Sigma}_\tau \cdot \mathbf{L}_\tau \, d\Omega_\tau + \int_{\Gamma_\tau} (\sigma \mathcal{C} + \bar{p})(\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau + \int_{\Omega_\tau} j_0 \nabla_\tau u_\tau \cdot \mathbf{v}_e \, d\Omega_\tau, \quad (2.114)$$

onde  $\boldsymbol{\Sigma}_\tau$  é o tensor simétrico definido na forma:

$$\boldsymbol{\Sigma}_\tau = \frac{1}{2\mu_0} [(\nabla_\tau u_\tau \cdot \nabla_\tau u_\tau) \mathbf{I} - 2\nabla_\tau u_\tau \otimes \nabla_\tau u_\tau].$$

Através de um cálculo direto mostra-se que  $\text{div} \boldsymbol{\Sigma}_\tau = j_0 \nabla_\tau u_\tau$ . Assim, visto que  $\boldsymbol{\Sigma}_\tau \cdot \mathbf{L}_\tau = \text{div}(\boldsymbol{\Sigma}_\tau^T \mathbf{v}_e) - \text{div}(\boldsymbol{\Sigma}_\tau) \cdot \mathbf{v}_e$  e supondo  $j_0$  com suporte compacto no interior de  $\Omega_\tau$  tem-se:

$$\int_{\Omega_\tau} \boldsymbol{\Sigma}_\tau \cdot \mathbf{L}_\tau \, d\Omega_\tau = \int_{\Gamma_\tau} (\boldsymbol{\Sigma}_\tau \mathbf{n}_\tau) \cdot \mathbf{v}_{\mathbf{n}_\tau} \, d\Gamma_\tau - \int_{\Omega_\tau} j_0 \nabla_\tau u_\tau \cdot \mathbf{v}_e \, d\Omega_\tau, \quad (2.115)$$

onde o primeiro termo pode ser calculado como:

$$\int_{\Gamma_\tau} (\boldsymbol{\Sigma}_\tau \mathbf{n}_\tau) \cdot \mathbf{v}_{\mathbf{n}_\tau} \, d\Gamma_\tau = \int_{\Gamma_\tau} \frac{1}{2\mu_0} [\|\nabla_\tau u_\tau\|^2 (\mathbf{v}_e \cdot \mathbf{n}_\tau) - 2(\nabla_\tau u_\tau \cdot \mathbf{v}_e)(\nabla_\tau u_\tau \cdot \mathbf{n}_\tau)] \, d\Gamma_\tau,$$

Mas, pela condição de contorno  $\mathbf{u}_\tau|_{\Gamma_\tau} = 0$ , portanto  $\nabla_\tau u_\tau|_{\Gamma_\tau} = \alpha_\tau \mathbf{n}_\tau$ . Assim:

$$(\nabla_\tau u_\tau \cdot \mathbf{v}_e)(\nabla_\tau u_\tau \cdot \mathbf{n}_\tau) = \alpha^2 (\mathbf{v}_e \cdot \mathbf{n}_\tau) = \|\nabla_\tau u_\tau\|^2 (\mathbf{v}_e \cdot \mathbf{n}_\tau),$$

Substituindo este resultado nas Eqs. (2.114) e (2.115), tem-se:

$$\frac{D}{D\tau} F(\Omega_\tau) = \int_{\Gamma_\tau} \left( -\frac{1}{2\mu_0} \|\nabla_\tau u_\tau\|^2 + \sigma \mathcal{C} + \bar{p} \right) (\mathbf{v}_e \cdot \mathbf{n}_\tau) \, d\Gamma_\tau, \quad (2.116)$$

Finalmente, considerando o movimento  $\mathcal{X}(x, \tau) = x + \tau V$ , chamando  $\mathbf{v}_0 = -\mathbf{n}_0$  à normal exterior a  $\omega_0$ , e chamando  $\mathcal{C}_\omega = -\mathcal{C}$  à curvatura de  $\omega_0$ , tem-se o resultado:

$$F'(\Omega_0)(V) = \int_{\Gamma_\tau} \left( \frac{1}{2\mu_0} \|\nabla_\tau u_\tau\|^2 + \sigma \mathcal{C}_\omega - \bar{p} \right) (V \cdot \mathbf{v}_0) \, d\Gamma_\tau, \quad (2.117)$$

## Capítulo 3

# Algoritmo de otimização FDIPA-H

Como visto na Seção 2.1.4, o algoritmo FDIPA é baseado na iteração de Newton aplicada ao sistema de equações definido pelas igualdades das condições de otimalidade de primeira ordem de KKT. Porém, para obter a convergência a pontos que satisfaçam estas condições, o FDIPA substitui a matriz hessiana que aparece no sistema linear que define a direção de Newton por uma matriz positiva definida. Esta substituição tem também a vantagem de evitar o cálculo da matriz hessiana, o qual é considerado de grande custo em grande parte dos problemas de otimização não linear da Engenharia Mecânica [27].

Contudo, existem casos onde a matriz hessiana do problema pode ser calculada com um custo computacional relativamente reduzido em comparação com a resolução dos sistemas lineares que definem a direção de busca do algoritmo. Nesses casos permanece o interesse de utilizar a matriz hessiana nos sistemas lineares, com o intuito de obter a solução do problema em um número de iterações menor, reduzindo desta forma o número de resoluções de sistemas lineares e, como consequência, o tempo total requerido pelo algoritmo.

Por exemplo, em problemas de otimização em Engenharia Mecânica para os quais seja utilizada a formulação SAND, a discretização de corpos materiais contínuos conduz a problemas de otimização de grande porte, onde o cálculo da matriz hessiana pode ser realizado eficientemente. Ver, por exemplo, o caso de otimização de forma em elasticidade linear apresentado no Capítulo 4. Com a motivação principal da resolução eficiente destes problemas de grande porte, foram estudadas as condições suficientes que permitem utilizar a matriz hessiana nos sistemas lineares, sem perder a convergência global.

Assim, neste capítulo é descrito um algoritmo de ponto interior, no que segue cha-

mado FDIPA-H, baseado no algoritmo FDIPA apresentado na Seção 2.1.4, e que, da mesma forma que o FDIPA, tem convergência global a pontos de KKT do problema de otimização não linear. Em cada iteração o algoritmo FDIPA-H verifica se a matriz hessiana do problema pode ser utilizada para o cálculo da direção de busca. No caso em que a matriz hessiana não possa ser utilizada, o algoritmo FDIPA-H modifica esta matriz com o intuito de preservar a convergência global. Uma característica importante do algoritmo FDIPA-H é que, nos problemas onde as condições suficientes de otimalidade forem satisfeitas, o algoritmo não alterará a matriz hessiana quando o ponto da iteração atual estiver suficientemente próximo da solução.

Na Seção 3.1 é revista a demonstração de convergência global do algoritmo FDIPA. A maior parte dos resultados apresentados nessa seção foram apresentados por Herskovits [14]. A diferença principal da demonstração apresentada na Seção 3.1 com relação a dada em [14] é a utilização do teorema de convergência global de algoritmos de Zangwill, o Teorema 2.14 da Seção 2.1.2. Este teorema permite estender de forma simples os resultados apresentados em [14], por exemplo, como feito na demonstração da convergência global no caso do problema de otimização que tem restrições de igualdade. Ver o Teorema 3.26 da Seção 3.1.1.

Na Seção 3.2 são apresentados alguns resultados obtidos para as matrizes relacionadas aos sistemas lineares que aparecem no cálculo da direção de busca. Estes resultados, principalmente os Teoremas 3.41 e 3.41 permitem estender os resultados de convergência do FDIPA para o caso do FDIPA-H.

Na Seção 3.3 é definido o algoritmo FDIPA-H, na Seção 3.3.1 são estudadas as condições que permitem utilizar a matriz hessiana para o cálculo da direção de busca, na Seção 3.3.3 é proposto o algoritmo de otimização e na Seção 3.3.4 é mostrada a convergência global do algoritmo para pontos de KKT do problema de otimização.

## 3.1 Convergência global do FDIPA

Considere o problema de otimização sem restrições de igualdade:

$$\begin{aligned} & \text{minimize} && f(\mathbf{x}), \\ & \text{sujeito a:} && \mathbf{g}(\mathbf{x}) \geq 0. \end{aligned} \tag{3.1}$$

**Definição 3.1** *Um ponto  $\mathbf{x} \in \Omega$  é dito estacionário se existe um multiplicador de La-*

grange  $\lambda$  tal que satisfaz as condições:

$$\begin{aligned}\nabla f(\mathbf{x}) - \sum_{i=1}^m \lambda_i \nabla \mathbf{g}_i(\mathbf{x}) &= 0, \\ \mathbf{g}_i(\mathbf{x}) \lambda_i &= 0, \quad 1 \leq i \leq m.\end{aligned}$$

**Definição 3.2** Um ponto  $\mathbf{x} \in \Omega$  é dito ponto de KKT se é estacionário e o multiplicador de Lagrange associado satisfaz  $\lambda \geq 0$ .

São feitas as seguintes suposições sobre o problema de otimização da Eq. (3.1):

**Suposição 3.3** Existe um número real  $a$  tal que o conjunto  $\Omega_a = \{\mathbf{x} \in \Omega \mid f(\mathbf{x}) \leq a\}$  é compacto.

**Suposição 3.4** O requisito RILR dado pela Definição 2.4 é satisfeito para todo ponto  $\mathbf{x}$ , isto é, o conjunto  $\{\nabla \mathbf{g}_i(\mathbf{x}) \mid \mathbf{g}_i(\mathbf{x}) = 0, 1 \leq i \leq m\}$  é linearmente independente para todo  $\mathbf{x}$ .

**Suposição 3.5** As funções  $f$  e  $\mathbf{g}$  são continuamente diferenciáveis em  $\Omega_a$  e as suas derivadas satisfazem a condição de Lipschitz em  $\mathbb{R}^n$ , isto é, existem constantes  $k_f$  e  $k_{\mathbf{g}_i}$ ,  $1 \leq i \leq m$ , tal que para quaisquer  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  tem-se  $\|\nabla f(\mathbf{y}) - \nabla f(\mathbf{x})\| \leq k_f \|\mathbf{y} - \mathbf{x}\|$  e  $\|\nabla \mathbf{g}_i(\mathbf{y}) - \nabla \mathbf{g}_i(\mathbf{x})\| \leq k_{\mathbf{g}_i} \|\mathbf{y} - \mathbf{x}\|$ ,  $1 \leq i \leq m$ .

**Suposição 3.6** Existem números positivos  $\lambda^m, \lambda^M$  tal que:

$$\lambda^m \leq \lambda_i \leq \lambda^M, \quad 1 \leq i \leq m.$$

**Suposição 3.7** Existem números positivos  $\sigma_m$  e  $\sigma_M$  tal que:

$$\sigma_m \leq \sigma_1(\mathbf{B}) \leq \sigma_n(\mathbf{B}) \leq \sigma_M.$$

**Suposição 3.8** Existem números positivos  $\omega_1$  e  $\omega_2$  tal que:

$$\omega_1 \leq \omega_i^D \leq \omega_2, \quad 1 \leq i \leq m.$$

Considerando as suposições anteriores e assumindo  $\mathbf{x} \in \Omega_a$ , são válidos os seguintes lemas:

**Lema 3.9** A função  $(\mathbf{x}, \lambda, \mathbf{B}, \omega^D) \rightarrow (\mathbf{d}_x^\alpha, \lambda^\alpha, \mathbf{d}_x^\beta, \lambda^\beta)$  é contínua, portanto, o ponto  $(\mathbf{d}_x^\alpha, \lambda^\alpha, \mathbf{d}_x^\beta, \lambda^\beta)$  pertence a um compacto.

Prova: Os vetores  $\mathbf{d}_x^\alpha$ ,  $\boldsymbol{\lambda}^\alpha$ ,  $\mathbf{d}_x^\beta$ ,  $\boldsymbol{\lambda}^\beta$  são definidos pela equação:

$$\begin{pmatrix} \mathbf{B} & -\nabla \mathbf{g}^T(\mathbf{x}) \\ -\Lambda \nabla \mathbf{g}(\mathbf{x}) & -\mathbf{G}(\mathbf{x}) \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^\alpha & \mathbf{d}_x^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f^T(\mathbf{x}) & 0 \\ 0 & -\Lambda \boldsymbol{\omega}^D \end{pmatrix}. \quad (3.2)$$

Sem perda de generalidade, considere  $\mathbf{g}(\mathbf{x}) = (\mathbf{g}^I(\mathbf{x}), \mathbf{g}^A(\mathbf{x}))$  com  $\mathbf{g}^I(\mathbf{x}) > 0$  e  $\mathbf{g}^A(\mathbf{x}) = 0$ . De igual forma são divididos os vetores  $\boldsymbol{\lambda} = (\boldsymbol{\lambda}^I, \boldsymbol{\lambda}^A)$ ,  $\boldsymbol{\lambda}^\alpha = (\boldsymbol{\lambda}^{\alpha I}, \boldsymbol{\lambda}^{\alpha A})$  e  $\boldsymbol{\lambda}^\beta = (\boldsymbol{\lambda}^{\beta I}, \boldsymbol{\lambda}^{\beta A})$ . Visto que  $\boldsymbol{\lambda}^I > \boldsymbol{\lambda}^m$ , o sistema linear anterior pode ser escrito como:

$$\begin{pmatrix} \mathbf{B} & -\nabla \mathbf{g}^{IT}(\mathbf{x}) & -\nabla \mathbf{g}^{AT}(\mathbf{x}) \\ -\nabla \mathbf{g}^I(\mathbf{x}) & -\mathbf{S}^I(\mathbf{x}, \boldsymbol{\lambda}^I) & 0 \\ -\nabla \mathbf{g}^A(\mathbf{x}) & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^\alpha & \mathbf{d}_x^\beta \\ \boldsymbol{\lambda}^{\alpha I} & \boldsymbol{\lambda}^{\beta I} \\ \boldsymbol{\lambda}^{\alpha A} & \boldsymbol{\lambda}^{\beta A} \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x})^T & 0 \\ 0 & \boldsymbol{\omega}^{DI} \\ 0 & \boldsymbol{\omega}^{DA} \end{pmatrix},$$

onde  $\mathbf{S}^I$  é definido como  $\mathbf{S}^I(\mathbf{x}, \boldsymbol{\lambda}^I) = (\Lambda^I)^{-1} \mathbf{G}^I(\mathbf{x})$ . Pela Suposição 3.4 e por ser  $\mathbf{S}^I$  positiva definida é válido o Lema 3.39 da Seção 3.2. Então, a matriz do sistema anterior é não singular em todo o domínio compacto definido pelas Suposições 3.3, 3.6, 3.7, e 3.8. Isto prova que a função  $(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{B}, \boldsymbol{\omega}^D) \rightarrow (\mathbf{d}_x^\alpha, \boldsymbol{\lambda}^\alpha, \mathbf{d}_x^\beta, \boldsymbol{\lambda}^\beta)$  é contínua e, por ser o domínio compacto, os pontos  $(\mathbf{d}_x^\alpha, \boldsymbol{\lambda}^\alpha, \mathbf{d}_x^\beta, \boldsymbol{\lambda}^\beta)$  estão também em um domínio compacto.  $\square$

**Lema 3.10** A direção  $\mathbf{d}_x^\alpha$  satisfaz:

$$\nabla f(\mathbf{x}) \mathbf{d}_x^\alpha \leq -\sigma_m \|\mathbf{d}_x^\alpha\|^2. \quad (3.3)$$

Prova: Da primeira linha do sistema Eq. (3.2) tem-se:

$$\mathbf{d}_x^{\alpha T} \mathbf{B} \mathbf{d}_x^\alpha - \mathbf{d}_x^{\alpha T} \nabla \mathbf{g}^T(\mathbf{x}) \boldsymbol{\lambda}^\alpha = -\nabla f(\mathbf{x}) \mathbf{d}_x^\alpha. \quad (3.4)$$

Da segunda linha tem-se:

$$\nabla \mathbf{g}(\mathbf{x}) \mathbf{d}_x^\alpha = -\Lambda^{-1} \mathbf{G}(\mathbf{x}) \boldsymbol{\lambda}^\alpha.$$

Substituindo este valor na Eq. (3.4) fica:  $-\nabla f(\mathbf{x}) \mathbf{d}_x^\alpha = \mathbf{d}_x^{\alpha T} \mathbf{B} \mathbf{d}_x^\alpha + \boldsymbol{\lambda}^{\alpha T} \mathbf{S}(\mathbf{x}, \boldsymbol{\lambda}) \boldsymbol{\lambda}^\alpha$  com  $\mathbf{S}(\mathbf{x}, \boldsymbol{\lambda}) = \Lambda^{-1} \mathbf{G}(\mathbf{x})$ . A Eq. (3.3) obtém-se por ser  $\mathbf{S}(\mathbf{x}, \boldsymbol{\lambda})$  positiva definida e pela Suposição 3.7.  $\square$

**Lema 3.11** A função  $(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{B}, \boldsymbol{\omega}^D) \rightarrow \rho$  é contínua e portanto  $\rho$  pertence a um compacto. Além disso  $\rho$  satisfaz:

$$\varphi_0 \|\mathbf{d}_x^\alpha\|^2 \leq \rho \leq \varphi \|\mathbf{d}_x^\alpha\|^2, \quad (3.5)$$

para algum valor  $\varphi_0 > 0$ .

Prova: A desigualdade  $\rho \leq \varphi \|\mathbf{d}_x^\alpha\|^2$  é válida visto que  $\rho$  é definido como:

$$\rho = \min \left( \varphi \|\mathbf{d}_x^\alpha\|^2, (\xi - 1) \nabla f(\mathbf{x}) \mathbf{d}_x^\alpha / \nabla f(\mathbf{x}) \mathbf{d}_x^\beta \right), \quad \text{caso } \nabla f(\mathbf{x}) \mathbf{d}_x^\beta > 0, \quad (3.6)$$

$$\rho = \varphi \|\mathbf{d}_x^\alpha\|^2, \quad \text{caso contrário.} \quad (3.7)$$

Esta desigualdade prova que a função é contínua em  $\mathbf{d}_x^\alpha = 0$ , onde se obtém  $\rho = 0$ . Para  $\|\mathbf{d}_x^\alpha\| > 0$  tem-se a dificuldade do termo  $\nabla f(\mathbf{x}) \mathbf{d}_x^\beta$  que aparece dividindo na Eq. (3.6). Porém, o termo  $(\xi - 1) \nabla f(\mathbf{x}) \mathbf{d}_x^\alpha / \nabla f(\mathbf{x}) \mathbf{d}_x^\beta$  somente é considerado se for  $\varphi \|\mathbf{d}_x^\alpha\|^2 \geq (\xi - 1) \nabla f(\mathbf{x}) \mathbf{d}_x^\alpha / \nabla f(\mathbf{x}) \mathbf{d}_x^\beta$ , o que somente acontece se for:

$$\begin{aligned} \nabla f(\mathbf{x}) \mathbf{d}_x^\beta &\geq (\xi - 1) \nabla f(\mathbf{x}) \mathbf{d}_x^\alpha / (\varphi \|\mathbf{d}_x^\alpha\|^2), \\ &\geq -(\xi - 1) \sigma_m / \varphi > 0. \end{aligned}$$

onde na última desigualdade foi usado o Lema 3.8. Por outro lado, utilizando o Lema 3.10 tem-se:

$$\rho \geq \min \left( \varphi, -(\xi - 1) \sigma_m / \nabla f(\mathbf{x}) \mathbf{d}_x^\beta \right) \|\mathbf{d}_x^\alpha\|^2.$$

É claro que  $\nabla f(\mathbf{x}) \mathbf{d}_x^\beta$  é limitado superiormente por pertencer a um compacto, portanto existe um valor positivo  $\varphi_0$  tal que  $\rho \geq \varphi_0 \|\mathbf{d}_x^\alpha\|^2$ .  $\square$

**Lema 3.12** A função  $(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{B}, \boldsymbol{\omega}^D) \rightarrow (\mathbf{d}_x, \bar{\boldsymbol{\lambda}})$  é contínua e portanto  $(\mathbf{d}_x, \bar{\boldsymbol{\lambda}})$  pertence a um compacto. Além disso, existe  $\delta > 0$  tal que  $\mathbf{d}_x$  satisfaz:

$$\nabla f(\mathbf{x}) \mathbf{d}_x \leq -\xi \sigma_m \|\mathbf{d}_x^\alpha\|^2, \quad (3.8)$$

$$\|\mathbf{d}_x\| \leq \delta \|\mathbf{d}_x^\alpha\|. \quad (3.9)$$

Prova: A continuidade é consequência dos Lemas 3.9 e 3.11 e da definição de  $\mathbf{d}_x$  e  $\bar{\boldsymbol{\lambda}}$ :

$$\mathbf{d}_x = \mathbf{d}_x^\alpha + \rho \mathbf{d}_x^\beta, \quad (3.10)$$

$$\bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^\alpha + \rho \boldsymbol{\lambda}^\beta. \quad (3.11)$$

Da Eq. (3.10) tem-se:

$$\nabla f(\mathbf{x}) \mathbf{d}_x = \nabla f(\mathbf{x}) \mathbf{d}_x^\alpha + \rho \nabla f(\mathbf{x}) \mathbf{d}_x^\beta,$$

Se for  $\nabla f(\mathbf{x}) \mathbf{d}_x^\beta \leq 0$ , a Eq. (3.8) resulta diretamente da Eq. (3.10) e do Lema 3.10. No caso de ser  $\nabla f(\mathbf{x}) \mathbf{d}_x^\beta > 0$  a Eq. (3.8) resulta das Eqs. (3.6) e (3.10) e do Lema 3.10. Pelo Lema 3.6,  $\rho \leq \varphi \|\mathbf{d}_x^\alpha\|^2$ . Portanto, da Eq. (3.10) tem-se:

$$\|\mathbf{d}_x\| \leq (1 + \varphi \|\mathbf{d}_x^\alpha\| \|\mathbf{d}_x^\beta\|) \|\mathbf{d}_x^\alpha\|.$$

Portanto, a Eq. (3.9) resulta do fato de  $\mathbf{d}_x^\alpha$  e  $\mathbf{d}_x^\beta$  serem limitados.  $\square$

**Lema 3.13** Considere uma função  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  continuamente diferenciável e tal que  $\nabla\phi$  satisfaça a condição de Lipschitz em  $D \subset \mathbb{R}^n$ , isto é, existe uma constante  $k$  tal que para quaisquer  $\mathbf{x}, \mathbf{y} \in D$  tem-se  $\|\nabla\phi(\mathbf{y}) - \nabla\phi(\mathbf{x})\| \leq k\|\mathbf{y} - \mathbf{x}\|$ . Então,  $\phi$  satisfaz:

$$a) \phi(\mathbf{y}) \leq \phi(\mathbf{x}) + \nabla\phi(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) + k\|\mathbf{y} - \mathbf{x}\|^2, \quad (3.12)$$

$$b) \phi(\mathbf{y}) \geq \phi(\mathbf{x}) + \nabla\phi(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) - k\|\mathbf{y} - \mathbf{x}\|^2. \quad (3.13)$$

Prova: Pelo teorema do valor médio tem-se:

$$\begin{aligned} \phi(\mathbf{y}) &= \phi(\mathbf{x}) + \nabla\phi(\mathbf{x} + \xi(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x}), \\ &= \phi(\mathbf{x}) + \nabla\phi(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) + [\nabla\phi(\mathbf{x} + \xi(\mathbf{y} - \mathbf{x})) - \nabla\phi(\mathbf{x})] \cdot (\mathbf{y} - \mathbf{x}), \end{aligned}$$

para algum  $\xi \in (0, 1)$ . Então, tem-se:

$$a) \phi(\mathbf{y}) \leq \phi(\mathbf{x}) + \nabla\phi(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) + \|\nabla\phi(\mathbf{x} + \xi(\mathbf{y} - \mathbf{x})) - \nabla\phi(\mathbf{x})\| \|\mathbf{y} - \mathbf{x}\|,$$

$$b) \phi(\mathbf{y}) \geq \phi(\mathbf{x}) + \nabla\phi(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) - \|\nabla\phi(\mathbf{x} + \xi(\mathbf{y} - \mathbf{x})) - \nabla\phi(\mathbf{x})\| \|\mathbf{y} - \mathbf{x}\|,$$

e as Eq. (3.12)-(3.13) são obtidas pela aplicação da condição de Lipschitz.  $\square$

**Lema 3.14** (Condição uniforme de descida) Existe  $\tau_1 > 0$  tal que para todo  $\mathbf{x} \in \Omega_a$  a condição

$$f(\mathbf{x} + t \mathbf{d}_x) \leq f(\mathbf{x}) + t\eta \nabla f(\mathbf{x}) \mathbf{d}_x, \quad (3.14)$$

é satisfeita para todo  $t \in [0, \tau_1]$ .

Prova: Pela Suposição 3.5 e pelo Lema 3.13 tem-se:

$$f(\mathbf{x} + t \mathbf{d}_x) \leq f(\mathbf{x}) + t \nabla f(\mathbf{x}) \mathbf{d}_x + t^2 k_f \|\mathbf{d}_x\|^2.$$

Então, a condição da Eq. (3.14) é satisfeita para todo  $t$  tal que:

$$f(\mathbf{x}) + t \nabla f(\mathbf{x}) \mathbf{d}_x + t^2 k_f \|\mathbf{d}_x\|^2 \leq f(\mathbf{x}) + t\eta \nabla f(\mathbf{x}) \mathbf{d}_x.$$

ou, equivalentemente:

$$t \leq (\eta - 1) \nabla f(\mathbf{x}) \mathbf{d}_x / (k_f \|\mathbf{d}_x\|^2).$$

Pelo Lema 3.8, esta condição é satisfeita de forma automática se for:

$$t \leq (1 - \eta) \sigma_m / (k_f \delta^2).$$

Portanto,  $\tau_1$  pode ser definido como  $\tau_1 = (1 - \eta) \sigma_m / (k_f \delta^2)$ .  $\square$

**Lema 3.15** (Condição uniforme de viabilidade) Existe  $\tau_2 > 0$  tal que para todo  $\mathbf{x} \in \Omega_a$  onde  $\mathbf{d}_x$  não é nulo, a condição

$$\mathbf{g}_i(\mathbf{x} + t \mathbf{d}_x) > 0, \quad \text{se } \bar{\lambda}_i \geq 0 \text{ ou} \quad (3.15)$$

$$\mathbf{g}_i(\mathbf{x} + t \mathbf{d}_x) \geq \mathbf{g}_i(\mathbf{x}), \quad \text{caso contrário,} \quad (3.16)$$

é satisfeita para todo  $t \in [0, \tau_2]$ .

Prova: Pela Suposição 3.5 e pelo Lema 3.13 tem-se:

$$\mathbf{g}_i(\mathbf{x} + t \mathbf{d}_x) \geq \mathbf{g}_i(\mathbf{x}) + t \nabla \mathbf{g}_i(\mathbf{x}) \mathbf{d}_x - t^2 k_{\mathbf{g}_i} \|\mathbf{d}_x\|^2.$$

Então, as condições das Eqs. (3.15)-(3.16) são satisfeitas para todo  $t$  tal que:

$$\mathbf{g}_i(\mathbf{x}) + t \nabla \mathbf{g}_i(\mathbf{x}) \mathbf{d}_x - t^2 k_{\mathbf{g}_i} \|\mathbf{d}_x\|^2 > 0, \quad \text{se } \bar{\lambda}_i \geq 0 \text{ ou} \quad (3.17)$$

$$\mathbf{g}_i(\mathbf{x}) + t \nabla \mathbf{g}_i(\mathbf{x}) \mathbf{d}_x - t^2 k_{\mathbf{g}_i} \|\mathbf{d}_x\|^2 \geq \mathbf{g}_i(\mathbf{x}), \quad \text{caso contrário.} \quad (3.18)$$

Além disso, das Eqs. (3.2) e (3.10)-(3.11) tem-se:

$$-\lambda_i \nabla \mathbf{g}_i(\mathbf{x}) \mathbf{d}_x - \mathbf{g}_i(\mathbf{x}) \bar{\lambda}_i = -\rho \lambda_i \omega_i^D.$$

Calculando  $\nabla \mathbf{g}_i(\mathbf{x}) \mathbf{d}_x$  a partir desta expressão e substituindo nas Eqs. (3.17)-(3.18) tem-se:

$$\mathbf{g}_i(\mathbf{x})(1 - t \bar{\lambda}_i / \lambda_i) + t(\rho \omega_i^D - t k_{\mathbf{g}_i} \|\mathbf{d}_x\|^2) > 0, \quad \text{se } \bar{\lambda}_i \geq 0 \text{ ou}$$

$$\mathbf{g}_i(\mathbf{x})(1 - t \bar{\lambda}_i / \lambda_i) + t(\rho \omega_i^D - t k_{\mathbf{g}_i} \|\mathbf{d}_x\|^2) \geq \mathbf{g}_i(\mathbf{x}), \quad \text{caso contrário.}$$

A primeira condição é satisfeita se for:

$$\text{a) } 1 - t \bar{\lambda}_i / \lambda_i \geq 0,$$

$$\text{b) } \rho \omega_i^D - t k_{\mathbf{g}_i} \|\mathbf{d}_x\|^2 > 0.$$

Se for  $\bar{\lambda}_i \leq 0$ , basta ver somente a segunda condição. Considere então o caso que  $\bar{\lambda}_i > 0$ .

Então,  $t$  deve satisfazer:

$$\text{a) } t \leq \lambda_i / \bar{\lambda}_i,$$

$$\text{b) } t < \rho \omega_i^D / (k_{\mathbf{g}_i} \|\mathbf{d}_x\|^2).$$

Como  $\lambda_i \geq \lambda_m$ , o Lema 3.9 garante que existe  $\tilde{\lambda}_i$  tal que  $\bar{\lambda}_i \leq \tilde{\lambda}_i$ , e pelos Lemas 3.11 e 3.12 tem-se que  $\tau_2$  pode ser definido, por exemplo, como  $\tau_2 = \min_i \min\{\lambda_m / \tilde{\lambda}_i, 0.8 \varphi_0 \omega_1 / (k_{\mathbf{g}_i} \delta)\}$ .  $\square$

**Lema 3.16** *A direção  $\mathbf{d}_x$  é nula se e somente se  $\mathbf{x}$  é ponto estacionário do problema de otimização. Nesse caso o vetor  $\bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^\alpha$  é o multiplicador de Lagrange associado.*

Prova: Se  $\mathbf{d}_x$  for nulo, então pela Eq. (3.8) deve ser  $\mathbf{d}_x^\alpha = 0$ . Portanto pela Eq. (3.2) o vetor  $\boldsymbol{\lambda}^\alpha$  satisfaz as equações  $\nabla f(\mathbf{x}) = \boldsymbol{\lambda}^{\alpha T} \nabla \mathbf{g}(\mathbf{x})$  e  $\mathbf{G}(\mathbf{x})\boldsymbol{\lambda}^\alpha = 0$ . Isto mostra que  $\mathbf{x}$  é estacionário. Reciprocamente, se  $\mathbf{x}$  for estacionário e  $\tilde{\boldsymbol{\lambda}}$  o multiplicador de Lagrange associado, então os vetores  $\mathbf{d}_x^\alpha = 0$  e  $\boldsymbol{\lambda}^\alpha = \tilde{\boldsymbol{\lambda}}$  serão solução do sistema da Eq. (3.2). Pela Eq. (3.9) tem-se  $\mathbf{d}_x = 0$ . Também, como nesse caso  $\rho = 0$ , tem-se  $\bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^\alpha$ .  $\square$

**Teorema 3.17** *Toda subsequência convergente da seqüência gerada pelo algoritmo FDIPA converge para um ponto estacionário do problema de otimização.*

Prova: Considere o mapeamento ponto-conjunto  $\mathcal{A}$  definido como:

$$\mathcal{A}(\mathbf{x}) = \{(\mathbf{x}, \mathbf{d}_x) \mid \mathbf{d}_x \text{ é definido pelas Eqs. (3.2), (3.6)-(3.7) e (3.10),} \\ \text{e } \mathbf{B}, \boldsymbol{\lambda}, \text{ e } \boldsymbol{\omega}^D \text{ satisfazem as Suposições 3.6-3.8} \}.$$

Pela Suposição 3.3 e pelo Lema 3.12, o mapeamento  $\mathcal{A}$  é fechado e os pontos  $(\mathbf{x}, \mathbf{d}_x)$  pertencem a um compacto. Considere o mapeamento ponto-conjunto  $\mathcal{B}$  definido como:

$$\mathcal{B}(\mathbf{x}, \mathbf{d}_x) = \{\mathbf{x} + t \mathbf{d}_x \mid t \geq \nu \min\{\tau_1, \tau_2\}, \\ f(\mathbf{x} + t \mathbf{d}_x) \leq f(\mathbf{x}) + t\eta \nabla f(\mathbf{x}) \mathbf{d}_x, \\ \mathbf{g}_i(\mathbf{x} + t \mathbf{d}_x) \geq 0, \\ \text{com } \tau_1 \text{ e } \tau_2 \text{ como definidos pelos Lemas 3.14-3.15} \}.$$

Claramente  $\mathcal{B}$  é um mapeamento fechado em  $\mathcal{A}(\mathbf{x})$  e, pelo Corolário 2.18, o mapeamento  $\mathcal{B}\mathcal{A}$  também é. Por outro lado, pelos Lemas 3.12 e 3.16, a função  $f$  é uma função de descida para o conjunto  $E$  dos pontos estacionários e para o mapeamento  $\mathcal{B}\mathcal{A}$ . Pelo Teorema 2.14, qualquer subsequência convergente de uma seqüência gerada pelo algoritmo converge para um ponto no conjunto  $E$ . Como o FDIPA é um caso particular do algoritmo  $\mathcal{B}\mathcal{A}$ , o mesmo acontece com as seqüências geradas por ele.  $\square$

Se o ponto estacionário em  $\Omega_a$  for único, pelo Corolário 2.15 tem-se que a seqüência gerada pelo FDIPA converge para esse ponto. Como a seqüência  $\{f(\mathbf{x}_k)\}_{k \in \mathbb{N}}$  é monótona, obtém-se o mesmo resultado se o ponto estacionário  $\bar{\mathbf{x}}$  para o qual converge uma dada subsequência for único na região  $\{\mathbf{x} \in \Omega \mid f(\mathbf{x}) \leq f(\bar{\mathbf{x}}) + \epsilon\}$  para algum valor  $\epsilon > 0$ .

O seguinte teorema mostra que o limite das subseqüências convergentes são, na verdade, pontos de KKT.

**Teorema 3.18** *Toda subseqüência convergente da seqüência gerada pelo algoritmo FDIPA converge para um ponto de KKT do problema de otimização.*

Prova: Considere uma seqüência  $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$  gerada pelo FDIPA e uma subseqüência  $\{\mathbf{x}_j\}_{j \in \mathbb{K}_1}$ , com  $\mathbb{K}_1 \subset \mathbb{N}$ , convergente para um ponto estacionário  $\bar{\mathbf{x}}$  do problema de otimização. Basta provar que o multiplicador de Lagrange  $\bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^\alpha$  nesse ponto é não negativo. Se for  $\mathbf{g}_i(\bar{\mathbf{x}}) > 0$  tem-se  $\bar{\lambda}_i = 0$ , então resta somente ver o caso  $\mathbf{g}_i(\bar{\mathbf{x}}) = 0$ . Como o FDIPA é estritamente viável, existe uma subseqüência  $\{\mathbf{x}_\ell\}_{\ell \in \mathbb{K}_2}$ , com  $\mathbb{K}_2 \subset \mathbb{K}_1$ , tal que  $\mathbf{g}_i(\mathbf{x}_\ell) < \mathbf{g}_i(\mathbf{x}_{\ell-1})$  e, portanto,  $\bar{\lambda}_i^{\ell-1} \geq 0$ . Note que  $\ell - 1$  pode não pertencer a  $\mathbb{K}_2$ . Porém, pelo Teorema 3.17, existe uma subseqüência  $\{\mathbf{x}_{m-1}\}_{m \in \mathbb{K}_3}$ , com  $\mathbb{K}_3 \subset \mathbb{K}_2$ , convergente a um ponto estacionário  $\tilde{\mathbf{x}}$ . Como  $\mathbf{d}_{\mathbf{x}}(\mathbf{x}, \mathbf{B}, \boldsymbol{\lambda}, \boldsymbol{\omega}^i)$  é uma função contínua e no ponto  $\tilde{\mathbf{x}}$  é nula, então  $\mathbf{d}_{\mathbf{x}}^{m-1} \rightarrow 0$ , mas como  $t \leq 1$  tem-se  $\|\mathbf{x}_m - \mathbf{x}_{m-1}\| \rightarrow 0$ , portanto,  $\tilde{\mathbf{x}} = \bar{\mathbf{x}}$ . Por conseguinte, deve ser  $\bar{\boldsymbol{\lambda}} \geq 0$  no ponto  $\bar{\mathbf{x}}$ , visto que a função  $\bar{\boldsymbol{\lambda}}(\mathbf{x}, \mathbf{B}, \boldsymbol{\lambda}, \boldsymbol{\omega}^i)$  é contínua,  $\bar{\lambda}_i^{\ell-1} \rightarrow \bar{\lambda}$  e  $\bar{\lambda}_i^{\ell-1} \geq 0$ .  $\square$

### 3.1.1 Problema com restrições de igualdade

Neste caso é considerado o problema de otimização com restrições de igualdade:

$$\begin{aligned} & \text{minimize} && f(\mathbf{x}), \\ & \text{sujeito a:} && \mathbf{g}(\mathbf{x}) \geq 0, \\ & && \mathbf{h}(\mathbf{x}) = 0. \end{aligned} \tag{3.19}$$

Para o problema da Eq. (3.19) um ponto  $\mathbf{x} \in \Omega$  é considerado estacionário se existem multiplicadores de Lagrange  $\boldsymbol{\lambda}$  e  $\boldsymbol{\mu}$  tal que satisfazem as condições:

$$\begin{aligned} \nabla f(\mathbf{x}) - \sum_{i=1}^m \lambda_i \nabla \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^p \mu_i \nabla \mathbf{h}_i(\mathbf{x}) &= 0, \\ \mathbf{g}_i(\mathbf{x}) \lambda_i &= 0, \quad 1 \leq i \leq m, \\ \mathbf{h}_i(\mathbf{x}) &= 0, \quad 1 \leq i \leq p, \end{aligned}$$

e é considerado ponto de KKT se, além se ser estacionário, o multiplicador de Lagrange associado às restrições de desigualdade satisfaz  $\boldsymbol{\lambda} \geq 0$ . De forma análoga à Suposição 3.3, é exigido que o conjunto  $\Delta_a = \{\mathbf{x} \in \Delta \mid f(\mathbf{x}) \leq a\}$  seja compacto onde  $\Delta$  é

o conjunto  $\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0, \mathbf{h}(\mathbf{x}) \geq 0\}$ . O requisito RILR dado pela Definição 2.4 é satisfeito para todo ponto  $\mathbf{x}$  considerando, além dos gradientes das restrições de desigualdade ativas, os gradientes de todas as restrições de igualdade. As funções  $f$ ,  $\mathbf{g}$  e  $\mathbf{h}$  são continuamente diferenciáveis em  $\Delta_a$  e as suas derivadas satisfazem a condição de Lipschitz em  $\mathbb{R}^n$ , com constantes de Lipschitz  $k_f, k_{g_i}, 1 \leq i \leq m$ , e  $k_{h_i}, 1 \leq i \leq p$ . Por último, de igual forma que o parâmetro  $\boldsymbol{\omega}^D$ , o parâmetro  $\boldsymbol{\omega}^E$  satisfaz  $\omega_1 \leq \omega_i^E \leq \omega_2, 1 \leq i \leq p$ .

Considerando estas suposições, são válidos os seguintes lemas:

**Lema 3.19** *A função  $(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{B}, \boldsymbol{\omega}^D, \boldsymbol{\omega}^E) \rightarrow (\mathbf{d}_x^\alpha, \boldsymbol{\lambda}^\alpha, \boldsymbol{\mu}^\alpha, \mathbf{d}_x^\beta, \boldsymbol{\lambda}^\beta, \boldsymbol{\mu}^\beta)$  é contínua e, portanto, o ponto  $(\mathbf{d}_x^\alpha, \boldsymbol{\lambda}^\alpha, \boldsymbol{\mu}^\alpha, \mathbf{d}_x^\beta, \boldsymbol{\lambda}^\beta, \boldsymbol{\mu}^\beta)$  pertence a um compacto.*

Prova: Neste caso o sistema linear fica:

$$\begin{pmatrix} \mathbf{B} & -\nabla \mathbf{g}^T(\mathbf{x}) & -\nabla \mathbf{h}^T(\mathbf{x}) \\ -\Lambda \nabla \mathbf{g}(\mathbf{x}) & -\mathbf{G}(\mathbf{x}) & 0 \\ -\nabla \mathbf{h}(\mathbf{x}) & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^\alpha & \mathbf{d}_x^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f^T(\mathbf{x}) & 0 \\ 0 & -\Lambda \boldsymbol{\omega}^D \\ \mathbf{h}(\mathbf{x}) & -\boldsymbol{\omega}^E \end{pmatrix}. \quad (3.20)$$

O resultado obtém-se de forma análoga à do Lema 3.9, separando as restrições ativas das inativas e utilizando o Lema 3.39.  $\square$

Pela existência das restrições de igualdade, o FDIPA em alguns casos será obrigado a aumentar o valor da função objetivo para satisfazer essas restrições. Porém, é possível definir uma função de descida para o algoritmo. No FDIPA é definida como:

$$\phi_c(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^p c_i |\mathbf{h}_i(\mathbf{x})|, \quad (3.21)$$

onde  $\mathbf{c} \in \mathbb{R}^p$  é definido em cada iteração pelo algoritmo.

**Lema 3.20** *Se o vetor  $\mathbf{c}$  satisfaz  $c_i \geq -\mu_i^\alpha, 1 \leq i \leq p$ , então a direção  $\mathbf{d}_x^\alpha$  satisfaz:*

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha \leq -\sigma_m \|\mathbf{d}_x^\alpha\|^2. \quad (3.22)$$

Prova: Da primeira linha do sistema Eq. (3.2) tem-se:

$$\mathbf{d}_x^{\alpha T} \mathbf{B} \mathbf{d}_x^\alpha - \mathbf{d}_x^{\alpha T} \nabla \mathbf{g}^T(\mathbf{x}) \boldsymbol{\lambda}^\alpha - \mathbf{d}_x^{\alpha T} \nabla \mathbf{h}^T(\mathbf{x}) \boldsymbol{\mu}^\alpha = -\nabla f(\mathbf{x}) \mathbf{d}_x^\alpha.$$

Da segunda e terceira linhas tem-se:

$$\begin{aligned}\nabla \mathbf{g}(\mathbf{x}) \mathbf{d}_x^\alpha &= -\Lambda^{-1} \mathbf{G}(\mathbf{x}) \boldsymbol{\lambda}^\alpha, \\ \nabla \mathbf{h}(\mathbf{x}) \mathbf{d}_x^\alpha &= -\mathbf{h}(\mathbf{x}).\end{aligned}$$

Substituindo estes valores na equação anterior fica:

$$\nabla f(\mathbf{x}) \mathbf{d}_x^\alpha = -\mathbf{d}_x^{\alpha T} \mathbf{B} \mathbf{d}_x^\alpha - \boldsymbol{\lambda}^{\alpha T} \mathbf{S}(\mathbf{x}, \boldsymbol{\lambda}) \boldsymbol{\lambda}^\alpha - \mathbf{h}^T(\mathbf{x}) \boldsymbol{\mu}^\alpha, \quad (3.23)$$

onde  $\mathbf{S}(\mathbf{x}, \boldsymbol{\lambda}) = \Lambda^{-1} \mathbf{G}(\mathbf{x})$ . Por outro lado, da Eq. (3.21) tem-se:

$$\begin{aligned}\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha &= \nabla f(\mathbf{x}) \mathbf{d}_x^\alpha + \sum_{i=1}^p \mathbf{c}_i \operatorname{sign}(\mathbf{h}_i(\mathbf{x})) \nabla \mathbf{h}_i \mathbf{d}_x^\alpha, \\ &= \nabla f(\mathbf{x}) \mathbf{d}_x^\alpha - \sum_{i=1}^p \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})|.\end{aligned} \quad (3.24)$$

Combinando as Eqs. (3.23) e (3.24) tem-se:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha = -\mathbf{d}_x^{\alpha T} \mathbf{B} \mathbf{d}_x^\alpha - \boldsymbol{\lambda}^{\alpha T} \mathbf{S}(\mathbf{x}, \boldsymbol{\lambda}) \boldsymbol{\lambda}^\alpha - \sum_{i=1}^p (\mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| + \boldsymbol{\mu}_i^\alpha \mathbf{h}_i(\mathbf{x})). \quad (3.25)$$

A Eq. (3.22) obtém-se por ser  $\mathbf{S}(\mathbf{x}, \boldsymbol{\lambda})$  positiva definida e pela Suposição 3.7.  $\square$

**Lema 3.21** A função  $(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{B}, \boldsymbol{\omega}^D, \boldsymbol{\omega}^E) \rightarrow \rho$  é contínua e portanto  $\rho$  pertence a um compacto. Além disso  $\rho$  satisfaz:

$$\varphi_0 \|\mathbf{d}_x^\alpha\|^2 \leq \rho \leq \varphi \|\mathbf{d}_x^\alpha\|^2, \quad (3.26)$$

para algum valor  $\varphi_0 > 0$ .

Prova: Neste caso  $\rho$  é definido como:

$$\rho = \min(\varphi \|\mathbf{d}_x^\alpha\|^2, (\xi - 1) \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha / \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\beta), \quad \text{caso } \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\beta > 0, \quad (3.27)$$

$$\rho = \varphi \|\mathbf{d}_x^\alpha\|^2, \quad \text{caso contrário}, \quad (3.28)$$

e a demonstração é idêntica à prova do Lema 3.11.  $\square$

**Lema 3.22** A função  $(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{B}, \boldsymbol{\omega}^D, \boldsymbol{\omega}^E) \rightarrow (\mathbf{d}_x, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$  é contínua, e  $(\mathbf{d}_x, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$  pertence a um compacto. Além disso, existe  $\delta > 0$  tal que  $\mathbf{d}_x$  satisfaz:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x \leq -\xi \sigma_m \|\mathbf{d}_x^\alpha\|^2, \quad (3.29)$$

$$\|\mathbf{d}_x\| \leq \delta \|\mathbf{d}_x^\alpha\|. \quad (3.30)$$

Prova: Idêntica à prova do Lema 3.12.  $\square$

**Lema 3.23** (*Condição uniforme de descida*) Existe  $\tau_1 > 0$  tal que para todo  $\mathbf{x} \in \Omega_a$  a condição

$$\phi_{\mathbf{c}}(\mathbf{x} + t \mathbf{d}_{\mathbf{x}}) \leq \phi_{\mathbf{c}}(\mathbf{x}) + t\eta \nabla \phi_{\mathbf{c}}(\mathbf{x}) \mathbf{d}_{\mathbf{x}}, \quad (3.31)$$

é satisfeita para todo  $t \in [0, \tau_1]$ .

Prova: Idêntica à prova do Lema 3.14.  $\square$

**Lema 3.24** (*Condição uniforme de viabilidade*) Existe  $\tau_2 > 0$  tal que para todo  $\mathbf{x} \in \Omega_a$  onde  $\mathbf{d}_{\mathbf{x}}$  não é nulo, as condições

$$\mathbf{g}_i(\mathbf{x} + t \mathbf{d}_{\mathbf{x}}) > 0, \quad \text{se } \bar{\lambda}_i \geq 0 \text{ ou} \quad (3.32)$$

$$\mathbf{g}_i(\mathbf{x} + t \mathbf{d}_{\mathbf{x}}) \geq \mathbf{g}_i(\mathbf{x}), \quad \text{caso contrário,} \quad (3.33)$$

$$\mathbf{h}_i(\mathbf{x} + t \mathbf{d}_{\mathbf{x}}) \geq 0, \quad (3.34)$$

são satisfeitas para todo  $t \in [0, \tau_2]$ .

Prova: Para o caso das restrições de desigualdade é idêntica à prova do Lema 3.15. Para as restrições de igualdade, pela Suposição 3.5 e pelo Lema 3.13 tem-se:

$$\mathbf{h}_i(\mathbf{x} + t \mathbf{d}_{\mathbf{x}}) \geq \mathbf{h}_i(\mathbf{x}) + t \nabla \mathbf{h}_i(\mathbf{x}) \mathbf{d}_{\mathbf{x}} - t^2 k_{\mathbf{h}_i} \|\mathbf{d}_{\mathbf{x}}\|^2.$$

Então, a condição da Eq. (3.34) é satisfeita para todo  $t$  tal que:

$$\mathbf{h}_i(\mathbf{x}) + t \nabla \mathbf{h}_i(\mathbf{x}) \mathbf{d}_{\mathbf{x}} - t^2 k_{\mathbf{h}_i} \|\mathbf{d}_{\mathbf{x}}\|^2 \geq 0. \quad (3.35)$$

Da Eq. (3.20) e por ser  $\mathbf{d}_{\mathbf{x}} = \mathbf{d}_{\mathbf{x}}^{\alpha} + \rho \mathbf{d}_{\mathbf{x}}^{\beta}$  tem-se:

$$-\nabla \mathbf{h}_i(\mathbf{x}) \mathbf{d}_{\mathbf{x}} = \mathbf{h}_i(\mathbf{x}) - \rho \boldsymbol{\omega}_i^E.$$

Calculando  $\nabla \mathbf{h}_i(\mathbf{x}) \mathbf{d}_{\mathbf{x}}$  a partir desta expressão e substituindo na Eq. (3.35) tem-se:

$$\mathbf{h}_i(\mathbf{x})(1 - t) + t(\rho \boldsymbol{\omega}_i^E - t k_{\mathbf{h}_i} \|\mathbf{d}_{\mathbf{x}}\|^2) \geq 0,$$

Esta condição é satisfeita se for:

$$\text{a) } 1 - t \geq 0,$$

$$\text{b) } \rho \boldsymbol{\omega}_i^E - t k_{\mathbf{h}_i} \|\mathbf{d}_{\mathbf{x}}\|^2 \geq 0.$$

Então,  $t$  deve satisfazer:

$$\text{a) } t \leq 1,$$

$$\text{b) } t \leq \rho \omega_i^E / (k_{h_i} \|\mathbf{d}_x\|^2).$$

Portanto, pelos Lemas 3.15, 3.21 e 3.22 tem-se que  $\tau_2$  pode ser definido, por exemplo, como  $\tau_2 = \min_i \min\{1, \lambda_m / \tilde{\lambda}_i, 0.8\varphi_0\omega_1 / (k_{g_i}\delta), \varphi_0\omega_1 / (k_{h_i}\delta)\}$ .  $\square$

**Lema 3.25** *A direção  $\mathbf{d}_x$  é nula se e somente se  $\mathbf{x}$  é ponto estacionário do problema de otimização. Nesse caso os vetores  $\bar{\lambda} = \lambda^\alpha$  e  $\bar{\mu} = \mu^\alpha$  são os multiplicadores de Lagrange associados.*

Prova: Idêntica à prova do Lema 3.16.  $\square$

**Teorema 3.26** *Toda subsequência convergente da seqüência gerada pelo algoritmo FDIPA converge para um ponto estacionário do problema de otimização.*

Prova: Neste caso as aplicações ponto-conjunto são definidas como:

$$\mathcal{A}(\mathbf{x}) = \{(\mathbf{x}, \mathbf{d}_x) \mid \mathbf{d}_x \text{ é definido pelas Eqs. (3.20), (3.27)-(3.28) e (3.10),} \\ \text{e } \mathbf{B}, \lambda, \omega^D, \text{ e } \omega^E \text{ satisfazem as Suposições 3.6-3.8} \}.$$

$$\mathcal{B}_c(\mathbf{x}, \mathbf{d}_x) = \{\mathbf{x} + t \mathbf{d}_x \mid t \geq \nu \min\{\tau_1, \tau_2\}, \\ \phi_c(\mathbf{x} + t \mathbf{d}_x) \leq \phi_c(\mathbf{x}) + t\eta \nabla \phi_c(\mathbf{x}) \mathbf{d}_x, \\ \mathbf{g}_i(\mathbf{x} + t \mathbf{d}_x) \geq 0, \\ \mathbf{h}_i(\mathbf{x} + t \mathbf{d}_x) \geq 0, \\ \text{com } \tau_1 \text{ e } \tau_2 \text{ como definidos pelos Lemas 3.23-3.24} \}.$$

Considere um vetor  $\mathbf{c}_k$  positivo e o conjunto  $E$  dos pontos estacionários ou que tem  $(\mathbf{c}_k)_i \leq -\mu_i^\alpha$ , para algum  $i$ ,  $1 \leq i \leq p$ . Visto que  $\phi_{\mathbf{c}_k}$  é uma função de descida para o conjunto  $E$  e para o mapeamento  $\mathcal{B}_{\mathbf{c}_k}\mathcal{A}$ , pelo Teorema 2.14 tem-se que qualquer subsequência convergente de uma seqüência gerada pelo algoritmo converge para um ponto no conjunto  $E$ . Portanto existem duas possibilidades: Ou a subsequência converge para um ponto estacionário ou converge para um ponto onde  $(\mathbf{c}_k)_i \leq -\mu_i^\alpha$ , para algum  $i$ ,  $1 \leq i \leq p$ . No segundo caso, após um número finito de iterações o FDIPA teria encontrado um ponto onde  $(\mathbf{c}_k)_i < -1.2\mu_i^\alpha$ , para algum  $i$ ,  $1 \leq i \leq p$ , e teria atualizado o vetor  $\mathbf{c}$  pela fórmula  $(\mathbf{c}_{k+1})_i = -2\mu_i^\alpha$ .

Falta mostrar que o algoritmo FDIPA atualiza o parâmetro  $\mathbf{c}_k$  em um número finito de iterações. Considere pelo absurdo que isso não acontece. Então, existe  $i$  tal que  $(\mathbf{c}_k)_i < -\mu_i^\alpha$ , para  $k \in \mathbb{K} \subset \mathbb{N}$ , com  $\mathbb{K}$  infinito. Porém, nesse caso o FDIPA teria atualizado  $(\mathbf{c}_{k+1})_i = -2\mu_i^\alpha \geq (2/1.2)(\mathbf{c}_k)_i$  e, portanto,  $(\mathbf{c}_k)_i \rightarrow +\infty$ . Então, deveria ser  $-\mu_i^\alpha \rightarrow +\infty$  o que não é possível por ser  $\mu^\alpha$  limitado.  $\square$

**Teorema 3.27** *Toda subsequência convergente da seqüência gerada pelo algoritmo FDIPA converge para um ponto de KKT do problema de otimização.*

Prova: Idêntica à prova do Teorema 3.18.  $\square$

Da mesma forma que no caso sem restrições de igualdade, se uma subsequência da seqüência gerada pelo FDIPA converge para um ponto de KKT  $\bar{\mathbf{x}}$ , então a seqüência toda converge para este ponto se ele for o único ponto de KKT na região  $\{\mathbf{x} \in \Delta \mid \phi_{\mathbf{c}}(\mathbf{x}) \leq \phi_{\mathbf{c}}(\bar{\mathbf{x}}) + \epsilon\}$  para algum valor  $\epsilon > 0$ .

## 3.2 Matrizes dos problemas de otimização

Nesta seção são estudadas as matrizes dos sistemas lineares derivados da aplicação do método de Newton às igualdades das condições de otimalidade de KKT de um problema de otimização. Estes sistemas lineares podem ser escritos de modo que a matriz de coeficientes pertença ao conjunto  $\mathcal{O}$  definido como:

$$\mathbf{A} \in \mathcal{O} \Leftrightarrow \mathbf{A} = \begin{matrix} & \begin{matrix} n & m & p \end{matrix} \\ \begin{matrix} n \\ m \\ p \end{matrix} & \begin{pmatrix} \mathbf{B} & \mathbf{C}^T & \mathbf{D}^T \\ \mathbf{C} & -\mathbf{G} & 0 \\ \mathbf{D} & 0 & 0 \end{pmatrix} \end{matrix}. \quad (3.36)$$

Nas seções seguintes será mostrado que algumas propriedades da direção de Newton poderão ser deduzidas da inércia da matriz de coeficientes:

**Definição 3.28** *(Inércia)* A inércia de uma matriz simétrica  $\mathbf{R}$  é a tripla ordenada  $\{i_+, i_-, i_0\}$  onde  $i_+$  é o número de valores próprios positivos de  $\mathbf{R}$ ,  $i_-$  é o número de valores próprios negativos, e  $i_0$  é o número de valores próprios nulos.

O seguinte teorema é muito útil para calcular a inércia de uma matriz:

**Teorema 3.29** (Sylvester) *Seja  $\mathbf{R}$  uma matriz simétrica e  $\mathbf{S}$  uma matriz arbitrária não singular de igual tamanho. Então a inércia da matriz  $\mathbf{R}$  é a mesma da matriz  $\mathbf{SRS}^T$ .*

Prova: Ver [23, Teorema 4.5.8]. □

**Corolário 3.30** *Qualquer permutação simétrica de linhas e colunas preserva a inércia de uma matriz.*

Prova: Considere  $\mathbf{S}$  como sendo uma matriz de permutação de linhas. □

O seguinte teorema é uma versão quantitativa do Teorema 3.29:

**Teorema 3.31** (Ostrowsky) *Seja  $\mathbf{R}$  uma matriz simétrica e  $\mathbf{S}$  uma matriz arbitrária não singular de igual tamanho. Considere os valores próprios ordenados em ordem crescente e seja  $\sigma_k(\mathbf{A})$  o  $k$ -ésimo valor próprio da matriz  $\mathbf{A}$ . Então se verifica:*

$$\sigma_k(\mathbf{SRS}^T) = \theta_k \sigma_k(\mathbf{R}),$$

onde o valor  $\theta_k$  verifica  $\sigma_1(\mathbf{S}) \leq \theta_k \leq \sigma_n(\mathbf{S})$ .

Prova: Ver [23, Teorema 4.5.9]. □

**Teorema 3.32** (Weyl) *Sejam  $\mathbf{R}, \mathbf{S} \in \mathbb{R}^{n \times n}$  simétricas. Então, para quaisquer inteiros  $j$  e  $k$  tal que  $1 \leq j, k \leq n$ , e  $j + k \geq n + 1$  tem-se:*

$$\sigma_{j+k-n}(\mathbf{R} + \mathbf{S}) \leq \sigma_j(\mathbf{R}) + \sigma_k(\mathbf{S}),$$

e, para quaisquer inteiros  $j$  e  $k$  tal que  $1 \leq j, k \leq n$ , e  $j + k \leq n + 1$  tem-se:

$$\sigma_{j+k-1}(\mathbf{R} + \mathbf{S}) \geq \sigma_j(\mathbf{R}) + \sigma_k(\mathbf{S}).$$

Prova: Ver [23, Teorema 4.3.7]. □

**Corolário 3.33** *As funções  $\sigma_j : \mathbb{S}^{n \times n} \rightarrow \mathbb{R}$ ,  $1 \leq j \leq n$ , são contínuas, onde  $\mathbb{S}^{n \times n} \subset \mathbb{R}^{n \times n}$  é o conjunto de matrizes simétricas.*

Prova: Considere a norma espectral  $\|\mathbf{R}\| = \max\{\sqrt{\lambda} \mid \lambda \text{ é valor próprio de } \mathbf{R}^T \mathbf{R}\}$ . Para matrizes simétricas  $\|\mathbf{R}\| = \max_{1 \leq j \leq n} |\sigma_j(\mathbf{R})|$ . Portanto, se  $\|\mathbf{S}\| \leq \epsilon$ , então  $\max(|\sigma_1(\mathbf{R})|, |\sigma_n(\mathbf{R})|) \leq \epsilon$ . Pelo Lema 3.32 com  $k = n$  e  $k = 1$ , tem-se:

$$\text{a) } \sigma_j(\mathbf{R} + \mathbf{S}) \leq \sigma_j(\mathbf{R}) + \epsilon,$$

$$\text{b) } \sigma_j(\mathbf{R} + \mathbf{S}) \geq \sigma_j(\mathbf{R}) - \epsilon,$$

o que prova a continuidade das funções  $\sigma_j$ .  $\square$

Antes de trabalhar no conjunto  $\mathcal{O}$  considere primeiro uma matriz simétrica  $\mathbf{A}$  do tipo:

$$\mathbf{A} = \begin{matrix} & \begin{matrix} n & m \end{matrix} \\ \begin{matrix} n \\ m \end{matrix} & \begin{pmatrix} \mathbf{B} & \mathbf{C}^T \\ \mathbf{C} & -\mathbf{G} \end{pmatrix} \end{matrix}, \quad (3.37)$$

com  $\mathbf{B} \in \mathbb{R}^{n \times n}$  simétrica,  $\mathbf{C} \in \mathbb{R}^{m \times n}$ , e  $\mathbf{G} \in \mathbb{R}^{m \times m}$  simétrica.

**Lema 3.34** *Seja a matriz  $\mathbf{A}$  da Eq. (3.37) com a matriz  $\mathbf{G}$  positiva definida. Então a matriz  $\mathbf{A}$  tem, como mínimo,  $m$  valores próprios negativos.*

Prova: Considere a decomposição:

$$\begin{pmatrix} -\mathbf{G} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & 0 \\ -\mathbf{C}\mathbf{G}^{-1} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} -\mathbf{G} & 0 \\ 0 & \mathbf{M} \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & -\mathbf{G}^{-1}\mathbf{C}^T \\ 0 & \mathbf{I}_n \end{pmatrix}. \quad (3.38)$$

onde  $\mathbf{M} = \mathbf{B} + \mathbf{C}^T\mathbf{G}^{-1}\mathbf{C}$ . Pelo Corolário 3.30, a inércia da matriz  $\mathbf{A}$  é a mesma da matriz à esquerda da Eq. (3.38). Aplicando o Teorema 3.29 tem-se que a inércia da matriz  $\mathbf{A}$  é igual à inércia da matriz central do lado direito da Eq. (3.38). Como a inércia desta matriz é igual à soma das inércias de  $-\mathbf{G}$  e de  $\mathbf{M}$ , portanto, independentemente de  $\mathbf{M}$ ,  $\mathbf{A}$  tem, como mínimo,  $m$  valores próprios negativos.  $\square$

**Lema 3.35** *Seja a matriz  $\mathbf{A}$  da Eq. (3.37) e  $\mathbf{M}$  o complemento de Schur  $\mathbf{M} = \mathbf{B} + \mathbf{C}^T\mathbf{G}^{-1}\mathbf{C}$ . Então, a matriz  $\mathbf{M}$  é positiva definida se e somente se a inércia da matriz  $\mathbf{A}$  é  $\{n, m, 0\}$ .*

Prova: Considere a decomposição da Eq. (3.38). A inércia de  $\mathbf{A}$  é igual à soma das inércias de  $-\mathbf{G}$  e de  $\mathbf{M}$ . Como a inércia de  $-\mathbf{G}$  é  $\{0, m, 0\}$ , se  $\mathbf{M}$  for positiva definida então a inércia de  $\mathbf{A}$  será  $\{n, m, 0\} = \{n, 0, 0\} + \{0, m, 0\}$ . Considere agora que a inércia de  $\mathbf{A}$  é  $\{n, m, 0\}$ . Então a inércia de  $\mathbf{M}$  deve ser  $\{n, 0, 0\} = \{n, m, 0\} - \{0, m, 0\}$  e, portanto,  $\mathbf{M}$  é positiva definida.  $\square$

**Corolário 3.36** *Seja a matriz  $\mathbf{A}$  da Eq. (3.37) e  $\mathbf{M}$  o complemento de Schur da matriz  $\mathbf{G}$ :  $\mathbf{M} = \mathbf{B} + \mathbf{C}^T\mathbf{G}^{-1}\mathbf{C}$ . Se a matriz  $\mathbf{M}$  é positiva definida então a matriz  $\mathbf{A}$  é não singular.*  $\square$

Considere agora uma matriz  $\mathbf{A}$  no conjunto  $\mathcal{O}$ , ou seja, definida como:

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{C}^T & \mathbf{D}^T \\ \mathbf{C} & -\mathbf{G} & \mathbf{0} \\ \mathbf{D} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad (3.39)$$

com  $\mathbf{B} \in \mathbb{R}^{n \times n}$  simétrica,  $\mathbf{C} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{G} \in \mathbb{R}^{m \times m}$  simétrica e  $\mathbf{D} \in \mathbb{R}^{p \times n}$ .

**Definição 3.37** O espaço tangente da matriz  $\mathbf{A}$  é definido como:

$$\mathcal{T} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{D}\mathbf{x} = \mathbf{0}\}.$$

**Definição 3.38** A matriz  $\mathbf{M} \in \mathbb{R}^{n \times n}$  é dita positiva definida no espaço tangente  $\mathcal{T}$  se para qualquer vetor  $\mathbf{x}$  não nulo pertencente a  $\mathcal{T}$  tem-se  $\mathbf{x}^T \mathbf{M} \mathbf{x} > 0$ .

**Lema 3.39** Seja a matriz  $\mathbf{A}$  da Eq. (3.39) com  $\mathbf{G}$  positiva definida. Seja  $\mathbf{M} = \mathbf{B} + \mathbf{C}^T \mathbf{G}^{-1} \mathbf{C}$  positiva definida no espaço tangente e as linhas da matriz  $\mathbf{D}$  linearmente independentes. Então, a matriz  $\mathbf{A}$  é não singular.

Prova: Considere o sistema linear:

$$\begin{pmatrix} \mathbf{B} & \mathbf{C}^T & \mathbf{D}^T \\ \mathbf{C} & -\mathbf{G} & \mathbf{0} \\ \mathbf{D} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

Da terceira equação tem-se que  $\mathbf{D}\mathbf{x} = \mathbf{0}$ , assim, o vetor  $\mathbf{x}$  pertence ao espaço tangente. Da segunda equação tem-se que  $\mathbf{y} = \mathbf{G}^{-1} \mathbf{C}\mathbf{x}$ . Substituindo este valor na primeira equação e multiplicando pela esquerda por  $\mathbf{x}^T$  resulta:

$$\mathbf{x}^T \mathbf{B} \mathbf{x} + \mathbf{x}^T \mathbf{C}^T \mathbf{G}^{-1} \mathbf{C} \mathbf{x} = \mathbf{x}^T \mathbf{M} \mathbf{x} = 0. \quad (3.40)$$

Como a matriz  $\mathbf{M}$  é positiva definida no espaço tangente, o vetor  $\mathbf{x}$  deve ser nulo. Então  $\mathbf{y}$  também é nulo e da primeira equação tem-se que  $\mathbf{D}^T \mathbf{z} = \mathbf{0}$ , mas pela independência linear das linhas da matriz  $\mathbf{D}$  tem-se que  $\mathbf{z} = \mathbf{0}$ . A solução trivial única indica que a matriz  $\mathbf{A}$  é não singular.  $\square$

**Teorema 3.40** Seja a matriz  $\mathbf{A}$  da Eq. (3.39) com  $\mathbf{G}$  positiva definida e as linhas da matriz  $\mathbf{D}$  linearmente independentes. Seja  $\mathbf{M} = \mathbf{B} + \mathbf{C}^T \mathbf{G}^{-1} \mathbf{C}$ . Então, a matriz  $\mathbf{M}$  é positiva definida no espaço tangente se e somente se a inércia da matriz  $\mathbf{A}$  é  $\{n, m+p, 0\}$ .

Prova: Considere a matriz  $\mathbf{A}(\epsilon)$ :

$$\mathbf{A}(\epsilon) = \begin{pmatrix} \mathbf{B} & \mathbf{C}^T & \mathbf{D}^T \\ \mathbf{C} & -\mathbf{G} & 0 \\ \mathbf{D} & 0 & -\epsilon \mathbf{I}_p \end{pmatrix}.$$

Considere o complemento de Schur  $\mathbf{B} + \mathbf{C}^T \mathbf{G}^{-1} \mathbf{C} + \frac{1}{\epsilon} \mathbf{D}^T \mathbf{D} = \mathbf{M} + \frac{1}{\epsilon} \mathbf{D}^T \mathbf{D}$ . É claro que a matriz  $\mathbf{M} + \frac{1}{\epsilon} \mathbf{D}^T \mathbf{D}$  é positiva definida no espaço tangente sempre que a matriz  $\mathbf{M}$  o for. Portanto, pelo Lema 3.35,  $\mathbf{M}$  é positiva definida no espaço tangente se e somente se a inércia de  $\mathbf{A}(\epsilon)$  é  $\{n, m + p, 0\}$ .

Considere  $\mathbf{M}$  positiva definida no espaço tangente. Então  $\sigma_j(\mathbf{A}(\epsilon)) < 0$ , sempre que  $1 \leq j \leq m + p$ , e  $\sigma_j(\mathbf{A}(\epsilon)) > 0$  sempre que  $m + p \leq j \leq n$ . Pelo Corolário 3.33,  $\sigma_j(\mathbf{A}) \leq 0$  sempre que  $1 \leq j \leq m + p$ , e  $\sigma_j(\mathbf{A}) \geq 0$  sempre que  $m + p \leq j \leq n$ . Pelo Lema 3.39 nenhum valor próprio da matriz  $\mathbf{A}$  é nulo, portanto, a sua inércia deve ser  $\{n, m + p, 0\}$ . Reciprocamente, se inércia de  $\mathbf{A}$  for  $\{n, m + p, 0\}$ , pelo Corolário 3.33 existe  $\epsilon > 0$  tal que a inércia de  $\mathbf{A}(\epsilon)$  é  $\{n, m + p, 0\}$  e, portanto,  $\mathbf{M}$  é positiva definida no espaço tangente.  $\square$

Uma outra propriedade das matrizes  $\mathbf{A} \in \mathcal{O}$  de inércia  $\{n, m + p, 0\}$  é dada pelo seguinte teorema:

**Teorema 3.41** *Seja a matriz  $\mathbf{A}$  da Eq. (3.39) com  $\mathbf{G}$  positiva definida e as linhas da matriz  $\mathbf{D}$  linearmente independentes. Seja o subespaço  $\mathcal{S} \subset \mathbb{R}^{n+m+p}$  definido como:*

$$\mathcal{S} = \{ \mathbf{v} = (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathbb{R}^{n+m+p} \mid \mathbf{D}\mathbf{x} = 0; \mathbf{C}\mathbf{x} - \mathbf{G}\mathbf{y} = 0 \}.$$

Então, se a matriz  $\mathbf{A}$  tem inércia  $\{n, m + p, 0\}$ , ela é positiva definida em  $\mathcal{S}$ . Além disso, o valor  $\sigma_1^{\mathcal{S}}(\mathbf{A})$  definido como:

$$\sigma_1^{\mathcal{S}}(\mathbf{A}) = \inf_{\substack{\mathbf{v} \in \mathcal{S} \\ \|\mathbf{v}\|=1}} \mathbf{v}^T \mathbf{A} \mathbf{v}, \quad (3.41)$$

é positivo e a função  $(\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{G}) \rightarrow \sigma_1^{\mathcal{S}}(\mathbf{A})$  é contínua no ponto  $(\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{G})$ .

Prova: Seja um vetor  $\mathbf{v} = (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathcal{S}$ . Para esse vetor considere o produto  $\mathbf{A} \mathbf{v} = (\mathbf{B}\mathbf{x} + \mathbf{C}^T \mathbf{y} + \mathbf{D}^T \mathbf{z}, 0, 0)$ . Portanto, tem-se:

$$\begin{aligned} \mathbf{v}^T \mathbf{A} \mathbf{v} &= \mathbf{x}^T \mathbf{B} \mathbf{x} + \mathbf{x}^T \mathbf{C}^T \mathbf{y} + \mathbf{x}^T \mathbf{D}^T \mathbf{z}, \\ &= \mathbf{x}^T \mathbf{B} \mathbf{x} + \mathbf{x}^T \mathbf{C}^T \mathbf{y}. \end{aligned}$$

Utilizando  $\mathbf{y} = \mathbf{G}^{-1}\mathbf{C}\mathbf{x}$ , tem-se:

$$\begin{aligned}\mathbf{v}^T \mathbf{A} \mathbf{v} &= \mathbf{x}^T \mathbf{B} \mathbf{x} + \mathbf{x}^T \mathbf{C}^T \mathbf{G}^{-1} \mathbf{C} \mathbf{x}, \\ &= \mathbf{x}^T \mathbf{M} \mathbf{x}.\end{aligned}\tag{3.42}$$

Pelo Teorema 3.40 a matriz  $\mathbf{M}$  é positiva definida no espaço tangente  $\mathcal{T}$ , assim, a matriz  $\mathbf{A}$  é positiva definida em  $\mathcal{S}$ . Considere os vetores  $\mathbf{f}^i = (\mathbf{x}^i, \mathbf{y}^i, \mathbf{z}^i)$  definidos na forma:

$$\begin{pmatrix} \mathbf{B} & \mathbf{C}^T & \mathbf{D}^T \\ \mathbf{C} & -\mathbf{G} & 0 \\ \mathbf{D} & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x}^i \\ \mathbf{y}^i \\ \mathbf{z}^i \end{pmatrix} = \begin{pmatrix} \mathbf{e}^i \\ 0 \\ 0 \end{pmatrix},$$

onde  $\{\mathbf{e}^i, 1 \leq i \leq n\}$  é a base canônica de  $\mathbb{R}^n$ . Como a matriz  $\mathbf{A}$  é não singular, o conjunto  $\{\mathbf{f}^i, 1 \leq i \leq n\}$  é uma base para  $\mathcal{S}$ . Seja  $\{\mathbf{v}^i, 1 \leq i \leq n\}$  a base ortonormal de  $\mathcal{S}$  obtida aplicando o processo de Gram-Schmidt à base  $\{\mathbf{f}^i, 1 \leq i \leq n\}$ .

Seja  $\mathbf{v} \in \mathcal{S}$  e a sua representação  $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{v}^i$ . Então, tem-se:

$$\begin{aligned}\mathbf{v}^T \mathbf{A} \mathbf{v} &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \mathbf{v}^{iT} \mathbf{A} \mathbf{v}^j \alpha_j, \\ &= \boldsymbol{\alpha}^T \mathbf{N} \boldsymbol{\alpha},\end{aligned}$$

onde o vetor  $\boldsymbol{\alpha}$  e a matriz  $\mathbf{N}$  satisfazem  $\alpha_i = \alpha_i$ ,  $\mathbf{N}_{ij} = \mathbf{v}^{iT} \mathbf{A} \mathbf{v}^j$ . Portanto, a matriz  $\mathbf{N}$  é positiva definida e como o conjunto  $\{\mathbf{v}^i, 1 \leq i \leq n\}$  é ortonormal tem-se que  $\|\mathbf{v}\| = \|\boldsymbol{\alpha}\|$ . Assim,  $\sigma_1^{\mathcal{S}}(\mathbf{A}) = \sigma_1(\mathbf{N})$  e a continuidade da função  $(\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{G}) \rightarrow \sigma_1(\mathbf{N})$  é assegurada pela continuidade do processo de Gram-Schmidt e o Corolário 3.33.  $\square$

**Corolário 3.42** *Para as hipóteses do teorema anterior, se  $\mathbf{v} = (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathcal{S}$  tem-se:*

$$\mathbf{x}^T \mathbf{M} \mathbf{x} \leq \sigma_1^{\mathcal{S}}(\mathbf{A}) \|\mathbf{x}\|^2.$$

Prova: O resultado sai diretamente das Eqs. (3.41) e (3.42) e por ser  $\|\mathbf{x}\| \leq \|\mathbf{v}\|$ .  $\square$

### 3.2.1 Decomposição de matrizes simétricas indefinidas

Considere uma matriz simétrica  $\mathbf{A} \in \mathbb{R}^{n \times n}$ . Ela é chamada de indefinida se a forma quadrática  $\mathbf{x}^T \mathbf{A} \mathbf{x}$  adota valores positivos e negativos. Os métodos diretos mais populares para a solução de um sistema linear com matriz simétrica indefinida são baseados na decomposição  $\mathbf{LDL}^T$  [23, 24, 43]:

$$\mathbf{PAP}^T = \mathbf{LDL}^T, \tag{3.43}$$

onde  $\mathbf{P}$  é uma matriz de permutação de linhas,  $\mathbf{L}$  é triangular inferior e  $\mathbf{D}$  é diagonal por blocos, com blocos diagonais de dimensões  $1 \times 1$  ou  $2 \times 2$ .

Para iniciar a decomposição, primeiro é escolhida uma permutação  $\mathbf{\Pi}$  e um inteiro  $s = 1$  ou  $2$  de forma tal que:

$$\mathbf{\Pi A \Pi}^T = \begin{matrix} & s & n-s \\ s & \begin{pmatrix} \mathbf{E} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{B} \end{pmatrix} \\ n-s & \end{matrix},$$

com a matriz  $\mathbf{E}$  não singular. Depois de escolher a permutação  $\mathbf{\Pi}$  e feita a decomposição:

$$\mathbf{\Pi A \Pi}^T = \begin{pmatrix} \mathbf{I}_s & 0 \\ \mathbf{C E}^{-1} & \mathbf{I}_{n-s} \end{pmatrix} \begin{pmatrix} \mathbf{E} & 0 \\ 0 & \mathbf{B} - \mathbf{C E}^{-1} \mathbf{C}^T \end{pmatrix} \begin{pmatrix} \mathbf{I}_s & \mathbf{E}^{-1} \mathbf{C}^T \\ 0 & \mathbf{I}_{n-s} \end{pmatrix}.$$

Este processo é repetido no complemento de Schur  $\mathbf{B} - \mathbf{C E}^{-1} \mathbf{C}^T$  e nos sucessivos complementos de Schur até obter a decomposição da Eq. (3.43).

O custo total da decomposição  $\mathbf{LDL}^T$  é  $n^3/3$  operações, igual ao custo da decomposição de Cholesky para uma matriz positiva definida, mais o custo de encontrar as permutações  $\mathbf{\Pi}$ . Existem vários métodos para determinar estas permutações, a estratégia Bunch-Parlett [44] faz uma busca em todo o complemento de Schur e permite encontrar um pivô  $1 \times 1$  ou  $2 \times 2$  não singular de forma tal que o máximo elemento da matriz  $\mathbf{L}$  seja menor em valor absoluto a 2.781. O inconveniente da estratégia Bunch-Parlett é que por buscar em todo o complemento de Schur realiza em total  $O(n^3)$  comparações. A estratégia Bunch-Kaufman [45] limita a busca em no máximo duas colunas da matriz, reduzindo desta forma o número total de comparações a  $O(n^2)$ . A desvantagem da estratégia Bunch-Kaufman é que não existe uma cota para os elementos da matriz  $\mathbf{L}$  como no caso anterior. Existe uma última estratégia proposta por Ashcraft et al. [46], conhecida como Bunch-Kaufman limitada ou BBK, que permite definir uma decomposição com os elementos de  $\mathbf{L}$  limitados pelo mesmo valor 2.781 da estratégia Bunch-Parlett, mas que realiza um número de comparações intermédio entre as estratégias Bunch-Parlett e Bunch-Kaufman. Esta estratégia tem, no pior caso, um custo computacional igual à estratégia Bunch-Parlett mas na prática é similar à estratégia de Bunch-Kaufman.

Existem várias aplicações da decomposição  $\mathbf{LDL}^T$ . A mais evidente é a solução de sistemas lineares. Para o sistema linear de equações  $\mathbf{Ax} = \mathbf{b}$ , se  $\mathbf{PAP}^T = \mathbf{LDL}^T$ , então o vetor  $\mathbf{x} = \mathbf{P}^T \mathbf{y}$ , com  $\mathbf{y} = \mathbf{L}^{-T} \mathbf{D}^{-1} \mathbf{L}^{-1} \mathbf{P} \mathbf{b}$ . Portanto, o vetor  $\mathbf{x}$  pode ser obtido com um custo aproximadamente igual à resolução de dois sistemas triangulares.

Uma outra aplicação desta decomposição é para achar a inércia de uma matriz simétrica, visto que pelo Teorema 3.29 a inércia da matriz  $\mathbf{A}$  é igual à inércia da matriz  $\mathbf{D}$ , sendo esta última muito fácil de calcular já que a matriz  $\mathbf{D}$  é diagonal por blocos.

### 3.3 Algoritmo FDIPA-H

#### 3.3.1 Estudo da direção de Newton

Nesta seção são apresentadas as condições que permitirão definir a matriz  $\mathbf{B}$  do algoritmo FDIPA como sendo a matriz hessiana  $\mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$  no ponto da iteração atual, ou uma matriz obtida a partir desta através da modificação da diagonal. As dificuldades que aparecem com este algoritmo são devidas, principalmente, ao fato de a matriz hessiana não ser sempre positiva definida. Assim, nesta seção são estudadas condições mais fracas que permitem que os sistemas lineares do FDIPA não sejam singulares e de forma tal que a direção  $\mathbf{d}_x$  do FDIPA seja de descida para uma função potencial convenientemente definida. Considere o problema de otimização:

$$\begin{aligned} & \text{minimize} && f(\mathbf{x}), \\ & \text{sujeito a:} && \mathbf{g}(\mathbf{x}) \geq 0, \\ & && \mathbf{h}(\mathbf{x}) = 0. \end{aligned} \tag{3.44}$$

Nesta seção são mostradas as condições suficientes para obter:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha < 0, \tag{3.45}$$

onde  $\phi_c$  é definida como:

$$\phi_c(\mathbf{x}) = f(\mathbf{x}) + c \Psi(\mathbf{h}(\mathbf{x})), \tag{3.46}$$

com o parâmetro  $c \in \mathbb{R}$  definido pelo algoritmo em cada iteração, e  $\Psi : \mathbb{R}^p \rightarrow \mathbb{R}$  é uma função que satisfaz:

$$\Psi(\mathbf{y}) \geq 0, \quad \forall \mathbf{y} \in \mathbb{R}^p, \tag{3.47}$$

$$\Psi(\mathbf{y}) = 0 \Leftrightarrow \mathbf{y} = 0. \tag{3.48}$$

Também será assumido que a função  $\Psi$  satisfaz a desigualdade:

$$\nabla \Psi(\mathbf{y}) \mathbf{y} > 0, \quad \forall \mathbf{y} \neq 0. \tag{3.49}$$

Considere a direção  $\mathbf{d}_x^\alpha$  do algoritmo FDIPA:

$$\begin{pmatrix} \mathbf{B} & -\nabla \mathbf{g}(\mathbf{x})^T & -\nabla \mathbf{h}(\mathbf{x})^T \\ -\Lambda \nabla \mathbf{g}(\mathbf{x}) & -\mathbf{G}(\mathbf{x}) & 0 \\ -\nabla \mathbf{h}(\mathbf{x}) & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^\alpha \\ \boldsymbol{\lambda}^\alpha \\ \boldsymbol{\mu}^\alpha \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x})^T \\ 0 \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}. \quad (3.50)$$

O vetor das restrições de desigualdade pode ser separado no vetor das restrições ativas e o vetor das restrições inativas. Se perda de generalidade, considere  $\mathbf{g}(\mathbf{x}) = (\mathbf{g}^I(\mathbf{x}), \mathbf{g}^A(\mathbf{x}))$  com  $\mathbf{g}^I(\mathbf{x}) > 0$  e  $\mathbf{g}^A(\mathbf{x}) = 0$ . De igual forma é dividido o vetor  $\boldsymbol{\lambda} = (\boldsymbol{\lambda}^I, \boldsymbol{\lambda}^A)$ .

**Definição 3.43** Define-se a matriz  $\mathbf{M} \in \mathbb{R}^{n \times n}$  como:

$$\mathbf{M} = \mathbf{B} + \nabla \mathbf{g}^I(\mathbf{x})^T \mathbf{G}^I(\mathbf{x})^{-1} \Lambda^I \nabla \mathbf{g}^I(\mathbf{x}). \quad (3.51)$$

**Lema 3.44** Se os gradientes das restrições de desigualdade ativas e das de igualdade são linearmente independentes, a matriz  $\mathbf{M}$  é positiva definida no espaço tangente definido por essas restrições,  $\boldsymbol{\lambda} > 0$  e  $\mathbf{g}(\mathbf{x}) \geq 0$ , então a matriz de coeficientes do sistema linear Eq. (3.50) é não singular.

Prova: Visto que  $\boldsymbol{\lambda} > 0$ , a matriz  $\Lambda$  é não singular. Multiplicando pela esquerda a segunda linha do sistema da Eq. (3.50) por  $\Lambda^{-1}$  e definindo  $\mathbf{S}^I(\mathbf{x}, \boldsymbol{\lambda}) = (\Lambda^I)^{-1} \mathbf{G}^I(\mathbf{x})$ , o sistema linear da Eq. (3.50) pode ser escrito como:

$$\begin{pmatrix} \mathbf{B} & -\nabla \mathbf{g}^I(\mathbf{x})^T & -\nabla \mathbf{g}^A(\mathbf{x})^T & -\nabla \mathbf{h}(\mathbf{x})^T \\ -\nabla \mathbf{g}^I(\mathbf{x}) & -\mathbf{S}^I(\mathbf{x}, \boldsymbol{\lambda}) & 0 & 0 \\ -\nabla \mathbf{g}^A(\mathbf{x}) & 0 & 0 & 0 \\ -\nabla \mathbf{h}(\mathbf{x}) & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^\alpha \\ \boldsymbol{\lambda}^{\alpha I} \\ \boldsymbol{\lambda}^{\alpha A} \\ \boldsymbol{\mu}^\alpha \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x})^T \\ 0 \\ 0 \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}.$$

Pelo Lema 3.39 da Seção 3.2, a matriz do sistema anterior é não singular. Portanto, também é não singular a matriz do sistema da Eq. (3.50).  $\square$

**Lema 3.45** O vetor  $\mathbf{d}_x^\alpha = 0$  se e somente se  $\mathbf{x}$  é ponto estacionário do Problema (3.44).

Prova: Se  $\mathbf{d}_x^\alpha$  for nulo, então pela Eq. (3.50) os vetores  $\boldsymbol{\lambda}^\alpha$  e  $\boldsymbol{\mu}^\alpha$  satisfazem as equações  $\nabla f(\mathbf{x}) = \boldsymbol{\lambda}^{\alpha T} \nabla \mathbf{g}(\mathbf{x}) + \boldsymbol{\mu}^{\alpha T} \nabla \mathbf{h}(\mathbf{x})$ ,  $\mathbf{G}(\mathbf{x}) \boldsymbol{\lambda}^\alpha = 0$  e  $\mathbf{h}(\mathbf{x}) = 0$ . Portanto,  $\mathbf{x}$  é estacionário. Reciprocamente, se  $\mathbf{x}$  for estacionário e  $\tilde{\boldsymbol{\lambda}}$  e  $\tilde{\boldsymbol{\mu}}$  forem os multiplicadores de Lagrange associados, então os vetores  $\mathbf{d}_x^\alpha = 0$ ,  $\boldsymbol{\lambda}^\alpha = \tilde{\boldsymbol{\lambda}}$  e  $\boldsymbol{\mu}^\alpha = \tilde{\boldsymbol{\mu}}$  são a solução do sistema da Eq. (3.50).  $\square$

O lema seguinte mostra que sempre pode ser escolhido um valor  $c$  positivo de forma tal que a direção  $\mathbf{d}_x$ , solução do sistema linear da Eq. (3.50), seja de descida para a função  $\phi_c$ .

**Teorema 3.46** *Se os gradientes das restrições ativas são linearmente independentes, a matriz  $\mathbf{M}$  é positiva definida no espaço tangente,  $\lambda > 0$ ,  $\mathbf{g}(\mathbf{x}) \geq 0$ , e  $\mathbf{x}$  não é ponto estacionário, então existe um valor  $c_0$  tal que para todo  $c > c_0$  a direção  $\mathbf{d}_x^\alpha$  é de descida para a função potencial  $\phi_c$  dada na Eq. (3.46).*

Prova: Considere os sistemas lineares:

$$\begin{pmatrix} \mathbf{B} & -\nabla \mathbf{g}^I(\mathbf{x})^T & -\nabla \mathbf{g}^A(\mathbf{x})^T & -\nabla \mathbf{h}(\mathbf{x})^T \\ -\nabla \mathbf{g}^I(\mathbf{x}) & -\mathbf{S}^I(\mathbf{x}, \lambda) & 0 & 0 \\ -\nabla \mathbf{g}^A(\mathbf{x}) & 0 & 0 & 0 \\ -\nabla \mathbf{h}(\mathbf{x}) & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^1 & \mathbf{d}_x^2 \\ \lambda^{I1} & \lambda^{I2} \\ \lambda^{A1} & \lambda^{A2} \\ \mu^1 & \mu^2 \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x})^T & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \mathbf{h}(\mathbf{x}) \end{pmatrix}. \quad (3.52)$$

Pelo Lema 3.44 os sistemas anteriores têm solução. Claramente  $\mathbf{d}_x^1 + \mathbf{d}_x^2 = \mathbf{d}_x^\alpha$ , e também  $\lambda^{I1} + \lambda^{I2} = \lambda^I$ ,  $\lambda^{A1} + \lambda^{A2} = \lambda^A$  e  $\mu^1 + \mu^2 = \mu$ . Então tem-se  $\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha = \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^1 + \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2$ . Da terceira e quarta linhas da Eq. (3.52) tem-se:

$$\nabla \mathbf{g}^A(\mathbf{x}) \mathbf{d}_x^1 = 0, \quad (3.53)$$

$$\nabla \mathbf{h}(\mathbf{x}) \mathbf{d}_x^1 = 0, \quad (3.54)$$

o que mostra que o vetor  $\mathbf{d}_x^1$  está no espaço tangente. Assim:

$$\begin{aligned} \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^1 &= \nabla f(\mathbf{x}) \mathbf{d}_x^1 + c \nabla \Psi(\mathbf{h}(\mathbf{x})) \nabla \mathbf{h}(\mathbf{x}) \mathbf{d}_x^1, \\ &= \nabla f(\mathbf{x}) \mathbf{d}_x^1, \end{aligned} \quad (3.55)$$

onde a Eq. (3.55) é obtida a partir da Eq. (3.54).

Da segunda linha da Eq. (3.52) tem-se  $\lambda^{I1} = -\mathbf{S}^I(\mathbf{x}, \lambda)^{-1} \nabla \mathbf{g}^I(\mathbf{x}) \mathbf{d}_x^1$ . Substituindo este valor na primeira linha e multiplicando por  $\mathbf{d}_x^1$  tem-se:

$$\begin{aligned} \nabla f(\mathbf{x}) \mathbf{d}_x^1 &= -\mathbf{d}_x^{1T} \mathbf{B} \mathbf{d}_x^1 - \mathbf{d}_x^{1T} \nabla \mathbf{g}^I(\mathbf{x})^T (\mathbf{S}^I(\mathbf{x}, \lambda))^{-1} \nabla \mathbf{g}^I(\mathbf{x}) \mathbf{d}_x^1 + \\ &+ \mathbf{d}_x^{1T} \nabla \mathbf{g}^A(\mathbf{x})^T \bar{\lambda}^{A1} + \mathbf{d}_x^{1T} \nabla \mathbf{h}(\mathbf{x})^T \bar{\mu}^1, \end{aligned} \quad (3.56)$$

$$= -\mathbf{d}_x^{1T} \mathbf{M} \mathbf{d}_x^1 \leq 0, \quad (3.57)$$

$$< 0 \quad \text{se } \|\mathbf{d}_x^1\| \neq 0, \quad (3.58)$$

onde a Eq. (3.57) é obtida a partir das Eqs. (3.56), (3.53) e (3.54) e a desigualdade nas Eqs. (3.57)-(3.58) é devida a estar  $\mathbf{d}_x^1$  no plano tangente.

Para a direção  $\mathbf{d}_x^2$ , da quarta linha da Eq. (3.52) tem-se:

$$\nabla \mathbf{h}(\mathbf{x}) \mathbf{d}_x^2 = -\mathbf{h}(\mathbf{x}).$$

Portanto, para a função  $\phi_c$  tem-se:

$$\begin{aligned} \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 &= \nabla f(\mathbf{x}) \mathbf{d}_x^2 + c \nabla \Psi(\mathbf{h}(\mathbf{x})) \nabla \mathbf{h}(\mathbf{x}) \mathbf{d}_x^2, \\ &= \nabla f(\mathbf{x}) \mathbf{d}_x^2 - c \nabla \Psi(\mathbf{h}(\mathbf{x})) \mathbf{h}(\mathbf{x}). \end{aligned} \quad (3.59)$$

As Eqs. (3.59) e (3.49) mostram que sempre é possível escolher  $c$  para ter  $\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 \leq 0$ , pois ainda no caso  $\mathbf{h}(\mathbf{x}) = 0$ , da Eq. (3.52) tem-se  $\mathbf{d}_x^2 = 0$  e qualquer valor para  $c$  fornece o mesmo resultado  $\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 = 0$ .

Assim, se  $\|\mathbf{d}_x^1\| > 0$ , basta escolher  $c_0$  tal que  $\nabla \phi_{c_0}(\mathbf{x}) \mathbf{d}_x^2 \leq 0$ . Se for  $\mathbf{d}_x^1 = 0$ , então como  $\mathbf{x}$  não é ponto de KKT,  $\mathbf{d}_x^2$  não pode ser nulo. Portanto, pela Eq. (3.52),  $\|\mathbf{h}(\mathbf{x})\| > 0$ , então  $c_0$  pode ser tomado igual a  $(\nabla f(\mathbf{x}) \mathbf{d}_x^2) / (\nabla \Psi(\mathbf{h}(\mathbf{x})) \mathbf{h}(\mathbf{x}))$ .  $\square$

**Nota:** O teorema anterior mostra que a direção de Newton pode ser decomposta em duas direções:  $\mathbf{d}_x^1$  que pode ser chamada de direção de otimalidade, pois tenta anular o gradiente da função lagrangiana, e  $\mathbf{d}_x^2$  que pode ser chamada de direção de viabilidade, pois tenta anular a função  $\mathbf{h}$ . O teorema mostra que a direção de otimalidade  $\mathbf{d}_x^1$  é de descida tanto para a função objetivo  $f$  quanto para a função potencial  $\phi_c$  para qualquer valor de  $c$ , sempre que  $\|\mathbf{d}_x^1\| > 0$  e a matriz  $\mathbf{M}$  for positiva definida no espaço tangente. A direção  $\mathbf{d}_x^2$  não é sempre de descida para a função objetivo como mostra a Figura 3.1. Nesta figura, a direção  $\mathbf{d}_x^2$  é de descida no ponto  $\mathbf{x}_l$  mas não é no ponto  $\mathbf{x}_k$ . No ponto  $\mathbf{x}_k$ , o requerimento de viabilidade exige que o valor da função objetivo aumente. Porém, visto que a direção  $\mathbf{d}_x^2$  tenta anular a função  $\mathbf{h}$ , ela é de descida para a função potencial  $\phi_c$  para um valor suficientemente grande do parâmetro  $c$ . A Figura 3.1 mostra também a importância do fato que a direção  $\mathbf{d}_x^1$  seja de descida para a função objetivo. No caso contrário, se a direção de otimalidade for a representada na figura por  $\mathbf{d}_x^{1'}$ , a direção de Newton seria a representada por  $\mathbf{d}_x^{1'}$  e tanto partindo do ponto  $\mathbf{x}_l$  quanto do ponto  $\mathbf{x}_k$  o algoritmo poderia convergir para o máximo local em vez do mínimo local.

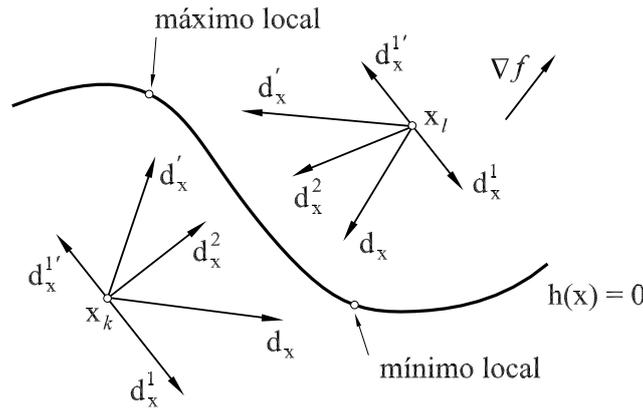


Figura 3.1: Direção do algoritmo de Newton.

### 3.3.2 Modificação do sistema linear

No caso onde a matriz  $M$  do Problema (3.44) não seja positiva definida no espaço tangente, a direção obtida como solução do sistema linear da Eq. (3.50) poderá não ser de descida para a função potencial. Para obter uma direção de descida, o sistema linear da Eq. (3.50) deve ser modificado. A forma mais simples de fazer isso é adicionando um múltiplo da matriz identidade à matriz  $B$ :

$$\begin{pmatrix} \mathbf{B} + \gamma \mathbf{I} & -\nabla \mathbf{g}(\mathbf{x})^T & -\nabla \mathbf{h}(\mathbf{x})^T \\ -\Lambda \nabla \mathbf{g}(\mathbf{x}) & -\mathbf{G}(\mathbf{x}) & 0 \\ -\nabla \mathbf{h}(\mathbf{x}) & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x \\ \bar{\lambda} \\ \bar{\mu} \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x})^T \\ 0 \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}.$$

Este procedimento é uma generalização ao proposto por Goldfeld et al. [47] para problemas sem restrições. Em problemas com restrições e utilizando um algoritmo Primal-dual foi utilizado por Vanderbei e Shanno [48]. Esta modificação da matriz  $B$  tem três grandes vantagens:

1. Sempre existe um valor real  $\gamma$  de forma tal que a matriz  $M$  seja positiva definida.
2. Se a matriz original é esparsa, a modificação proposta apenas incrementa o número de termos não nulos, permitindo que a esparsidade seja aproveitada na resolução eficiente dos sistemas lineares.
3. No caso sem restrições, a direção  $\mathbf{d}_x$  calculada é ótima se for considerada a aproximação quadrática da função objetivo do Problema (3.44) e for adicionada a restrição  $\|\mathbf{d}_x\| \leq \epsilon$  para algum valor  $\epsilon$  positivo [8, Teorema 4.3].

Por outro lado, é necessário que o valor de  $\gamma$  seja o menor possível. Considere, por exemplo, o problema seguinte:

$$\begin{aligned} \text{minimize} \quad & f(x, y, z) = \frac{1}{2}y^2 + z^2 - \frac{1}{2}x^2, \\ \text{sujeito a:} \quad & h(x, y, z) = x + 1 = 0. \end{aligned}$$

A solução  $(x^*, y^*, z^*)$  está dada por  $x^* = -1$  para satisfazer a restrição de igualdade e  $y^* = 0$  e  $z^* = 0$ , para minimizar a função objetivo. De acordo com a Eq. (3.50) com  $\mathbf{B}$  igual à matriz hessiana, a direção de Newton para este problema partindo do ponto  $(x_k, y_k, z_k)$  é:

$$\begin{pmatrix} -1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_x \\ d_y \\ d_z \\ \bar{\mu} \end{pmatrix} = \begin{pmatrix} x_k \\ -y_k \\ -2z_k \\ x_k + 1 \end{pmatrix}.$$

Portanto, após a solução do sistema linear tem-se:

$$d_x = -x_k - 1, \quad d_y = -y_k, \quad d_z = -z_k, \quad \bar{\mu} = 1.$$

Considerando o passo  $t = 1$  tem-se  $x^{k+1} = -1$ ,  $y^{k+1} = 0$  e  $z^{k+1} = 0$ , ou seja, tem-se a solução exata do problema. Por outro lado, o multiplicador de Lagrange calculado é também solução. Neste problema, tanto a matriz hessiana quanto a matriz  $\mathbf{M}$  são positivas definidas no espaço tangente e não é preciso modificá-las. Considere o caso em que o valor  $\gamma$  é tomado de forma tal que a matriz hessiana fique positiva definida, para isso deve ser tomado  $\gamma = 1 + \epsilon$  com  $\epsilon > 0$ . O sistema linear fica então:

$$\begin{pmatrix} \epsilon & 0 & 0 & -1 \\ 0 & 2 + \epsilon & 0 & 0 \\ 0 & 0 & 3 + \epsilon & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_x \\ d_y \\ d_z \\ \bar{\mu} \end{pmatrix} = \begin{pmatrix} x_k \\ -y_k \\ -2z_k \\ x_k + 1 \end{pmatrix}, \quad (3.60)$$

que tem como solução:

$$d_x = -x_k - 1, \quad d_y = -\frac{y_k}{2 + \epsilon}, \quad d_z = -\frac{2z_k}{3 + \epsilon}, \quad \bar{\mu} = -(1 + \epsilon)x_k - \epsilon.$$

Para um passo  $t = 1$  tem-se:

$$x_k = -1, \quad y_k = \frac{1 + \epsilon}{2 + \epsilon}y_k, \quad z_k = \frac{1 + \epsilon}{3 + \epsilon}z_k.$$

Portanto, considerando o passo  $t = 1$  em cada iteração, a convergência nas variáveis  $y$  e  $z$  é apenas linear. É claro que considerando uma escolha melhor do passo é possível melhorar a convergência da seqüência. De fato, após uma primeira iteração de passo unitário, a restrição de igualdade é satisfeita e nas seguintes iterações os pontos da seqüência estarão todos no plano  $x = -1$ . Considere a segunda e terceira linhas da Eq. (3.60):

$$\begin{pmatrix} 2 + \epsilon & 0 \\ 0 & 3 + \epsilon \end{pmatrix} \begin{pmatrix} d_y \\ d_z \end{pmatrix} = \begin{pmatrix} -y_k \\ -z_k \end{pmatrix}. \quad (3.61)$$

A Eq. (3.61) define a direção de busca no plano  $x = -1$ . Como a matriz do sistema é diferente da matriz hessiana restrita ao plano, o método sofrerá do mesmo problema que o método do gradiente: Ainda no caso de utilizar uma busca linear exata as iterações farão um ziguezague e a convergência do algoritmo será afetada. De fato, como o problema é quadrático, a análise de convergência do método de Newton modificado [11, pág. 261] mostra que a convergência do método será linear com velocidade igual a  $(\lambda^M - \lambda^m)^2 / (\lambda^M + \lambda^m)^2$  onde  $\lambda^M$  e  $\lambda^m$  são, respectivamente, o maior e o menor valores próprios da matriz  $\mathbf{S}(\epsilon)^{-1} \bar{\mathbf{H}}$  onde  $\mathbf{S}(\epsilon)$  é a matriz da Eq. (3.61) e  $\bar{\mathbf{H}}$  é a matriz hessiana do problema quando restrita ao plano  $x = -1$ .

Fica claro neste exemplo que a matriz hessiana somente deve ser modificada se for necessário, ou seja, no caso em que a matriz  $\mathbf{M}$  não seja positiva definida no plano tangente.

### 3.3.3 Definição do algoritmo FDIPA-H

O pseudocódigo do algoritmo de otimização proposto é o seguinte:

#### Algoritmo FDIPA-H

Parâmetros:  $\epsilon_1 > 0$ ,  $\epsilon_2 > 0$ ,  $\xi \in (0, 1)$ ,  $\eta \in (0, 1)$ ,  $\varphi > 0$  e  $\nu \in (0, 1)$ .

Dados:  $\mathbf{x}_0 \in \Delta^0$ ,  $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$  positivo,  $\boldsymbol{\mu}_0 \in \mathbb{R}^p$ ,  $\mathbf{B}_0 \in \mathbb{R}^{n \times n}$  simétrica,  $\boldsymbol{\omega}^I \in \mathbb{R}^m$  positivo, e  $c_0$  positivo.

**Passo 1:** Teste de convergência

**Passo 2:** Cálculo da direção de busca

2.1 Faça a decomposição  $\text{LDL}^T$  da matriz  $\mathbf{A}$ :

$$\mathbf{A} = \begin{pmatrix} \mathbf{B}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\nabla \mathbf{g}_k & -\mathbf{S}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix}.$$

2.2 Se  $\sigma_{m+p+1}(\mathbf{D}) \geq \epsilon^1$ , faça  $\gamma_k = 0$ . Se não, faça:

$$\gamma_k = \max_{1 \leq i \leq n} \left\{ \max \left( 1.2 \sum_{j \neq i} |(\mathbf{B}_k)_{ij}| - (\mathbf{B}_k)_{ii}, 0 \right) \right\}. \quad (3.62)$$

2.3 Calcule  $(\mathbf{d}_x^\alpha, \boldsymbol{\lambda}^\alpha, \boldsymbol{\mu}^\alpha)$  e  $(\mathbf{d}_x^\beta, \boldsymbol{\lambda}^\beta, \boldsymbol{\mu}^\beta)$  como a solução do sistema:

$$\begin{pmatrix} \mathbf{B}_k + \gamma_k \mathbf{I} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\nabla \mathbf{g}_k & -\mathbf{S}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^1 & \mathbf{d}_x^2 & \mathbf{d}_x^\beta \\ \boldsymbol{\lambda}^1 & \boldsymbol{\lambda}^2 & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^1 & \boldsymbol{\mu}^2 & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 & 0 \\ 0 & 0 & -\boldsymbol{\omega}^D \\ 0 & \mathbf{h}_k & 0 \end{pmatrix}. \quad (3.63)$$

2.4 Defina  $\mathbf{d}_x^\alpha = \mathbf{d}_x^1 + \mathbf{d}_x^2$ ,  $\boldsymbol{\lambda}^\alpha = \boldsymbol{\lambda}^1 + \boldsymbol{\lambda}^2$  e  $\boldsymbol{\mu}^\alpha = \boldsymbol{\mu}^1 + \boldsymbol{\mu}^2$ .

2.5 Se  $\mathbf{d}_x^\alpha = 0$  pare.

2.6 Se  $\mathbf{h}_k \neq 0$ , faça  $c_k = \max(c_k, (\nabla f_k \mathbf{d}_x^2 + \epsilon_2 \|\mathbf{d}_x^2\|^2) / \sum_{i=1}^p |\mathbf{h}_i(\mathbf{x}_k)|)$ .

2.7 Se  $\nabla \phi_{c_k} \mathbf{d}_x^\beta > 0$ , faça:

$$\rho = \min(\varphi \|\mathbf{d}_x^\alpha\|^2, (\xi - 1) \nabla \phi_{c_k} \mathbf{d}_x^\alpha / \nabla \phi_{c_k} \mathbf{d}_x^\beta).$$

Se não, faça:

$$\rho = \varphi \|\mathbf{d}_x^\alpha\|^2.$$

2.8 Calcule a direção de busca  $\mathbf{d}_x$  e o vetor  $\bar{\boldsymbol{\lambda}}$ :

$$\mathbf{d}_x = \mathbf{d}_x^\alpha + \rho \mathbf{d}_x^\beta, \quad \bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^\alpha + \rho \boldsymbol{\lambda}^\beta.$$

### Passo 3: Busca linear

Calcule  $t_k$  como o primeiro número da seqüência  $\{1, \nu, \nu^2, \nu^3 \dots\}$  que satisfaz:

$$\begin{aligned} \phi_{c_k}(\mathbf{x}_k + t_k \mathbf{d}_x) &\leq \phi_{c_k}(\mathbf{x}_k) + t_k \eta \nabla \phi_{c_k} \mathbf{d}_x, \\ \mathbf{g}_i(\mathbf{x}_k + t_k \mathbf{d}_x) &> 0, & \text{se } \bar{\boldsymbol{\lambda}}_i \geq 0 \text{ ou} \\ \mathbf{g}_i(\mathbf{x}_k + t_k \mathbf{d}_x) &\geq \mathbf{g}_i(\mathbf{x}_k), & \text{caso contrário.} \end{aligned}$$

### Passo 4: Atualização

#### 4.1 Faça:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{d}_x,$$

$$c_{k+1} = c_k,$$

e defina novos valores:  $\omega^D > 0$ ,  $\lambda_{k+1} > 0$ , e  $\mu_{k+1}$ .

#### 4.2 Retorne ao Passo 1.

#### Notas:

1. No Passo 2.2, o parâmetro  $\gamma_k$  é definido suficientemente grande para que a inércia da matriz da Eq. (3.63) seja  $\{n, m + p, 0\}$ . No caso em que a inércia da matriz  $\mathbf{A}$  for  $\{n, m + p, 0\}$  e o menor dos valores próprios positivos for maior do que  $\epsilon_1$ ,  $\gamma_k$  é definida igual a zero. No caso contrário  $\gamma_k$  é definida de forma tal que a matriz  $\mathbf{B}_k + \gamma_k$  seja diagonal dominante e, portanto, positiva definida. Desta forma é garantido que a inércia da matriz da Eq. (3.63) seja  $\{n, m + p, 0\}$ . A desvantagem principal desta técnica é a de ter que fazer uma decomposição  $\mathbf{LDL}^T$  adicional nas iterações onde  $\gamma_k$  seja positivo para resolver os sistemas lineares da Eq. (3.63). Porém, a fórmula dada pela Eq. (3.62) tem a vantagem de ser simples de programar e de não modificar a matriz  $\mathbf{B}_k$  se não for necessário. Se as condições suficientes de otimalidade de segunda ordem forem satisfeitas no ponto ótimo do problema, então, de acordo com o Teorema 2.10, a matriz  $\mathbf{B}_k$  poderá ser definida igual à matriz hessiana. Desta forma, como a matriz hessiana será positiva definida no espaço tangente às restrições ativas em uma vizinhança do ponto ótimo, o algoritmo não a modificará nas iterações finais.
2. No Passo 2.6, o valor  $c_k$  é definido suficientemente grande para que a direção de busca seja de descida para a função potencial  $\phi_{c_k}$ . A existência deste valor suficientemente grande é garantida pelo Teorema 3.46. No caso particular da função potencial utilizada, a fórmula do Passo 2.6 é satisfatória de acordo com o apresentado na Seção 3.3.4.
3. O algoritmo FDIPA-H não considera um desvio com relação às restrições de igualdade do problema. Isto é, o vetor positivo  $\omega^E$  do algoritmo FDIPA foi considerado igual a zero para o algoritmo FDIPA-H, como se desprende da Eq. (3.63). Por esta

razão, a condição  $\mathbf{h}(\mathbf{x}) \geq 0$  não é requerida no Passo 3 do algoritmo. Esta modificação não altera a convergência global do algoritmo e tem a vantagem de permitir valores maiores do passo  $t$ . Na resolução dos problemas com um grande número de restrições de igualdade considerados na Seção 3.4 tornou-se evidente que esta modificação melhora a eficiência do algoritmo.

4. A busca linear do Passo 3 é a de Armijo. As buscas lineares de Goldstein ou Wolfe também poderão ser utilizadas, ver nas referências [8, 9, 11, 14].

### 3.3.4 Convergência global do FDIPA-H

Resultados de convergência global similares aos obtidos para o FDIPA na Seção 3.1 podem ser obtidos para o algoritmo FDIPA-H. A diferença fundamental é que agora a matriz  $\mathbf{B}$  pode não ser positiva definida e, portanto, a Suposição 3.7 deve ser substituída pela seguinte:

**Suposição 3.47** *A matriz  $\mathbf{B}$  pertence a um compacto no qual a matriz  $\mathbf{A}$  tem inércia  $\{n, m + p, 0\}$ .*

Note que a matriz  $\mathbf{B}$  da suposição anterior corresponde à matriz modificada  $\mathbf{B}_k + \gamma \mathbf{I}$  definida pelo algoritmo FDIPA-H. Portanto, se a matriz  $\mathbf{B}_k$  é tomada igual à matriz Hessiana do problema, a Suposição 3.47 é satisfeita se as derivadas segundas das funções que definem o problema forem limitadas, visto que a modificação diagonal satisfaz automaticamente a condição na inércia.

Com isso, pelo Teorema 3.44 tem-se que a matriz  $\mathbf{A}$  sempre é não singular e o algoritmo é bem definido. Além disso, também é possível obter um resultado similar ao Lema 3.20. Porém, ele deve ser obtido de forma diferente. Considere a Eq. (3.25):

$$\begin{aligned} \nabla \phi_{\mathbf{c}}(\mathbf{x}) \mathbf{d}_{\mathbf{x}}^{\alpha} &= -\mathbf{d}_{\mathbf{x}}^{\alpha T} \mathbf{B} \mathbf{d}_{\mathbf{x}}^{\alpha} - \boldsymbol{\lambda}^{\alpha T} \mathbf{S}(\mathbf{x}, \boldsymbol{\lambda}) \boldsymbol{\lambda}^{\alpha} - \sum_{i=1}^p (\mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| + \boldsymbol{\mu}_i^{\alpha} \mathbf{h}_i(\mathbf{x})), \\ &= -\mathbf{d}_{\mathbf{x}}^{\alpha T} \mathbf{M} \mathbf{d}_{\mathbf{x}}^{\alpha} - \sum_{i=1}^p (\mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| + \boldsymbol{\mu}_i^{\alpha} \mathbf{h}_i(\mathbf{x})). \end{aligned}$$

A dificuldade surge do fato de a direção  $\mathbf{d}_{\mathbf{x}}^{\alpha}$  não pertencer ao espaço tangente  $\mathcal{T}$  e, portanto, o valor  $\mathbf{d}_{\mathbf{x}}^{\alpha T} \mathbf{M} \mathbf{d}_{\mathbf{x}}^{\alpha}$  pode não ser positivo. Uma outra dificuldade é que  $\mathbf{M}$  não é uma função contínua do ponto  $\mathbf{x}$ . Portanto, para limitar o valor  $\mathbf{d}_{\mathbf{x}}^{\alpha T} \mathbf{M} \mathbf{d}_{\mathbf{x}}^{\alpha}$  deve-se

proceder com cuidado. O seguinte lema é o equivalente ao Lema 3.20 para o algoritmo FDIPA-H:

**Lema 3.48** *Seja  $\sigma_m^1$  o valor mínimo de  $\sigma_1^S(A)$  e  $\sigma_m^2$  um valor positivo. Se o parâmetro  $c$  satisfaz:*

$$c \geq 0, \quad \text{se } \mathbf{h}(\mathbf{x}) = 0, \text{ ou} \quad (3.64)$$

$$c \geq (\nabla f_k \mathbf{d}_x^2 + \epsilon_2 \|\mathbf{d}_x^2\|^2) / \sum_{i=1}^p |\mathbf{h}_i(\mathbf{x}_k)|, \quad \text{caso contrário.} \quad (3.65)$$

então, a direção  $\mathbf{d}_x^\alpha$  satisfaz:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha \leq -\frac{1}{2} \min(\sigma_m^1, \sigma_m^2) \|\mathbf{d}_x^\alpha\|^2. \quad (3.66)$$

Prova: Na Seção 3.3.1 mostrou-se que  $\mathbf{d}_x^\alpha = \mathbf{d}_x^1 + \mathbf{d}_x^2$  onde  $\mathbf{d}_x^1$  satisfaz:

$$\begin{aligned} \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^1 &= \nabla f(\mathbf{x}) \mathbf{d}_x^1, \\ &= -\mathbf{d}_x^{1T} \mathbf{M} \mathbf{d}_x^1. \end{aligned}$$

Pelo Corolário 3.42 tem-se:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^1 \leq -\sigma_m^1 \|\mathbf{d}_x^1\|^2. \quad (3.67)$$

Para a direção  $\mathbf{d}_x^2$ , da Eq. (3.59) tem-se:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 = \nabla f(\mathbf{x}) \mathbf{d}_x^2 - c \sum_{i=1}^p |\mathbf{h}_i(\mathbf{x})|. \quad (3.68)$$

Assim, se  $\mathbf{h}(\mathbf{x}) = 0$ , da Eq. (3.52) tem-se  $\mathbf{d}_x^2 = 0$ , e a Eq. (3.67) prova o resultado. Se  $\mathbf{h}(\mathbf{x}) \neq 0$ , da Eq. (3.65) tem-se:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 \leq -\sigma_m^2 \|\mathbf{d}_x^2\|^2. \quad (3.69)$$

Assim, das Eqs. (3.67) e (3.69), tem-se:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha \leq -(\sigma_m^1 \|\mathbf{d}_x^1\|^2 + \sigma_m^2 \|\mathbf{d}_x^2\|^2),$$

e a Eq. (3.66) surge de considerar  $\min_{\mathbf{u}+\mathbf{v}=\mathbf{w}} \{\alpha \|\mathbf{u}\|^2 + \beta \|\mathbf{v}\|^2\} = \frac{\alpha\beta}{\alpha+\beta} \|\mathbf{w}\|^2$  e da desigualdade  $\frac{\alpha\beta}{\alpha+\beta} \geq \frac{1}{2} \min(\alpha, \beta)$ .  $\square$

Note que o lema anterior é análogo ao Lema 3.20 com o valor  $\frac{1}{2} \min(\sigma_m^1, \sigma_m^2)$  no lugar de  $\sigma_m$ . Note também que a atualização do parâmetro  $c_k$  do Passo 2.6 do algoritmo FDIPA-H satisfaz as Eqs. (3.64)-(3.65).

Note também que nenhum outro lema da Seção 3.1.1 utiliza a Suposição 3.7. Contudo, para que o Teorema 3.26 de convergência global do FDIPA seja válido neste caso, é necessário mostrar que o parâmetro  $c$  é atualizado somente em um número finito de iterações e para isso basta mostrar que existe um valor real acima do qual as condições das Eqs. (3.64)-(3.65) são sempre satisfeitas. Isso é mostrado pelo seguinte lema:

**Lema 3.49** *Se o parâmetro  $c$  satisfaz:*

$$c \geq (\sigma_m^2 \|\mathbf{d}_x^2\| + \|\nabla f(\mathbf{x})\|) \|\mathbf{A}^{-1}\|, \quad (3.70)$$

*então, a direção  $\mathbf{d}_x^\alpha$  satisfaz:*

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^\alpha \leq -\frac{1}{2} \min(\sigma_m^1, \sigma_m^2) \|\mathbf{d}_x^\alpha\|^2. \quad (3.71)$$

Prova: Para a direção  $\mathbf{d}_x^2$ , da Eq. (3.59) tem-se:

$$\begin{aligned} \nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 &= \nabla f(\mathbf{x}) \mathbf{d}_x^2 - c \sum_{i=1}^p |\mathbf{h}_i(\mathbf{x})|, \\ &\leq \|\nabla f(\mathbf{x})\| \|\mathbf{d}_x^2\| - c \|\mathbf{h}_i(\mathbf{x})\|. \end{aligned}$$

Da Eq. (3.52) tem-se  $\|\mathbf{d}_x^2\| \leq \|(\mathbf{d}_x^2, \boldsymbol{\lambda}^2, \boldsymbol{\mu}^2)\| \leq \|\mathbf{A}^{-1}\| \|\mathbf{h}_i(\mathbf{x})\|$ . Assim:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 \leq (\|\nabla f(\mathbf{x})\| - c/\|\mathbf{A}^{-1}\|) \|\mathbf{d}_x^2\|.$$

Portanto, se  $c$  satisfaz a Eq. (3.70), tem-se:

$$\nabla \phi_c(\mathbf{x}) \mathbf{d}_x^2 \leq -\sigma_m^2 \|\mathbf{d}_x^2\|^2, \quad (3.72)$$

e o resultado surge da mesma forma que no Lema 3.48.  $\square$

**Teorema 3.50** *Toda subsequência convergente da seqüência gerada pelo algoritmo FDIPA-H converge para um ponto de KKT do problema de otimização.*

Prova: O resultado obtém-se diretamente utilizando os mesmos argumentos dos Teoremas 3.26 e 3.27.  $\square$

## 3.4 Exemplos

Nesta seção são apresentados alguns dos resultados obtidos com o algoritmo FDIPA-H para uma coleção de problemas de otimização utilizados para avaliação de algoritmos de otimização. Em primeiro lugar foram testados os problemas de Hock e Schittkowski [49], uma coleção de problemas de pequeno tamanho. Em segundo lugar foram testados os problemas do CUTE [50], neste caso somente uma parte da coleção completa, considerando os problemas de maior número de variáveis e/ou restrições.

### 3.4.1 Alguns detalhes da implementação do algoritmo

Na Seção 3.3.4 provou-se a convergência global do algoritmo FDIPA-H sem definir precisamente as regras de atualização dos multiplicadores de Lagrange. As mesmas fórmulas utilizadas no FDIPA e descritas em [14] podem ser utilizadas no FDIPA-H:

$$(\lambda_{k+1})_i = \max\{\epsilon \|\mathbf{d}_x^\alpha\|^2, \lambda_i^\alpha\}, \quad 1 \leq i \leq m, \quad (3.73)$$

$$(\mu_{k+1})_i = \mu_i^\alpha, \quad 1 \leq i \leq p. \quad (3.74)$$

Note que a regra de atualização dada pela Eq. (3.73) pode não satisfazer a Suposição 3.6 no caso em que  $\|\mathbf{d}_x^\alpha\|^2 \rightarrow 0$ . Porém, como visto no Capítulo 3, esta condição implica que as subsequências convergentes da gerada pelo algoritmo convergem para um ponto estacionário do problema e, portanto, essa regra de atualização não compromete a convergência global do algoritmo.

Nenhuma regra especial foi considerada para a definição do vetor  $\omega_D$ . No programa utilizado para resolver os exemplos apresentados nesta seção foi definido na forma:  $\omega_i^D = 1, 1 \leq i \leq m$ .

Para a decomposição  $\text{LDL}^T$  foi utilizada a rotina MA57 [51, 52]. Para reduzir o número de decomposições, a primeira decomposição da matriz dos sistemas lineares é utilizada sempre que a matriz dos sistemas lineares não for singular. No caso em que a matriz  $M(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$  não for positiva definida no espaço tangente, a modificação da matriz hessiana é realizada previamente à primeira decomposição da próxima iteração. Por causa desta iteração e pelo Teorema 2.19, a convergência global não é comprometida por este procedimento. A grande vantagem deste procedimento é a de precisar, na prática, de uma única decomposição  $\text{LDL}^T$  por iteração. Nas iterações onde a matriz  $M(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$  não

seja positiva definida no espaço tangente, a direção  $\mathbf{d}_x^1$  pode não ser de descida para a função objetivo do problema. Neste caso, a direção  $\mathbf{d}_x^1$  é substituída pela direção  $\tilde{\mathbf{d}}_x^1 = -\mathbf{d}_x^1$ . Esta direção tem a vantagem de ser de descida para a função objetivo e de ser uma direção de *curvatura negativa* para a função objetivo da aproximação quadrática do problema de otimização. Isto é, considere a aproximação  $\tilde{f}(\mathbf{x} + t\tilde{\mathbf{d}}_x^1) = f(\mathbf{x}) + t\nabla f(\mathbf{x})\tilde{\mathbf{d}}_x^1 + t^2\tilde{\mathbf{d}}_x^{1T}\mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})\tilde{\mathbf{d}}_x^1$ , tanto o primeiro termo  $t\nabla f(\mathbf{x})\tilde{\mathbf{d}}_x^1$  quanto o segundo  $t^2\tilde{\mathbf{d}}_x^{1T}\mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})\tilde{\mathbf{d}}_x^1$  são negativos, visto que, da Eq. (3.52) com  $\mathbf{B} = \mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ , tem-se:  $\tilde{\mathbf{d}}_x^{1T}\mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})\tilde{\mathbf{d}}_x^1 = \nabla f(\mathbf{x})\tilde{\mathbf{d}}_x^1 - \boldsymbol{\lambda}^{1T}\mathbf{S}(\mathbf{x}, \boldsymbol{\lambda})\boldsymbol{\lambda}^1$ .

### 3.4.2 Resultados

Os problemas de Hock e Schittkowski foram resolvidos utilizando o algoritmo FDIPA, o algoritmo FAIPA e o algoritmo FDIPA-H. O algoritmo FDIPA é o descrito na Seção 2.1.4 com a diferença que nesta seção é utilizada a busca linear de Wolfe [8, 9, 11] que é também utilizada pelos algoritmos FAIPA e FDIPA-H. A matriz  $\mathbf{B}$  dos sistemas lineares do FDIPA é definida por um método quasi-Newton BFGS [14]. O algoritmo FAIPA também utiliza o método quasi-Newton BFGS para definir a matriz  $\mathbf{B}$ . A diferença fundamental com o FDIPA é que em cada iteração o FAIPA resolve um sistema linear adicional para definir um arco sobre o qual realiza a busca linear. Esta modificação é de grande importância quando as restrições são não lineares, pois contribui para evitar o efeito de Maratos. O efeito de Maratos acontece quando a não linearidade das restrições evita que o algoritmo aceite o passo unitário para pontos arbitrariamente próximos da solução [8, 9, 11, 53]. Adicionalmente, o FAIPA realiza um ajuste da escala das restrições. O algoritmo FDIPA-H é o apresentado na Seção 3.3.3 com as modificações apresentadas na Seção 3.4.1. Para todos os problemas de Hock e Schittkowski foi considerada uma tolerância de  $1.0 \times 10^{-8}$  para a norma da função que define as restrições de igualdade e, como critério de parada, o valor  $1.0 \times 10^{-7}$  na norma da direção  $\mathbf{d}_x^\alpha$  ou na norma do gradiente da função lagrangiana.

A Tabela 3.1 apresenta o resumo dos resultados obtidos que são mostrados de forma completa na Tabela A.1 do Apêndice A. Como mostra a Tabela 3.1 o algoritmo FDIPA-H resolve uma quantidade similar de problemas e, em geral, realiza um número menor de iterações. Em relação com o número de iterações, dos 112 problemas teste ele foi o melhor algoritmo em 84 casos, ou seja, para esses problemas ele realizou menos iterações que os

outros dois algoritmos. Em 19 casos o FDIPA-H foi o pior algoritmo, ou seja, ele realizou mais iterações que os outros dois. Note que para os problemas de Hock e Schittkowski o FDIPA-H foi tão robusto quanto o FAIPA e ainda apresentou um número menor de iterações para uma grande parte do conjunto de problemas, ainda sem contar com uma técnica especial para evitar o efeito de Maratos nem ajuste de escala das restrições. Estes resultados sugerem que o algoritmo FDIPA-H deve ser utilizado em vez do FDIPA ou o FAIPA nos casos onde o custo de cálculo da matriz hessiana não é muito elevado.

A Tabela 3.2 apresenta o resumo dos resultados para a coleção de problemas do CUTE. Estes resultados são mostrados de forma mais completa nas Tabelas A.2 e A.3 do Apêndice A. Para estes problemas de otimização de grande porte o FDIPA-H é comparado com dois dos mais conhecidos algoritmos de otimização para este tipo de problemas: O algoritmo SQP KNITRO [54] e o algoritmo Primal-dual LOQO [48]. Estes algoritmos também utilizam derivadas segundas em sua formulação e os resultados obtidos com eles para a lista de problemas testados foram obtidos de [4]. Os algoritmos FDIPA e FAIPA não foram competitivos para esta coleção de problemas de grande porte, principalmente por causa da matriz quasi-Newton cheia, que aumenta consideravelmente o requerimento de memória e o tempo de resolução dos sistemas lineares. As técnicas de memória limitada LM-BFGS implementadas no algoritmo FAIPA também não apresentaram bons resultados. Esta implementação não superou os inconvenientes associados aos requerimentos de memória nos problemas com grande número de restrições. Além disso, o número de iterações requerido para satisfazer o critério de parada do FAIPA LM-BFGS foi usualmente alto. Também não foi competitivo para estes problemas de grande porte o algoritmo SQP SNOPT [33], que também utiliza técnicas quasi-Newton em sua formulação e cujos resultados também são oferecidos em [4]. Para todos os problemas do CUTE foi considerada uma tolerância de  $1.0 \times 10^{-8}$  para a norma da função que define as restrições de igualdade e, como critério de parada, o valor  $1.0 \times 10^{-5}$  na norma da direção  $d_x^\alpha$  ou na norma do gradiente da função lagrangiana. Os resultados da Tabela A.2 em muitos casos permitem inferir que este critério de parada é tão rigoroso quanto ou mais rigoroso ainda que o utilizado para os algoritmos KNITRO e LOQO.

Para esta coleção de 150 problemas teste o algoritmo LOQO aparece em primeiro lugar em quanto à robustez. Ele somente falha em 12 casos. Em segundo lugar aparece o FDIPA-H, que falha em 36 casos, e em terceiro lugar o algoritmo KNITRO, que falha em

57 dos problemas testados. O motivo principal de falha do algoritmo FDIPA-H é o número elevado de iterações que necessita para resolver alguns dos problemas, produzindo a falha por excesso de número de iterações ou tempo de cálculo. Em muitos destes problemas pode-se perceber um efeito similar ao descrito por Maratos [53]. Portanto, para obter um desempenho similar ao algoritmo LOQO, técnicas especiais deverão ser incorporadas no FDIPA-H visando a diminuição da importância deste efeito. Por exemplo, podem ser consideradas para trabalhos futuros as técnicas de busca em arco como a utilizada no algoritmo FAIPA, técnicas de ajuste de escala das restrições, uso de barreiras logarítmicas para evitar a aproximação prematura à fronteira da região viável e técnicas de redução não monótona da função potencial. Outro problema do algoritmo FDIPA-H foi reconhecido: A solução do problema auxiliar no caso de não ter um ponto inicial viável. Este problema auxiliar pode levar ao algoritmo a um ponto inicial muito desfavorável por não considerar a função objetivo ou as restrições de igualdade na sua formulação. Uma implementação não viável que considere um problema equivalente ao original, como realizado pelo algoritmo LOQO [48], deverá ser estudada no futuro.

No relativo à eficiência, o algoritmo FDIPA-H aparece claramente em primeiro lugar, com relação ao número de iterações ele foi o melhor algoritmo em 81 casos, e o pior somente em 40 dos problemas testados. Isto mostra que o FDIPA-H é promissor, e os estudos mencionados anteriormente visando melhorar a robustez do algoritmo poderão vir a fornecer uma ferramenta computacional muito apropriada para problemas de grande porte.

Tabela 3.1: Resumo dos resultados obtidos para os problemas de Hock e Schittkowski.

	Resolveu	Falhou	Foi o melhor*	Foi o pior*
FDIPA	104 (92.9%)	8 (7.1%)	13 (8.7)	71 (51.2)
FAIPA	103 (92.0%)	9 (8.0%)	24 (19.2)	62 (42.2)
FDIPA-H	104 (92.9%)	8 (7.1%)	84 (80.2)	19 (14.7)

(\*) O valor entre parênteses considera 1/2 para o caso de empate duplo e 1/3 para o caso de empate triplo.

Tabela 3.2: Resumo dos resultados obtidos para os problemas do CUTE.

	Resolveu	Falhou	Foi o melhor*	Foi o pior*
KNITRO	93 (62.0%)	57 (38.0%)	31 (27.0)	76 (66.5)
LOQO	138 (92.0%)	12 (8.0%)	39 (37.0)	43 (42.5)
FDIPA-H	114 (76.0%)	36 (24.0%)	81 (76.0)	40 (31.0)
SNOPT	62 (41.3%)	88 (58.7%)	-	-
FDIPA	61 (40.6%)	89 (59.3%)	-	-
FAIPA	59 (39.3%)	91 (60.7%)	-	-

(\*) A comparação compreende os algoritmos KNITRO, LOQO e FDIPA-H.

O valor entre parênteses considera 1/2 para o caso de empate duplo e 1/3 para o caso de empate triplo.

## Capítulo 4

# Otimização usando o MEC e a formulação SAND

O MEC é um método ideal para utilizar nos problemas de otimização de forma, visto que somente necessita da discretização do contorno do domínio. Esta discretização é, usualmente, suficiente tanto para o cálculo das funções que definem o problema de otimização quanto para o cálculo das suas derivadas. Desta forma evita-se a custosa tarefa de discretização do interior do domínio que são obrigatórias quando o MEF ou o MDF são utilizados. Por outro lado, problemas de domínio infinito são difíceis de tratar utilizando o MEF ou o MDF. A discretização de uma região suficientemente grande que confine o problema em questão é computacionalmente custosa e obriga a tratar com um número enorme de variáveis. Porém, estes problemas podem ser facilmente resolvidos utilizando o MEC, sempre que o contorno do domínio for limitado. Por estas razões existem numerosos trabalhos sobre otimização de forma e análise de sensibilidade utilizando o MEC [55–65].

O inconveniente principal do MEC é que as matrizes do método não são esparsas. Como foi comentado na Seção 2.3, a esparsidade das matrizes do MEF é responsável pelo sucesso da técnica SAND na solução de problemas de otimização estrutural. Esta esparsidade produz uma redução importante do custo computacional de três formas diferentes: a mais óbvia é a redução dos requerimentos de memória para armazenamento das próprias matrizes dos sistemas. Em segundo lugar, permite uma redução significativa do tempo de solução quando rotinas específicas para solução de sistemas lineares esparsos são utilizadas. Por último, as matrizes destes sistemas devem ser calculadas. Tratando-se de

problemas de otimização, estas matrizes provêm da análise de sensibilidade do problema. Quanto maior a esparsidade, menor o número de elementos que devem ser calculados e, portanto, menor o tempo de cálculo.

As matrizes cheias do MEC configuram um problema existente já na análise, que é compensado pela redução importante do número de variáveis. A nova dificuldade que aparece na otimização é o cálculo das matrizes de derivadas. Esta última dificuldade é talvez a razão pela qual a grande maioria dos trabalhos apresentados sobre otimização de forma utilizando o MEC não utilizam a formulação SAND. Por outro lado, os algoritmos utilizados requerem somente de derivadas primeiras, como nos trabalhos [55–61], ou não utilizam informação relativa às derivadas como nos trabalhos [58–61]. A maioria dos trabalhos sobre análise de sensibilidade utilizando o MEC são também dedicados somente à formulação implícita e ao cálculo das derivadas primeiras [57, 62–65].

No entanto, neste capítulo é mostrado que a propriedade de separabilidade parcial presente nas formulações que utilizam o MEF está ainda presente nas formulações que utilizam o MEC. Esta propriedade permite o cálculo eficiente das matrizes de derivadas primeiras e também das derivadas segundas que requerem os algoritmos de otimização apropriados para utilizar junto à formulação SAND.

O método para o cálculo eficiente das derivadas das funções da formulação SAND que surgem da aplicação do MEC é ilustrado nas Seções 4.3 e 4.4 através do estudo do problema de otimização de forma na elasticidade linear. Nas Seções 4.3 e 4.4 é descrito um método para análise de sensibilidade do modelo discreto. Este método permite o cálculo das matrizes de derivadas primeiras e segundas em um tempo proporcional ao quadrado do número de elementos  $n_e$  da discretização. Para isso é mostrado que as funções que definem o problema de otimização têm a propriedade de separabilidade parcial. Como a solução dos sistemas lineares do algoritmo de otimização é uma tarefa de complexidade superior, nos problemas de grande número de variáveis o cálculo das derivadas segundas não representará um custo excessivo. Portanto, os algoritmos de otimização que utilizarem derivadas segundas serão os mais convenientes nesses problemas, visto que, em geral, requerem de um menor número de iterações para obter a solução.

A Seção 4.1 define precisamente o conceito de separabilidade parcial. O apresentado nas Seções 4.3 e 4.4 é baseado nos trabalhos [36, 37]. A Seção 4.3 apresenta o método para o caso de otimizar o volume com uma restrição na complacência da estrutura e a

Seção 4.4 apresenta uma comparação com outros problemas de otimização.

## 4.1 Funções parcialmente separáveis

Nesta seção é descrito um método eficiente para o cálculo de matrizes de derivadas primeiras e segundas de funções em  $\mathbb{R}^n$  que têm a propriedade de separabilidade parcial.

**Definição 4.1** *Uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  tem a propriedade de separabilidade parcial se a mesma pode ser expressada como:*

$$f(\mathbf{z}) = \sum_{j=1}^{n_t} f_j(\mathbf{z}),$$

onde cada termo  $f_j(\mathbf{z})$  depende somente de um número pequeno das variáveis que integram o vetor  $\mathbf{z}$  [8].

Em geral, se a função  $f$  não tiver a propriedade de separabilidade parcial, o custo computacional da avaliação de  $\nabla f$  e  $\nabla^2 f$  é, aproximadamente,  $n$  e  $n^2$  vezes, respectivamente, o custo da avaliação da função  $f$ . Em problemas estruturais  $n$  depende da discretização utilizada, da mesma forma que o número de graus de liberdade da estrutura. Por esta razão o custo da avaliação das derivadas segundas pode ser muito grande quando a malha é refinada.

Se a função  $f$  tiver a propriedade de separabilidade parcial, o custo computacional da avaliação de  $f$ ,  $\nabla f$  e das derivadas de ordem superior será independente de  $n$ . Ao invés disso, o custo dependerá do número  $n_t$  dos termos que definem  $f$ . Por exemplo, considere que cada termo  $f_j$  depende de  $m$  variáveis, com  $m \ll n$ . Então, para calcular  $\nabla f$  devem ser calculadas as  $m$  componentes para as  $n_t$  funções  $f_j$ . Assim, para calcular  $\nabla f$  devem ser calculadas  $mn_t$  componentes. De forma análoga, para calcular  $\nabla^2 f$  devem ser calculadas  $m^2 n_t$  componentes. Para  $n_t$  da ordem de  $n$ , este número é consideravelmente menor que as  $n^2$  componentes de  $\nabla^2 f$ .

Considere, por exemplo, a função  $v : \mathbb{R}^{2n} \rightarrow \mathbb{R}$  definida como o volume de um prisma de base definida por  $n_t$  elementos  $\Gamma^{(j)}$  em  $\mathbb{R}^2$ . Sejam  $\xi^i$ ,  $i \in \{1, \dots, n_t\}$ , os vértices da base,  $\xi^{(j)}$ ,  $j \in \{1, \dots, n_t\}$ , o vetor das coordenadas dos nós do elemento  $\Gamma^{(j)}$ , e  $\xi \in \mathbb{R}^{2n_t}$

o vetor das coordenadas de todos os nós. Para uma altura unitária a expressão de  $v$  é:

$$v(\boldsymbol{\xi}) = \frac{1}{2} \sum_{j=1}^{n_t} \boldsymbol{\xi}^{(j)T} \mathbf{A} \boldsymbol{\xi}^{(j)}, \quad \text{com } A = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Portanto, a função  $v$  tem a propriedade de separabilidade parcial com os termos  $v_j$  definidos como  $v_j(\boldsymbol{\xi}) = \frac{1}{2} \boldsymbol{\xi}^{(j)T} \mathbf{A} \boldsymbol{\xi}^{(j)}$ . Usando a convenção  $\nabla f(\mathbf{z})(i, j) = \frac{\partial f_i}{\partial z_j}(\mathbf{z})$  para uma função vetorial e  $\nabla^2 f(\mathbf{z})(i, j) = \frac{\partial^2 f}{\partial z_i \partial z_j}(\mathbf{z})$  para uma função escalar, o seguinte pseudo-código utiliza a propriedade de separabilidade parcial de  $v$  para calcular as suas matrizes de derivadas primeiras e segundas:

**Passo 1:** Inicialização

$$\begin{aligned} v &= 0, \\ \nabla_{\boldsymbol{\xi}} v &= 0, \\ \nabla_{\boldsymbol{\xi}\boldsymbol{\xi}}^2 v &= 0. \end{aligned}$$

**Passo 2:** Montagem

Para  $j = 1 : n_t$

$$\begin{aligned} \mathbf{v}_u &:= \text{índices do elemento } \Gamma^{(j)}, \\ v &= v + \frac{1}{2} \boldsymbol{\xi}^{(j)T} \mathbf{A} \boldsymbol{\xi}^{(j)}, \\ \nabla_{\boldsymbol{\xi}} v(\mathbf{v}_u) &= \nabla_{\boldsymbol{\xi}} v(\mathbf{v}_u) + \boldsymbol{\xi}^{(j)T} \mathbf{A}, \\ \nabla_{\boldsymbol{\xi}\boldsymbol{\xi}}^2 v(\mathbf{v}_u, \mathbf{v}_u) &= \nabla_{\boldsymbol{\xi}\boldsymbol{\xi}}^2 v(\mathbf{v}_u, \mathbf{v}_u) + \mathbf{A}. \end{aligned}$$

Fim

O pseudo-código anterior mostra que o custo computacional para o cálculo de  $v$ ,  $\nabla v$  é  $\nabla^2 v$  é proporcional ao número de elementos  $n_t$ . De fato, para este exemplo simples tem-se  $m = 4$  e, considerando as linhas finais tem-se, tanto para  $f$  quanto para  $\nabla f$  e  $\nabla^2 f$  um total da ordem de  $m^2 n_t$  operações aritméticas.

## 4.2 Detalhes da implementação do MEC

O programa baseado no MEC utilizado nesta seção emprega nós duplos para modelar situações em que a normal ao contorno não é definida em um ponto, ou é dada uma

condição de contorno descontínua nas forças de superfície, ver na Seção 2.2. Os nós duplos têm associadas as mesmas variáveis de deslocamento para impor a continuidade da função que aproxima a solução do problema e para reduzir o número de variáveis. No caso em que a mesma condição de contorno de deslocamentos é estabelecida para ambos os nós, as variáveis de deslocamentos são consideradas diferentes. Neste caso a condição de contorno garante a continuidade da solução aproximada. Para ter um sistema de equações lineares com matriz definida é necessário, neste caso, agregar equações associadas a pontos no interior dos elementos adjacentes ao nó duplo.

Em qualquer situação, independentemente da presença de nós duplos ou do tipo de condição de contorno, o número de equações necessárias será igual ao número de variáveis de deslocamento.

Cada nó contribui com duas equações integrais associadas às duas soluções fundamentais para esse nó, a solução fundamental correspondente a uma carga unitária pontual horizontal e a correspondente a uma carga pontual vertical.

A montagem das equações integrais de nós nas matrizes  $\mathbf{H}^n$  e  $\mathbf{G}^n$  pode ser representada como:

$$\mathbf{H}^n = \sum_{i=1}^{n_n} \sum_{j=1}^{n_e} \mathcal{H}^n(\mathbf{h}_{ij}), \quad (4.1)$$

$$\mathbf{G}^n = \sum_{i=1}^{n_n} \sum_{j=1}^{n_e} \mathcal{G}^n(\mathbf{g}_{ij}), \quad (4.2)$$

onde  $n_n$  é o número de nós,  $n_e$  é o número de elementos, e os operadores  $\mathcal{H}^n$  e  $\mathcal{G}^n$  colocam os elementos das matrizes  $\mathbf{h}_{ij}$  e  $\mathbf{g}_{ij}$  em matrizes nulas de dimensões iguais a  $\mathbf{H}^n$  e  $\mathbf{G}^n$  respectivamente. Os elementos das linhas 1 e 2 das matrizes  $\mathbf{h}_{ij}$  e  $\mathbf{g}_{ij}$  são levados às linhas correspondentes às variáveis  $\mathbf{u}_1^i$  e  $\mathbf{u}_2^i$ . Os elementos das colunas 1, 2, 3 e 4 da matriz  $\mathbf{h}_{ij}$  são levados às colunas correspondentes às variáveis  $\mathbf{u}_1^{(j)}$ ,  $\mathbf{u}_2^{(j)}$ ,  $\mathbf{u}_3^{(j)}$  e  $\mathbf{u}_4^{(j)}$ . Os mesmos elementos da matriz  $\mathbf{g}_{ij}$  são levados às colunas correspondentes às variáveis  $\mathbf{p}_1^{(j)}$ ,  $\mathbf{p}_2^{(j)}$ ,  $\mathbf{p}_3^{(j)}$  e  $\mathbf{p}_4^{(j)}$ .

No caso dos elementos da matriz  $\mathbf{h}_{ij}$ , os que estão nas colunas 1 e 2 são trocados de sinal e levados às colunas correspondentes às variáveis  $\mathbf{u}_1^i$  e  $\mathbf{u}_2^i$ , os que estão nas colunas 3 e 4 são também trocados de sinal e levados às mesmas colunas. Esta última operação corresponde aos elementos definidos indiretamente pelos deslocamentos rígidos, Eq. (2.49).

Quando é necessário utilizar nós duplos com condições de contorno de deslocamentos segundo a mesma direção é necessário agregar equações integrais obtidas para pontos

fontes no interior dos elementos adjacentes a estes nós duplos. Estas equações integrais são associadas às variáveis de deslocamento do nó duplo adjacente ao elemento.

Neste caso, a montagem das equações integrais de elementos nas matrizes  $\mathbf{H}^e$  e  $\mathbf{G}^e$  pode ser representada como:

$$\mathbf{H}^e = \sum_{i \in I} \sum_{j=1}^{n_e} \mathcal{H}^e(\mathbf{h}_{ij}), \quad (4.3)$$

$$\mathbf{G}^e = \sum_{i \in I} \sum_{j=1}^{n_e} \mathcal{G}^e(\mathbf{g}_{ij}), \quad (4.4)$$

onde  $I$  é o conjunto de índices dos elementos adjacentes aos nós duplos.

Os operadores  $\mathcal{H}^e$  e  $\mathcal{G}^e$  são similares aos operadores  $\mathcal{H}^n$  e  $\mathcal{G}^n$  descritos na seção anterior com a diferença que neste caso não é necessário definir elementos por deslocamentos rígidos.

As matrizes  $\mathbf{H}$  e  $\mathbf{G}$  são definidas como:

$$\mathbf{H} = \mathbf{H}^n + \mathbf{H}^e, \quad (4.5)$$

$$\mathbf{G} = \mathbf{G}^n + \mathbf{G}^e. \quad (4.6)$$

### 4.3 Minimização do volume com restrição na complacência

Considere o seguinte problema de otimização:

$$\begin{aligned} & \text{minimize} \quad v(\boldsymbol{\xi}(\mathbf{q})), \\ & \text{sujeito a:} \quad \mathbf{h}(\boldsymbol{\xi}(\mathbf{q}), \mathbf{u}, \mathbf{p}) = 0, \\ & \quad \mathbf{u}_i = \bar{\mathbf{u}}_i, \quad i \in I_{\mathbf{u}}, \\ & \quad \mathbf{p}_i = \bar{\mathbf{p}}_i, \quad i \in I_{\mathbf{p}}, \\ & \quad g(\mathbf{u}) \geq 0. \end{aligned} \quad (4.7)$$

Aqui,  $\mathbf{q}$  é o vetor das variáveis que definem a geometria do contorno. Para descrever esta geometria, curvas B-spline são utilizadas [26, 66]. Estas curvas aparecem representadas na Eq. (4.7) pela função  $\boldsymbol{\xi}(\mathbf{q})$  que define as coordenadas dos nós da malha dado o vetor de parâmetros  $\mathbf{q}$ . Os vetores  $\mathbf{u}$  e  $\mathbf{p}$  definem os deslocamentos e forças de superfície no contorno, e os conjuntos de índices  $I_{\mathbf{u}}$  e  $I_{\mathbf{p}}$  definem, respectivamente, os nós onde são

conhecidos os deslocamentos ou as forças de superfície, sujeito ao tipo de condição de contorno.

A função  $v$  representa o volume da estrutura. Para o problema bidimensional, onde  $\Omega$  é o domínio ocupado pelo corpo elástico e  $\Gamma$  é o seu contorno, o mesmo é definido como:

$$v(\Omega) = \int_{\Omega} d\Omega. \quad (4.8)$$

A equação  $\mathbf{h}(\boldsymbol{\xi}(\mathbf{q}), \mathbf{u}, \mathbf{p}) = 0$ , aqui chamada de equação de equilíbrio, deriva da equação integral do MEC como foi apresentado na Seção 2.2. A função  $\mathbf{h}$  é definida como:

$$\mathbf{h}(\boldsymbol{\xi}, \mathbf{u}, \mathbf{p}) = \mathbf{H}(\boldsymbol{\xi})\mathbf{u} - \mathbf{G}(\boldsymbol{\xi})\mathbf{p}, \quad (4.9)$$

Por último, a restrição de desigualdade do problema de otimização é uma restrição na complacência da estrutura. Esta restrição prescreve um valor máximo ao trabalho realizado pelas forças externas. A função  $g$  é definida como:

$$g(\mathbf{u}) = \bar{c} - c(\mathbf{u}), \quad (4.10)$$

onde  $\bar{c}$  é um dado do problema, por exemplo, a complacência da estrutura inicial, e a função  $c$  é a complacência da estrutura, ou seja, o trabalho das forças externas. No caso em que as únicas forças externas são as forças de superfície  $\bar{\mathbf{p}}$ , tem-se:

$$c(\mathbf{u}) = \int_{\Gamma_p} \mathbf{u}(x) \cdot \bar{\mathbf{p}}(x) d\Gamma \quad (4.11)$$

Note que na Eq. (4.11) é assumido que a região do contorno onde as forças de superfície  $\bar{\mathbf{p}}$  estão aplicadas não depende das variáveis de forma  $e$ , portanto, a função  $c$  somente depende somente das variáveis do vetor  $\mathbf{u}$ .

No seguinte será mostrado que as funções  $v$ ,  $h$  e  $g$  que definem o problema de otimização têm a propriedade de separabilidade parcial.

Para a função  $v$  tem-se:

$$\begin{aligned} v(\Omega) &= \int_{\Omega} d\Omega, \\ &= \int_{\Gamma} \mathbf{w} \cdot \mathbf{n} d\Gamma, \end{aligned}$$

onde o campo vetorial  $\mathbf{w}$  pode ser definido como  $\mathbf{w}(\boldsymbol{\xi}) = \frac{1}{2}(\boldsymbol{\xi} - \mathbf{o})$  com o ponto  $\mathbf{o}$  fixo. Discretizando o contorno  $\Gamma$  em  $n_e$  elementos lineares  $\Gamma^{(j)}$ , e considerando a representação

paramétrica de  $\Gamma^{(j)}$  dada pela Eq. (2.36), tem-se:

$$v(\boldsymbol{\xi}) = \frac{1}{2} \sum_{j=1}^{n_e} \boldsymbol{\xi}^{(j)T} \mathbf{A} \boldsymbol{\xi}^{(j)}, \quad \text{com } A = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (4.12)$$

onde  $\boldsymbol{\xi}$  é o vetor de todas as coordenadas nodais da discretização e  $\boldsymbol{\xi}^{(j)}$  é o vetor das coordenadas nodais do elemento  $j$ , como definido na Eq. (2.36).

Portanto, a função  $v$  tem a propriedade de separabilidade parcial onde os termos  $v_j(\boldsymbol{\xi})$  são definidos como:  $v_j(\boldsymbol{\xi}) = \frac{1}{2} \boldsymbol{\xi}^{(j)T} \mathbf{A} \boldsymbol{\xi}^{(j)}$ . O custo do cálculo da função  $v$  e de suas primeiras e segundas derivadas é proporcional ao número  $n_e$  de elementos da discretização. As derivadas de  $v$  com relação às variáveis  $\mathbf{q}$  podem ser facilmente obtidas aplicando a regra da cadeia.

Considere agora a equação de equilíbrio. Esta é definida como:

$$\mathbf{h}(\boldsymbol{\xi}, \mathbf{u}, \mathbf{p}) = \mathbf{H}(\boldsymbol{\xi})\mathbf{u} - \mathbf{G}(\boldsymbol{\xi})\mathbf{p} = 0, \quad (4.13)$$

onde as matrizes  $\mathbf{H}$  e  $\mathbf{G}$  são definidas pelas Eqs. (4.5)-(4.6). No problema de otimização da Eq. (4.7) a função  $\mathbf{h}$  define  $2n_e$  restrições de igualdade que têm associados  $2n_e$  multiplicadores de Lagrange. Seja  $\boldsymbol{\mu}_h$  o vetor dos multiplicadores de Lagrange associados a estas restrições. A contribuição de  $\mathbf{h}$  à função lagrangiana do problema de otimização é:

$$\begin{aligned} L_h(\boldsymbol{\xi}, \mathbf{u}, \mathbf{p}, \boldsymbol{\mu}_h) &= \boldsymbol{\mu}_h^T \mathbf{h}(\boldsymbol{\xi}, \mathbf{u}, \mathbf{p}), \\ &= \boldsymbol{\mu}_h^T \mathbf{H}(\boldsymbol{\xi})\mathbf{u} - \boldsymbol{\mu}_h^T \mathbf{G}(\boldsymbol{\xi})\mathbf{p}. \end{aligned} \quad (4.14)$$

Considere o caso em que as equações de equilíbrio são todas resultantes da aplicação da função fundamental nos nós do problema. Neste caso  $\mathbf{H} = \mathbf{H}^n$  e  $\mathbf{G} = \mathbf{G}^n$ . Portanto, das Eqs. (4.1)-(4.2) tem-se:

$$\begin{aligned} L_h(\boldsymbol{\xi}, \mathbf{u}, \mathbf{p}, \boldsymbol{\mu}_h) &= \boldsymbol{\mu}_h^T \left[ \sum_{i=1}^{n_n} \sum_{j=1}^{n_e} \mathcal{H}^n(\mathbf{h}_{ij}) \right] \mathbf{u} - \boldsymbol{\mu}_h^T \left[ \sum_{i=1}^{n_n} \sum_{j=1}^{n_e} \mathcal{G}^n(\mathbf{g}_{ij}) \right] \mathbf{p}, \\ &= \sum_{i=1}^{n_e} \sum_{j=1}^{n_e} [(\boldsymbol{\mu}_h^i)^T \mathbf{h}_{ij}(\boldsymbol{\xi})\mathbf{u}^{(j)} - (\boldsymbol{\mu}_h^i)^T \mathbf{g}_{ij}(\boldsymbol{\xi})\mathbf{p}^{(j)}], \end{aligned} \quad (4.15)$$

onde o vetor  $\boldsymbol{\mu}_h^i$  contém as componentes do vetor  $\boldsymbol{\mu}_h$  que são associadas às equações de equilíbrio correspondentes ao nó  $\xi^i$ . Da Eq. (2.41) é fácil ver que as matrizes  $\mathbf{h}_{ij}$  e  $\mathbf{g}_{ij}$  dependem somente das coordenadas do nó  $\xi^i$  e das coordenadas  $\boldsymbol{\xi}^{(j)}$  do elemento  $\Gamma^{(j)}$ .

A Eq. (4.15) mostra que a função  $L_h$  tem a propriedade de separabilidade parcial, porque cada termo depende somente das variáveis associadas ao nó  $\xi^i$  e ao elemento  $\Gamma^{(j)}$ .  $L_h$  é a soma de  $n_n \times n_e$  termos  $L_h^{ij}$  definidos como:  $L_h^{ij}(\boldsymbol{\xi}, \mathbf{u}, \mathbf{p}, \boldsymbol{\mu}_h) = (\boldsymbol{\mu}_h^i)^T \mathbf{h}_{ij}(\boldsymbol{\xi}) \mathbf{u}^{(j)} - (\boldsymbol{\mu}_h^i)^T \mathbf{g}_{ij}(\boldsymbol{\xi}) \mathbf{p}^{(j)}$ . Cada termo depende das 16 variáveis dadas por  $\xi^i$ ,  $\boldsymbol{\xi}^{(j)}$ ,  $\mathbf{u}^{(j)}$ ,  $\mathbf{p}^{(j)}$  e  $\boldsymbol{\mu}_h^i$ . A Eq. (4.15) mostra também que  $L_h$ ,  $\nabla L_h$  e  $\nabla^2 L_h$  podem ser calculados com um custo computacional proporcional a  $n_n \times n_e$ . As matrizes de derivadas dos termos  $L_h^{ij}$  devem ser calculados considerando a definição das matrizes  $\mathbf{h}_{ij}$  e  $\mathbf{g}_{ij}$ , Eq. (2.41).

Algumas das derivadas de  $L_h$  são ainda mais fáceis de calcular. De fato, o vetor das derivadas primeiras de  $L_h$  é:

$$\nabla L_h = \left( \nabla_{\boldsymbol{\xi}} L_h \quad \nabla_{\mathbf{u}} L_h \quad \nabla_{\mathbf{p}} L_h \quad \nabla_{\boldsymbol{\mu}} L_h \right).$$

onde:

$$\nabla_{\boldsymbol{\xi}} L_h = \boldsymbol{\mu}_h^T \nabla_{\boldsymbol{\xi}} [\mathbf{H}(\boldsymbol{\xi}) \mathbf{u} - \mathbf{G}(\boldsymbol{\xi}) \mathbf{p}],$$

$$\nabla_{\mathbf{u}} L_h = \boldsymbol{\mu}_h^T \mathbf{H}(\boldsymbol{\xi}),$$

$$\nabla_{\mathbf{p}} L_h = -\boldsymbol{\mu}_h^T \mathbf{G}(\boldsymbol{\xi}),$$

$$\nabla_{\boldsymbol{\mu}} L_h = \mathbf{u}^T \mathbf{H}(\boldsymbol{\xi})^T - \mathbf{p}^T \mathbf{H}(\boldsymbol{\xi})^T.$$

A matriz das derivadas segundas de  $L_h$  é:

$$\nabla^2 L_h = \begin{pmatrix} \nabla_{\boldsymbol{\xi}\boldsymbol{\xi}}^2 L_h & \nabla_{\boldsymbol{\xi}\mathbf{u}}^2 L_h & \nabla_{\boldsymbol{\xi}\mathbf{p}}^2 L_h & \nabla_{\boldsymbol{\xi}\boldsymbol{\mu}}^2 L_h \\ \nabla_{\mathbf{u}\boldsymbol{\xi}}^2 L_h & \nabla_{\mathbf{u}\mathbf{u}}^2 L_h & \nabla_{\mathbf{u}\mathbf{p}}^2 L_h & \nabla_{\mathbf{u}\boldsymbol{\mu}}^2 L_h \\ \nabla_{\mathbf{p}\boldsymbol{\xi}}^2 L_h & \nabla_{\mathbf{p}\mathbf{u}}^2 L_h & \nabla_{\mathbf{p}\mathbf{p}}^2 L_h & \nabla_{\mathbf{p}\boldsymbol{\mu}}^2 L_h \\ \nabla_{\boldsymbol{\mu}\boldsymbol{\xi}}^2 L_h & \nabla_{\boldsymbol{\mu}\mathbf{u}}^2 L_h & \nabla_{\boldsymbol{\mu}\mathbf{p}}^2 L_h & \nabla_{\boldsymbol{\mu}\boldsymbol{\mu}}^2 L_h \end{pmatrix},$$

onde:

$$\begin{aligned}
\nabla_{\xi\xi}^2 L_h &= \nabla_{\xi\xi}^2 [\boldsymbol{\mu}_h^T \mathbf{H}(\boldsymbol{\xi}) \mathbf{u} - \boldsymbol{\mu}_h^T \mathbf{G}(\boldsymbol{\xi}) \mathbf{p}], \\
\nabla_{\xi\mathbf{u}}^2 L_h &= (\nabla_{\mathbf{u}\xi}^2 L_h)^T = \nabla_{\xi} [\mathbf{H}(\boldsymbol{\xi})^T \boldsymbol{\mu}_h], \\
\nabla_{\xi\mathbf{p}}^2 L_h &= (\nabla_{\mathbf{p}\xi}^2 L_h)^T = -\nabla_{\xi} [\mathbf{G}(\boldsymbol{\xi})^T \boldsymbol{\mu}_h], \\
\nabla_{\xi\boldsymbol{\mu}}^2 L_h &= (\nabla_{\boldsymbol{\mu}\xi}^2 L_h)^T = \nabla_{\xi} [\mathbf{H}(\boldsymbol{\xi}) \mathbf{u} - \mathbf{G}(\boldsymbol{\xi}) \mathbf{p}], \\
\nabla_{\mathbf{u}\mathbf{u}}^2 L_h &= 0, \\
\nabla_{\mathbf{u}\mathbf{p}}^2 L_h &= (\nabla_{\mathbf{p}\mathbf{u}}^2 L_h)^T = 0, \\
\nabla_{\mathbf{u}\boldsymbol{\mu}}^2 L_h &= (\nabla_{\boldsymbol{\mu}\mathbf{u}}^2 L_h)^T = \mathbf{H}(\boldsymbol{\xi}), \\
\nabla_{\mathbf{p}\mathbf{p}}^2 L_h &= 0, \\
\nabla_{\mathbf{p}\boldsymbol{\mu}}^2 L_h &= (\nabla_{\boldsymbol{\mu}\mathbf{p}}^2 L_h)^T = -\mathbf{G}(\boldsymbol{\xi}), \\
\nabla_{\boldsymbol{\mu}\boldsymbol{\mu}}^2 L_h &= 0.
\end{aligned}$$

Então, para obter  $L_h$  somente é necessário calcular:

- 1)  $\mathbf{H}(\boldsymbol{\xi})$ ,
- 2)  $\mathbf{G}(\boldsymbol{\xi})$ .

(4.16)

Além disso, para obter  $\nabla L_h$  é preciso calcular:

- 3)  $\nabla_{\xi} [\mathbf{H}(\boldsymbol{\xi}) \mathbf{u} - \mathbf{G}(\boldsymbol{\xi}) \mathbf{p}]$ .

(4.17)

Por último, para calcular  $\nabla^2 L_h$  são também necessárias as matrizes:

- 4)  $\nabla_{\xi\xi}^2 [\boldsymbol{\mu}_h^T \mathbf{H}(\boldsymbol{\xi}) \mathbf{u} - \boldsymbol{\mu}_h^T \mathbf{G}(\boldsymbol{\xi}) \mathbf{p}]$ ,
- 5)  $\nabla_{\xi} [\mathbf{H}(\boldsymbol{\xi})^T \boldsymbol{\mu}_h]$ ,
- 6)  $\nabla_{\xi} [\mathbf{G}(\boldsymbol{\xi})^T \boldsymbol{\mu}_h]$ .

(4.18)

As expressões das Eqs. (4.17) e (4.18) são calculadas utilizando o método de integração de Gauss, no caso das integrais regulares, e derivando diretamente as expressões analíticas no caso das integrais singulares, utilizando as expressões das Eqs. (2.45)-(2.48) e (2.53)-(2.54).

Os tempos de CPU requeridos para a montagem de  $L_h$  e de suas matrizes de derivadas são mostrados pela Figura 4.1. Na figura, “Tempo da função” é o tempo para o cálculo da Eq. (4.16), “Derivadas primeiras” é o tempo para calcular Eq. (4.17) e “Derivadas segundas” é o tempo necessário para Eq. (4.18). A última curva é o tempo requerido

para realizar uma análise, ou seja, a solução do sistema linear da Eq. (2.56) (o tempo de análise é sempre menor que o requerido por uma iteração do algoritmo de otimização). Os tempos foram obtidos para um número de elementos crescente considerando um problema de geometria fixa.

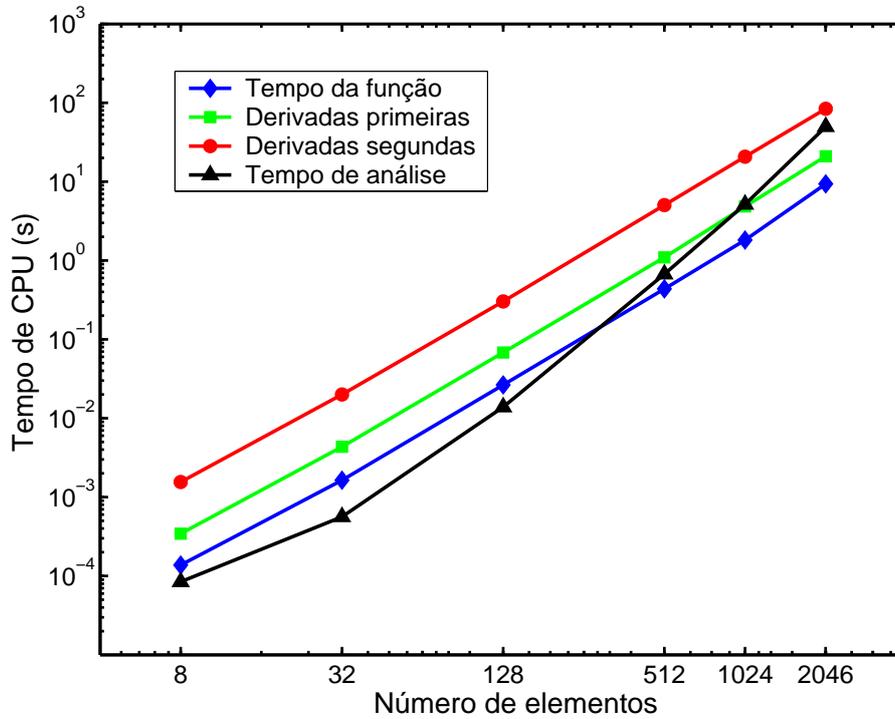


Figura 4.1: Tempo para o cálculo de  $L_h$  e de suas derivadas.

A Figura 4.1 mostra que para problemas com um grande número de elementos a solução dos sistemas de equações lineares é a tarefa de maior custo computacional. Uma redução do custo computacional nesses problemas poderá ser obtida reduzindo o número de sistemas lineares resolvidos, ou seja, através da redução do número de iterações requerido para encontrar a solução. Neste caso os algoritmos que utilizam derivadas segundas serão os mais convenientes porque, em geral, eles requerem de menos iterações para obter a solução do problema de otimização.

Finalmente, considere a complacência da estrutura:

$$\begin{aligned}
c(\mathbf{u}) &= \int_{\Gamma_p} \mathbf{u}(x) \cdot \bar{\mathbf{p}}(x) \, d\Gamma, \\
&= \sum_{j \in J_p} \int_{\Gamma^{(j)}} (\Phi \mathbf{u}^{(j)})^T (\Phi \bar{\mathbf{p}}^{(j)}) \, d\Gamma, \\
&= \sum_{j \in J_p} \mathbf{u}^{(j)T} \mathbf{m}_j \bar{\mathbf{p}}^{(j)}, \tag{4.19}
\end{aligned}$$

onde  $\Phi$  é definido nas Eqs. (2.37) e (2.38),  $J_p$  é o conjunto de índices de elemento onde são prescritas as forças de superfície, e a matriz  $\mathbf{m}_j$  é:

$$\mathbf{m}_j = \int_{\Gamma^{(j)}} \Phi^T \Phi \, d\Gamma. \tag{4.20}$$

A Eq. (4.19) mostra que a função  $c$  tem a propriedade de separabilidade parcial, onde os termos do somatório dependem somente das variáveis do vetor  $\mathbf{u}^{(j)}$ .

Pela Eq. (4.10), a função  $g$ , tal como a função  $c$ , tem também a propriedade de separabilidade parcial.

Até aqui foi mostrado como calcular derivadas das funções com relação às coordenadas, deslocamentos e forças de superfície nodais. No entanto, é conveniente descrever o contorno utilizando curvas spline. Neste trabalho foram utilizadas curvas B-spline cúbicas, como apresentado na referência [66]. As coordenadas nodais, dadas pelo vetor  $\xi$  dependem dos parâmetros das curvas B-spline agrupados no vetor  $\mathbf{q}$ . As derivadas das funções consideradas em relação aos parâmetros das curvas spline podem ser obtidas facilmente aplicando a regra da cadeia.

Algumas das vantagens da descrição do contorno por curvas B-spline são:

- As coordenadas nodais são funções lineares dos parâmetros das B-spline. Esta propriedade facilita o cálculo das derivadas.
- Para qualquer valor do vetor  $\mathbf{q}$  a geometria consiste na união de curvas suaves.

Os exemplos apresentados aqui consistem na otimização de uma geometria previamente obtida utilizando um programa de otimização topológica. O algoritmo de otimização Primal-dual descrito na referência [34] foi utilizado. Para a otimização de topologia foi utilizado o programa proposto por Sigmund [67] e a interface de usuário descrita na referência [68].

Um termo de penalidade foi introduzido para prevenir grandes variações de geometria nas primeiras iterações. À função objetivo  $v$  foi adicionado o termo  $\alpha(\boldsymbol{\xi}(q) - \boldsymbol{\xi}(q_0))^2$  onde  $q_0$  é o vetor de parâmetros de spline da geometria inicial. Este termo de penalidade mostrou ser útil nos problemas onde se produzem grandes variações de geometria e tem a vantagem de ser fácil de calcular, tal como as suas derivadas. O processo de otimização é dividido em etapas, começando com um valor positivo para o parâmetro  $\alpha$  que é reduzido quando as iterações convergem. Na etapa final o valor de  $\alpha$  é nulo, retornando desta forma ao problema original de minimização do volume.

### Exemplo 1

A Figura 4.2 mostra a configuração inicial para a otimização de topologia neste exemplo. O problema consiste em uma placa de dimensões  $2.0 \text{ m} \times 1.0 \text{ m} \times 1.0 \text{ mm}$ , respectivamente de base, altura e espessura. Uma malha de  $40 \times 20$  elementos finitos foi utilizada. A Figura 4.3 mostra o resultado da otimização de topologia e a configuração inicial adotada para a otimização de forma. A cor escura na Figura 4.3.a indica as regiões do domínio onde se tem o material sólido, e a cor clara indica as regiões vazias ou de densidade de material desprezível. A configuração inicial adotada para a otimização de forma corresponde a uma malha de 198 elementos com 47 parâmetros de forma. A Figura 4.4 mostra dois resultados para diferentes valores máximos da complacência. Ambos os resultados foram obtidos após 38 iterações do algoritmo de otimização.

O módulo de Young do material é  $E = 100.0 \text{ GPa}$  e o coeficiente de Poisson é  $\nu = 0.3$ . O valor da força de superfície vertical é de  $1.0 \text{ kN/m}$  aplicada em uma região de  $0.2 \text{ m}$  de comprimento.



Figura 4.2: Exemplo 1, configuração inicial da otimização de topologia.

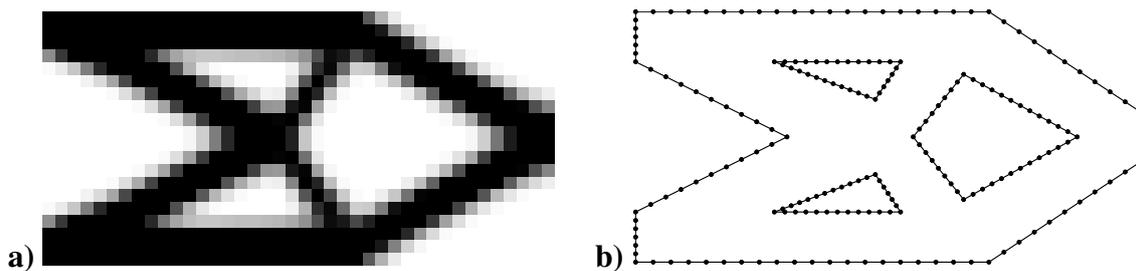


Figura 4.3: Exemplo 1, **a)** Resultado da otimização de topologia para um 50% do volume final, **b)** Configuração inicial para otimização de forma.

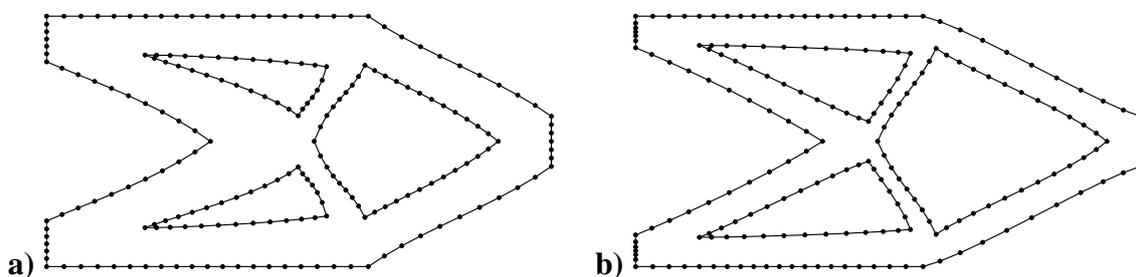


Figura 4.4: Exemplo 1, resultados para a complacência igual a **a)**  $2.22 \times 10^{-2}$  kNm e **b)**  $3.00 \times 10^{-2}$  kNm.

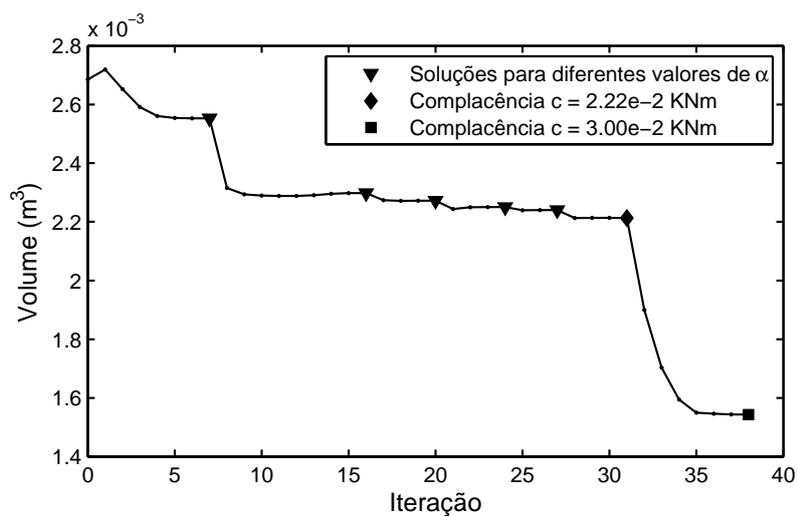


Figura 4.5: Exemplo 1, evolução da função objetivo.

## Exemplo 2

Para este exemplo a configuração inicial para a otimização de topologia é mostrado na Figura 4.6. Uma malha de  $40 \times 20$  elementos finitos foi utilizada. O resultado da otimização de topologia e a configuração inicial adotada para a otimização de forma são mostrados pela Figura 4.7. Esta configuração corresponde a uma malha de 180 elementos com 36 parâmetros de forma. Dois resultados para diferentes valores máximos da complacência são mostrados pela Figura 4.8. Ambos os resultados foram obtidos após 56 iterações do algoritmo de otimização.

As propriedades do material são as mesmas que no Exemplo 1. O valor da força de superfície vertical é de 1.0 kN/m aplicada em regiões de 0.2 m de comprimento.

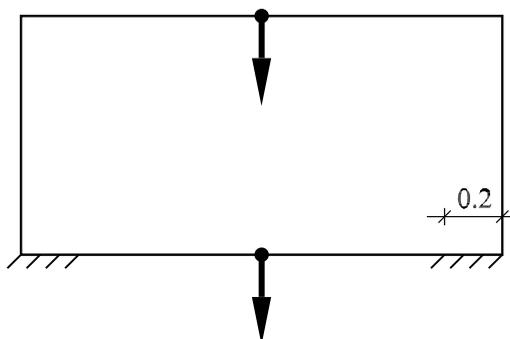


Figura 4.6: Exemplo 2, configuração inicial da otimização de topologia.

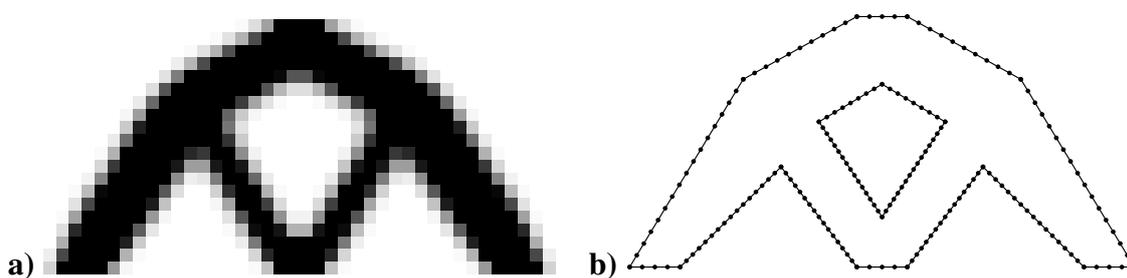


Figura 4.7: Exemplo 2, **a)** Resultado da otimização de topologia para um 40% do volume final, **b)** Configuração inicial para otimização de forma.

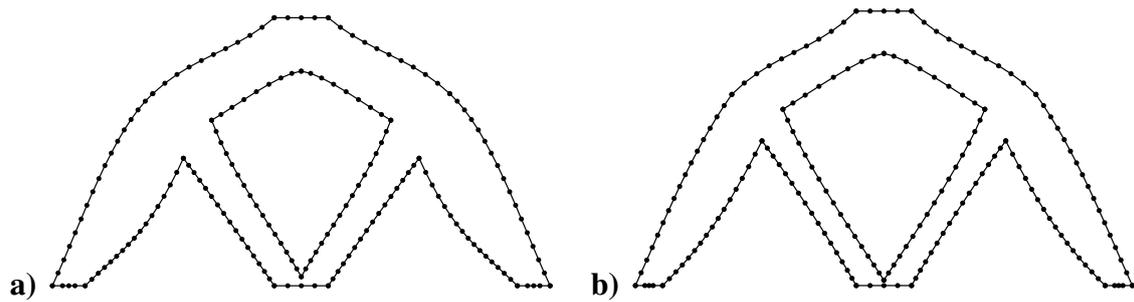


Figura 4.8: Exemplo 2, resultados para a complacência igual a **a)**  $0.80e-2$  kNm e **b)**  $1.00e-2$  kNm.

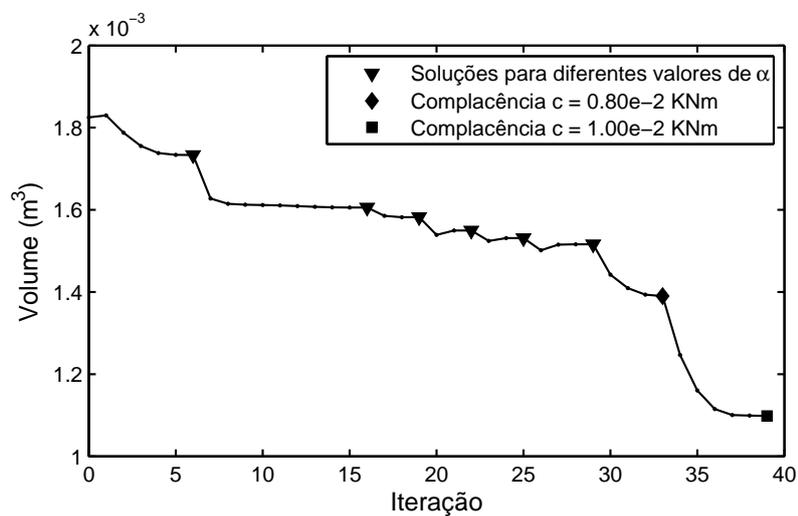


Figura 4.9: Exemplo 2, evolução da função objetivo.

### Exemplo 3

Neste caso a configuração inicial para a otimização de topologia é dada pela Figura 4.10. Uma malha de  $40 \times 20$  elementos finitos foi utilizada. A Figura 4.11 mostra o resultado da otimização de topologia e a configuração inicial adotada para a otimização de forma. Esta configuração corresponde a uma malha de 402 elementos com 82 parâmetros de forma. A Figura 4.12 mostra dois resultados para diferentes valores máximos da complacência. Ambos os resultados foram obtidos após 39 iterações do algoritmo de otimização.

As propriedades do material são as mesmas que no Exemplo 1. O valor da força de

superfície vertical é de 1.0 kN/m aplicada em uma região de 0.2 m de comprimento.

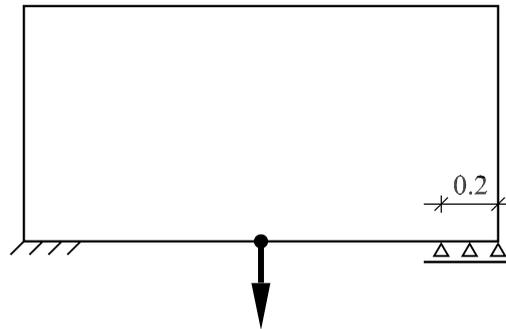


Figura 4.10: Exemplo 3, configuração inicial da otimização de topologia.

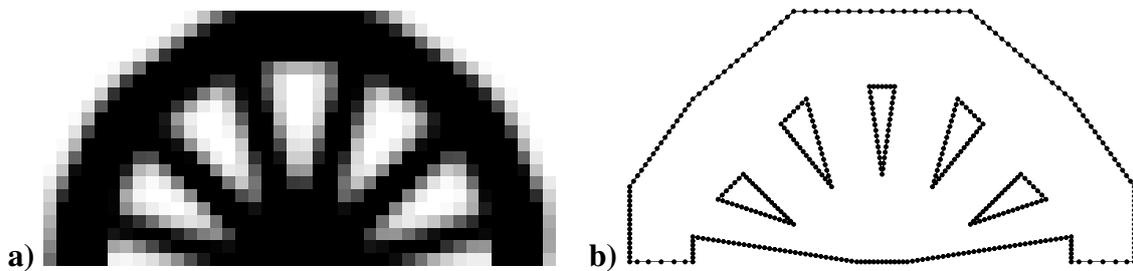


Figura 4.11: Exemplo 3, **a)** Resultado da otimização de topologia para um 60% do volume final, **b)** Configuração inicial para otimização de forma.

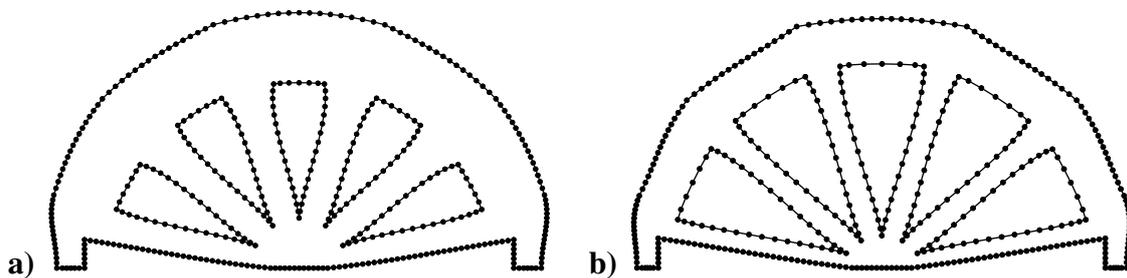


Figura 4.12: Exemplo 3, resultados para a complacência igual a **a)**  $2.52e-2$  kNm e **b)**  $3.60e-2$  kNm.

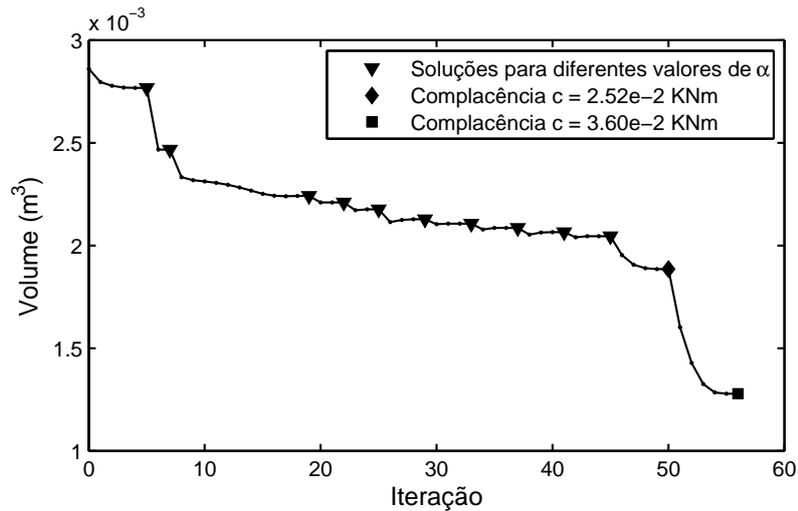


Figura 4.13: Exemplo 3, evolução da função objetivo.

#### Exemplo 4

Neste exemplo a configuração inicial para a otimização de topologia é dada pela Figura 4.14. Uma malha de  $80 \times 20$  elementos finitos foi utilizada. O resultado da otimização de topologia e a configuração inicial adotada para a otimização de forma são mostrados pela Figura 4.15. Esta configuração corresponde uma malha de 525 elementos com 105 parâmetros de forma. O resultado foi obtido após 37 iterações do algoritmo de otimização e é mostrado pela Figura 4.16.

As propriedades do material são as mesmas que no Exemplo 1. O valor da força de superfície vertical é de 1.0 kN/m aplicada em uma região de 0.1 m de comprimento.

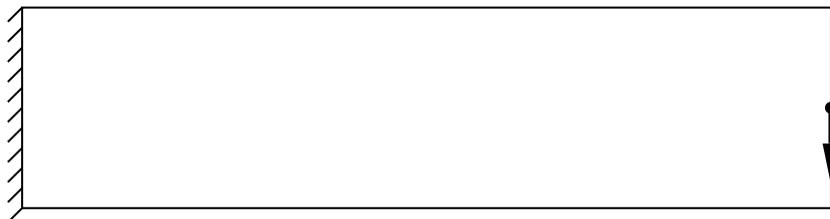


Figura 4.14: Exemplo 4, configuração inicial da otimização de topologia.

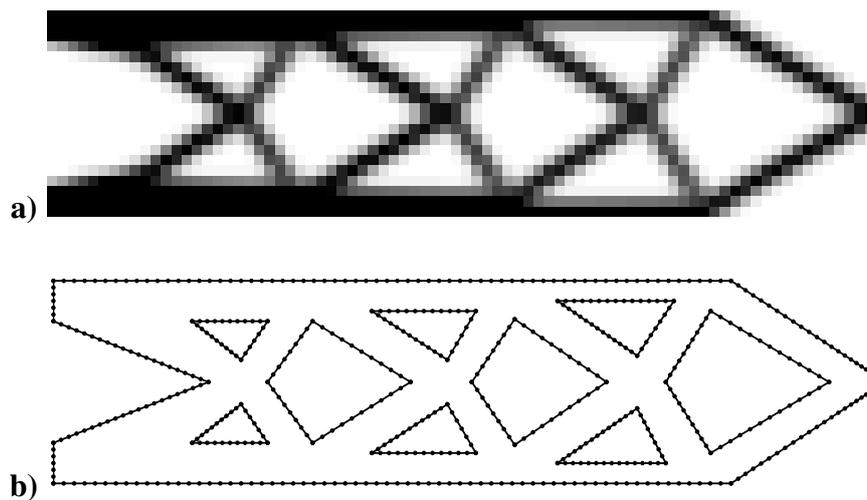


Figura 4.15: Exemplo 4, **a)** Resultado da otimização de topologia para um 40% do volume final, **b)** Configuração inicial para otimização de forma no exemplo 3.

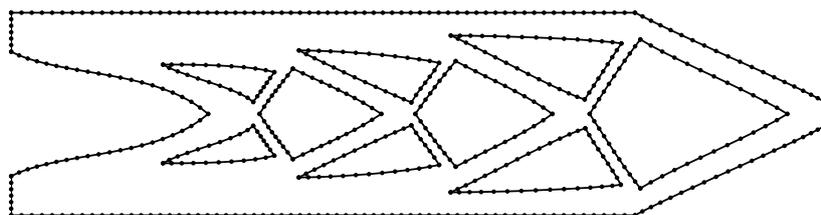


Figura 4.16: Exemplo 4, resultado para a complacência igual a  $3.42e-2$  kNm.

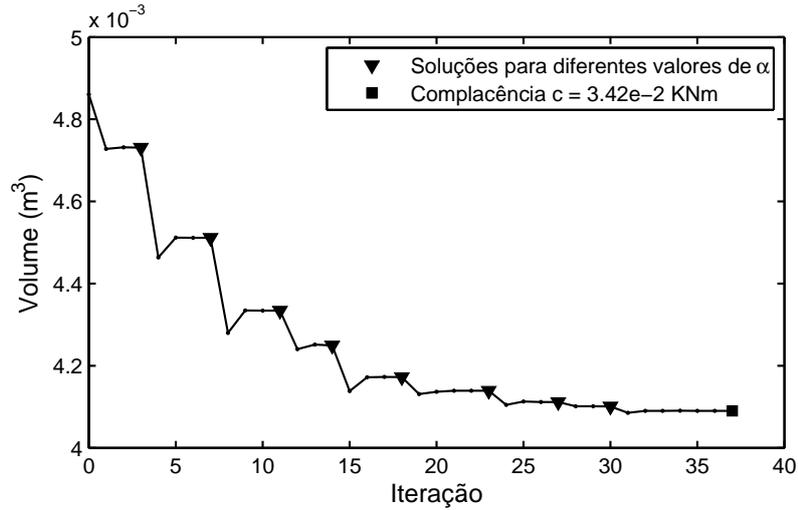


Figura 4.17: Exemplo 4, evolução da função objetivo.

## 4.4 Comparação com outros problemas de otimização

Nesta seção são comparados os resultados obtidos para três problemas diferentes de otimização de forma, dados pelas Eqs. (4.21), (4.22) e (4.23).

### Problema 1

O primeiro problema é equivalente ao considerado na seção anterior, para propósitos de comparação a função objetivo agora é a complacência  $c$  em quanto que o volume é considerado como restrição.

$$\begin{aligned}
 &\text{minimize } c(\mathbf{u}), \\
 &\text{sujeito a: } \mathbf{h}(\boldsymbol{\xi}(\mathbf{q}), \mathbf{u}, \mathbf{p}) = 0, \\
 &\quad \mathbf{u}_i = \bar{\mathbf{u}}_i, \quad i \in I_{\mathbf{u}}, \\
 &\quad \mathbf{p}_i = \bar{\mathbf{p}}_i, \quad i \in I_{\mathbf{p}}, \\
 &\quad v(\boldsymbol{\xi}(\mathbf{q})) \leq v_0,
 \end{aligned} \tag{4.21}$$

onde  $v_0$  é um valor real positivo previamente estabelecido e que define o volume máximo possível da estrutura ótima.

## Problema 2

O segundo problema considerado consiste em minimizar a máxima tensão de von Mises sobre o contorno da estrutura, considerando também a mesma restrição de desigualdade no volume. Este problema pode ser formulado como um problema diferenciável introduzindo uma variável auxiliar  $z$  na forma:

$$\begin{aligned} & \text{minimize } z, \\ & \text{sujeito a: } \mathbf{h}(\boldsymbol{\xi}(\mathbf{q}), \mathbf{u}, \mathbf{p}) = 0, \\ & \quad \mathbf{u}_i = \bar{\mathbf{u}}_i, \quad i \in I_{\mathbf{u}}, \\ & \quad \mathbf{p}_i = \bar{\mathbf{p}}_i, \quad i \in I_{\mathbf{p}}, \\ & \quad z \geq \sigma_j^M(\mathbf{u}, \mathbf{p}), \quad j \in I_M, \\ & \quad v(\boldsymbol{\xi}(\mathbf{q})) \leq v_0, \end{aligned} \tag{4.22}$$

onde  $\sigma_j^M$  representa a tensão de von Mises no ponto  $j$  que, nesta formulação SAND, depende dos vetores  $\mathbf{u}$  e  $\mathbf{p}$  como será mostrado posteriormente. O conjunto  $I_M$  contém os índices dos nós da malha considerados.

## Problema 3

O terceiro problema consiste em minimizar a media quadrática no contorno da diferença entre a tensão de von Mises e um valor de referência como proposto por Herskovits [69]. O Problema 3 considera também uma restrição de desigualdade no volume.

$$\begin{aligned} & \text{minimize } f(\boldsymbol{\xi}(\mathbf{q}), \mathbf{u}, \mathbf{p}) = \int_{\Gamma} (\sigma^M(x) - \sigma^R)^2 d\Gamma, \\ & \text{sujeito a: } \mathbf{h}(\boldsymbol{\xi}(\mathbf{q}), \mathbf{u}, \mathbf{p}) = 0, \\ & \quad \mathbf{u}_i = \bar{\mathbf{u}}_i, \quad i \in I_{\mathbf{u}}, \\ & \quad \mathbf{p}_i = \bar{\mathbf{p}}_i, \quad i \in I_{\mathbf{p}}, \\ & \quad v(\boldsymbol{\xi}(\mathbf{q})) \leq v_0. \end{aligned} \tag{4.23}$$

Na Eq. (4.23), a função  $f$  depende da variável  $\mathbf{q}$  através do domínio de integração  $\Gamma$ . O valor real  $\sigma^R$  é o valor de referência para as tensões de von Mises e pode ser, por exemplo, o valor admissível do material considerado.

Para o cálculo eficiente das matrizes de derivadas resta ver que as tensões de von Mises e a função objetivo do Problema 3 têm a propriedade de separabilidade parcial.

As tensões de von Mises são avaliadas somente no contorno  $\Gamma$  do corpo. Isto é feito expressando os deslocamentos e forças de superfície em coordenadas locais e empregando

a lei de Hooke para determinar os componentes do tensor de tensões. A descrição deste procedimento segue à dada em [15]. Empregando o sistema de coordenadas mostrado pela Figura 4.18, a componente de deformação  $\bar{\epsilon}_{11}$  pode ser obtida como:

$$\bar{\epsilon}_{11} = \frac{\partial \bar{u}_1}{\partial \bar{x}_1}, \quad (4.24)$$

onde a derivada é aplicada à Eq. (2.39) que define os deslocamentos no contorno. Para estado plano de deformação as três componentes requeridas do tensor de tensões são:

$$\bar{\sigma}_{11} = \frac{1}{1-\nu}(\nu \bar{\sigma}_{22} + 2G\bar{\epsilon}_{11}), \quad \bar{\sigma}_{12} = \bar{p}_1, \quad \bar{\sigma}_{22} = \bar{p}_2. \quad (4.25)$$

As mesmas fórmulas são válidas para estado plano de tensões se  $\nu$  é substituído por  $\bar{\nu}$ .

A tensão de von Mises  $\sigma^m$  no elemento pode ser calculada empregando a Eq. (4.26) onde a componente  $\bar{\sigma}_{33}$  deve ser tomada como  $\bar{\sigma}_{33} = \nu(\bar{\sigma}_{11} + \bar{\sigma}_{22})$  em estado plano de deformação e zero no caso de estado plano de tensão.

$$\sigma^m = \sqrt{\bar{\sigma}_{11}^2 + \bar{\sigma}_{22}^2 + \bar{\sigma}_{33}^2 - \bar{\sigma}_{11}\bar{\sigma}_{22} - \bar{\sigma}_{11}\bar{\sigma}_{33} - \bar{\sigma}_{22}\bar{\sigma}_{33} + 3\bar{\sigma}_{12}^2}. \quad (4.26)$$

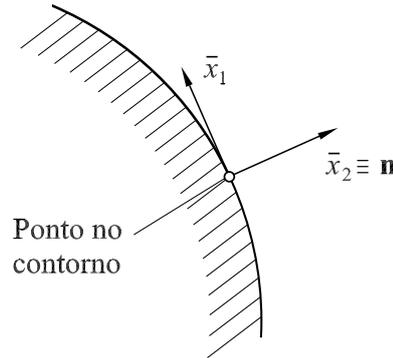


Figura 4.18: Sistema de coordenadas locais em um ponto do contorno.

As Eqs. (4.24), (4.25), e (4.26) mostram que a tensão de von Mises calculada em um ponto do elemento  $\Gamma^{(j)}$  depende das variáveis  $\xi^{(j)}$ ,  $\mathbf{u}^{(j)}$  e  $\mathbf{p}^{(j)}$ . A função objetivo do Problema 3 tem então a propriedade de separabilidade parcial, visto que pode ser expressa como soma de  $n_e$  integrais sobre os elementos. Cada integral depende somente das variáveis associadas ao elemento  $\Gamma^{(j)}$ ,  $\xi^{(j)}$ ,  $\mathbf{u}^{(j)}$  e  $\mathbf{p}^{(j)}$ .

Nesta seção são apresentados três exemplos de otimização de forma em estado plano de tensão. Os três problemas descritos são utilizados para cada exemplo. O algoritmo de otimização utilizado na seção anterior também foi utilizado neste caso. Nos três exemplos o material tem módulo de Young  $E = 1.0 \text{ N/m}^2$  e coeficiente de Poisson  $\nu = 0.3$ .

### Exemplo 1

Este exemplo consiste em encontrar a forma ótima de um orifício em uma placa como mostrado na Figura 4.19. Na figura as medidas estão expressadas em metros. Os valores das forças de superfície são:  $P_1 = 1.5 \text{ N/m}$  e  $P_2 = 1.0 \text{ N/m}$ . Para este exemplo foi utilizada uma malha de 68 elementos e 13 variáveis de forma.

Neste caso os três problemas de otimização forneceram resultados similares. A Figura 4.20 mostra as tensões de von Mises obtidas para um quarto da placa.

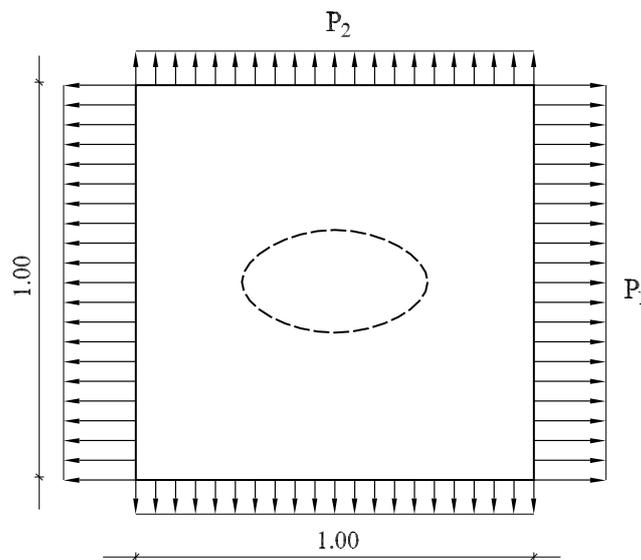


Figura 4.19: Exemplo 1, otimização da forma de um orifício.

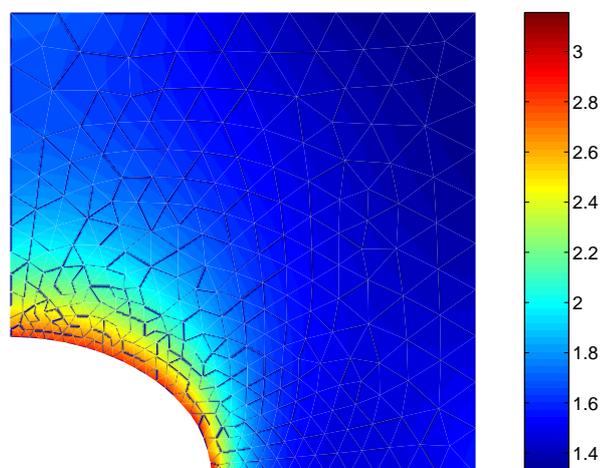


Figura 4.20: Exemplo 1, tensões de von Mises - Problemas 1–3.

## Exemplo 2

Este exemplo consiste em encontrar a forma ótica de um filete como mostrado na Figura 4.21. O filete não pode passar dos limites dados pelas das linhas verticais. As medidas na Figura 4.21 estão expressadas em metros. O valor da força de superfície é:  $P = 1.0 \text{ N/m}$ . Para este exemplo foi utilizada uma malha de 104 elementos e 3 variáveis de forma. Os resultados obtidos são mostrados pelas Figuras 4.22, 4.23 e 4.24 para os três problemas de otimização considerados.

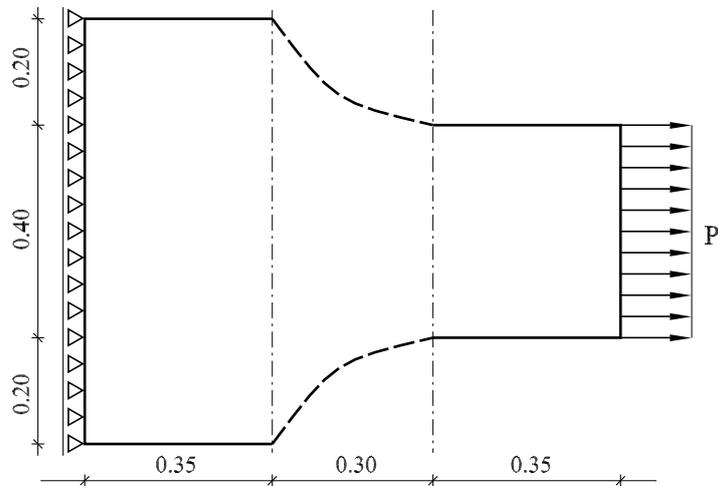


Figura 4.21: Exemplo 2, otimização do filete.

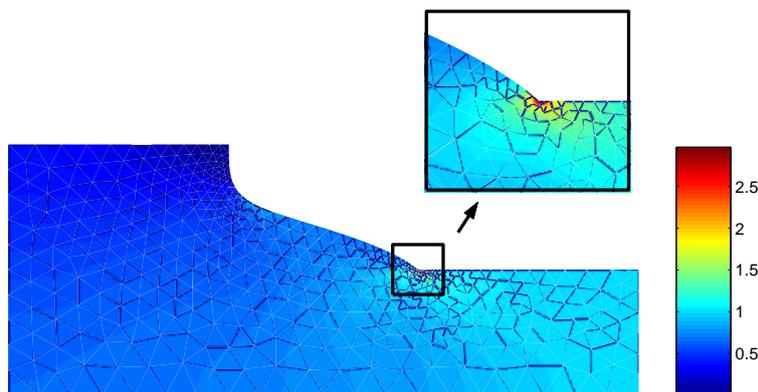


Figura 4.22: Exemplo 2, tensões de von Mises - Problema 1.

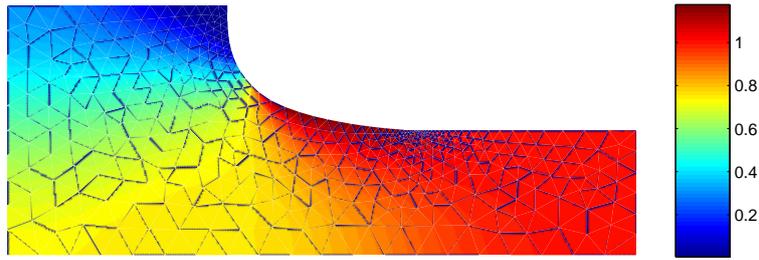


Figura 4.23: Exemplo 2, tensões de von Mises - Problema 2.

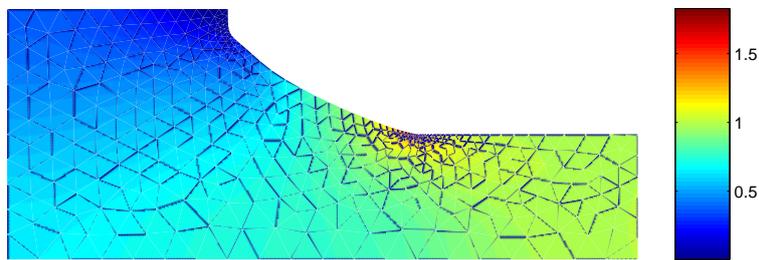


Figura 4.24: Exemplo 2, tensões de von Mises - Problema 3.

### Exemplo 3

Este último exemplo é descrito pela Figura 4.25 onde as medidas estão expressadas em metros. Neste caso é procurada a geometria de um pórtico. Os pontos A, B, C e D da Figura 4.25 são fixos em quanto que os pontos E e F podem ser movidos horizontalmente. Para este exemplo foi utilizada uma malha de 144 elementos e 16 variáveis de forma.

Os resultados obtidos para a metade do pórtico são mostrados pelas Figuras 4.26, 4.27 e 4.28.

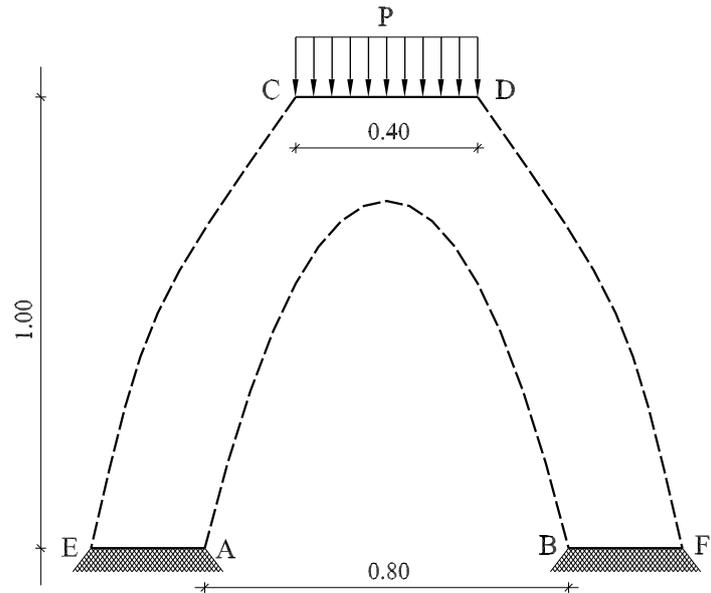


Figura 4.25: Exemplo 3, otimização de um pórtico plano.

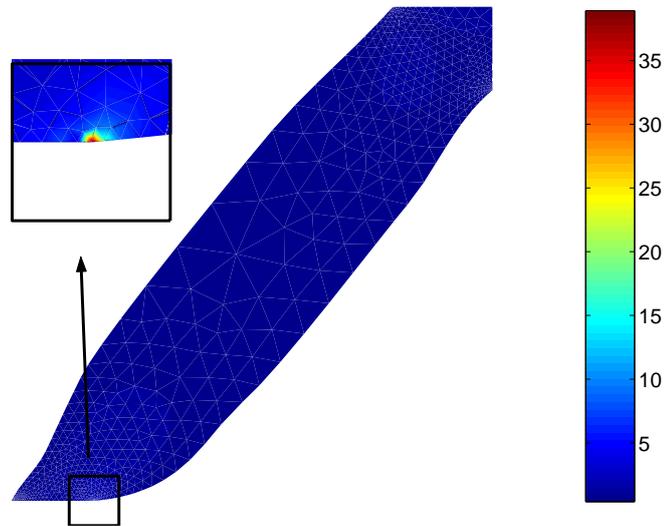


Figura 4.26: Exemplo 3, tensões de von Mises - Problema 1.

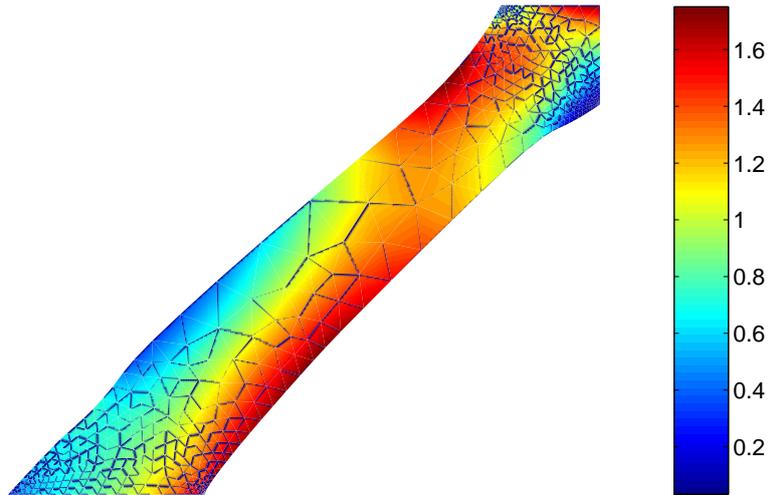


Figura 4.27: Exemplo 3, tensões de von Mises - Problema 2.

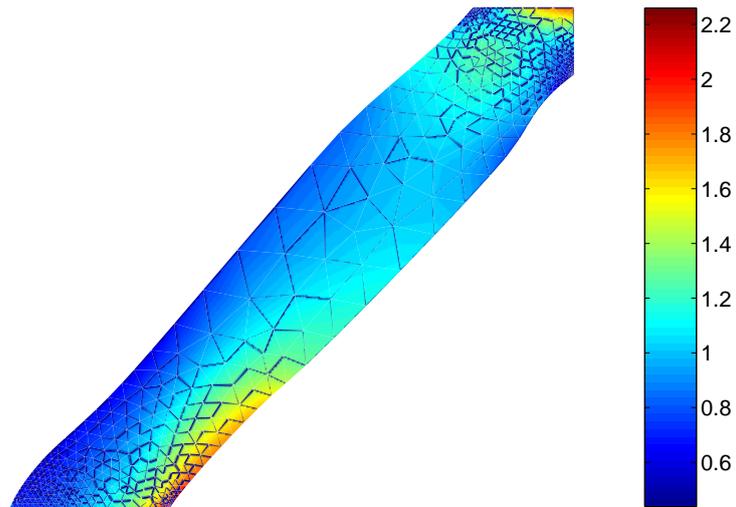


Figura 4.28: Exemplo 3, tensões de von Mises - Problema 3.

Os parâmetros dos exemplos e os resultados do volume e da tensão de von Mises máxima são mostrados pela Tabela 4.1. O valor de referência  $\sigma^R$  do Problema 3 foi escolhido igual a  $3.0 \text{ N/m}^2$  para a placa do Exemplo 1. Este é um valor intermédio entre os valores máximos obtidos para o Problema 1 e o Problema 2. Para o filete e o pórtico o valor de  $\sigma^R$  foi escolhido igual ao valor da força de superfície aplicada.

Tabela 4.1: Parâmetros e valores obtidos para os exemplos apresentados.

Exemplo	$v_0$ (m <sup>2</sup> )	$\sigma^R$ (N/m <sup>2</sup> )	Área final (m <sup>2</sup> )	$\sigma_{max}^m$ (N/m <sup>2</sup> )
Placa - Problema 1	3.60	-	3.600	3.16
Placa - Problema 2	3.60	-	3.600	2.83
Placa - Problema 3	3.60	3.0	3.427	3.02
Filete - Problema 1	0.58	-	0.580	2.97
Filete - Problema 2	0.58	-	0.559	1.15
Filete - Problema 3	0.58	1.0	0.578	1.83
Pórtico- Problema 1	0.70	-	0.700	38.92
Pórtico- Problema 2	0.70	-	0.700	1.70
Pórtico- Problema 3	0.70	1.0	0.671	2.26

As conclusões principais sobre os problemas de otimização considerados são:

- O Problema 1 funcionou bem no caso da placa, mas apresentou soluções com tensões de von Mises elevadas no caso do filete e o pórtico. Estas tensões elevadas parecem ser devidas às restrições geométricas impostas à geometria: O filete deve estar localizado entre as linhas verticais da Figura 4.21 e o pórtico tem os pontos A e B da Figura 4.25 fixos. O Problema 1 sempre utilizou todo o volume disponível como mostrado na Tabela 4.1 e colocou o material onde as maiores tensões de von Mises estavam localizadas. Isso reduz a complacência nestes problemas, mas não é a melhor estratégia para reduzir as tensões de von Mises. Por exemplo, o Problema 2 não utilizou todo o material disponível no caso do filete, ou seja, a restrição no volume não foi ativa na solução do problema. No caso do pórtico a situação é ainda mais expressiva, o Problema 1 colocou material em volta do ponto A da Figura 4.25, isto reduz significativamente a complacência mas ao mesmo tempo incrementa as tensões de von Mises nesse ponto. A melhor estratégia para esse problema, seguida pelo Problema 2, é colocar o material no lado oposto, em volta do ponto E onde estão as menores tensões de von Mises. Desta forma o Problema 2 consegue reduzir as tensões no ponto A.
- O Problema 2 apresentou as melhores soluções para todos os exemplos desde o ponto de vista da tensão de von Mises máxima. A maior desvantagem do Pro-

blema 2 é a de ser, computacionalmente, mais custoso que os outros dois. Principalmente por causa do grande número de restrições de desigualdade relativas às tensões de von Mises.

- O Problema 3 funcionou melhor que o Problema 1, ele apresentou boas soluções para os três exemplos. Comparado com o Problema 2, o Problema 3 apresentou tensões de von Mises maiores porque não reduziu completamente a concentração de tensões no filete e não utilizou todo o volume disponível no caso do pórtico. No entanto, ele encontrou boas soluções sendo computacionalmente bem menos custoso que o Problema 2.

# Capítulo 5

## Problema de Fundição Eletromagnética

A Fundição Eletromagnética (EMC do inglês *Electromagnetic Casting*) e o Processamento de Fundição por Suspensão Magnética (MSMP do inglês *Magnetic Suspension Melt Processing*) são importantes tecnologias na indústria metalúrgica. Elas utilizam campos eletromagnéticos para fundido, moldado e controle de solidificação de peças metálicas. Estas técnicas têm as vantagens de produzir superfícies de elevada qualidade, alta limpeza e baixa contaminação, além de permitir fabricação de alta precisão. A EMC foi empregada inicialmente no processo de fundição contínua sem utilização de forma, mas na atualidade é utilizada principalmente na fabricação de lingotes de ligas de alumínio [70]. Outra aplicação importante, muito utilizada na aeronáutica, na astronáutica, na energia e na engenharia química é na fabricação de componentes de motores feitos de materiais “superalloy” (Ni,Ti,...) [71].

Na Seção 5.1 é descrito o modelo matemático do problema EMC, e nas Seções 5.2, 5.3, 5.4 e 5.5 são descritos os métodos desenvolvidos para a resolução do problema de análise em EMC e dos problemas de otimização associados ao projeto dos indutores. O apresentado nas Seções 5.4 e 5.5 é baseado nos trabalhos [72, 73].

### 5.1 O modelo matemático

O modelo simplificado da fundição eletromagnética estudada aqui considera o caso de uma coluna vertical de metal líquido caindo dentro de um campo eletromagnético induzido por condutores elétricos verticais. Assume-se que a frequência da corrente elétrica é muito alta e portanto o campo eletromagnético não penetra dentro do metal líquido. Desta

forma, o efeito do campo eletromagnético é reduzido a uma pressão magnética atuando na superfície da interface.

Considere a seção transversal do problema representada na Figura 5.1. Na figura  $\omega$  representa o domínio compacto e simplesmente conexo no plano horizontal ocupado pelo metal líquido,  $\Gamma = \partial\omega$  é o seu contorno e  $\Omega = \mathbb{R}^2 \setminus \omega$  é o exterior do metal líquido.

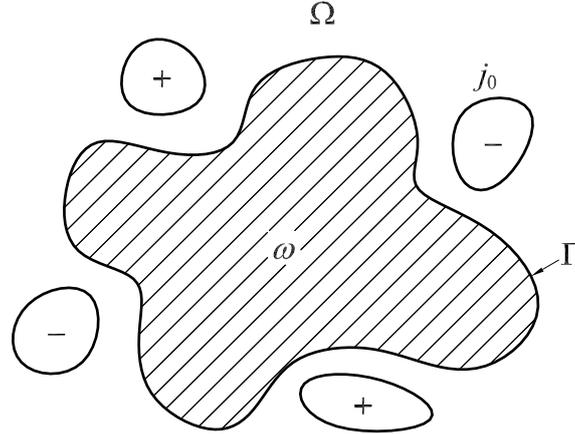


Figura 5.1: Problema EMC.

O campo magnético exterior pode ser calculado como a solução do seguinte problema de valores de contorno:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}_0 \quad \text{em } \Omega, \quad (5.1)$$

$$\nabla \cdot \mathbf{B} = 0 \quad \text{em } \Omega, \quad (5.2)$$

$$\mathbf{B} \cdot \mathbf{n} = 0 \quad \text{em } \Gamma, \quad (5.3)$$

$$\|\mathbf{B}(x)\| = O(\|x\|^{-1}) \quad \text{quando } \|x\| \rightarrow \infty \text{ em } \Omega. \quad (5.4)$$

onde os campos  $\mathbf{j}_0 = (0, 0, j_0)$  e  $\mathbf{B} = (B_1, B_2, 0)$  representam, respectivamente, as médias quadráticas do vetor de densidade de corrente elétrica e do vetor de campo magnético. A constante  $\mu_0$  é a permeabilidade magnética do vácuo e  $\mathbf{n}$  é o vetor unitário normal a superfície  $\Gamma$ . Aqui é assumido que  $j_0$  tem suporte compacto em  $\Omega$  e, além disso, verifica a seguinte igualdade:

$$\int_{\Omega} j_0 \, d\Omega = 0. \quad (5.5)$$

Por outra lado, o campo magnético causa uma pressão que movimenta o metal líquido até que este atinge uma configuração de equilíbrio. Pode ser visto que esta configuração

de equilíbrio é caracterizada pela seguinte condição:

$$\frac{1}{2\mu_0} \|\mathbf{B}\|^2 + \sigma\mathcal{C} = \bar{p} \quad \text{constante em } \Gamma, \quad (5.6)$$

onde  $\mathcal{C}$  é a curvatura de  $\Gamma$  (vista desde o metal líquido),  $\sigma$  é a tensão superficial do metal líquido e  $\bar{p}$  é um valor real que não depende do ponto considerado e é uma incógnita do problema, seu significado físico é a diferença entre a pressão interna e a externa do líquido. Assume-se também que a área do domínio  $\omega$  é conhecida e igual ao valor  $S_0$ :

$$\int_{\omega} d\Omega = S_0. \quad (5.7)$$

Seja  $\mathbf{B}_\omega$  o campo magnético solução das Eqs. (5.1)-(5.4). No problema direto a densidade de corrente elétrica  $j_0$  é conhecida e deve-se encontrar a geometria de  $\omega$  que verifique a Eq. (5.7) de forma tal que o campo magnético  $\mathbf{B}_\omega$  verifique a condição de equilíbrio dada pela Eq. (5.6) para uma determinada constante real  $\bar{p}$ .

As Eqs. (5.1)-(5.5) e a condição de  $j_0$  ter suporte compacto em  $\Omega$ , implicam que existe uma função potencial  $\varphi : \Omega \rightarrow \mathbb{R}$  tal que  $\mathbf{B} = (\frac{\partial\varphi}{\partial x_2}, -\frac{\partial\varphi}{\partial x_1}, 0)$  e  $\varphi$  é solução de:

$$\begin{aligned} -\Delta\varphi &= \mu_0 j_0 \quad \text{em } \Omega, \\ \varphi &= 0 \quad \text{em } \Gamma, \\ \varphi(x) &= O(1) \quad \text{quando } \|x\| \rightarrow \infty. \end{aligned} \quad (5.8)$$

Em termos do potencial  $\varphi$ , a condição de equilíbrio dada pela Eq. (5.6) fica:

$$\frac{1}{2\mu_0} \|\nabla\varphi\|^2 + \sigma\mathcal{C} = \bar{p} \quad \text{constante em } \Gamma. \quad (5.9)$$

O problema direto, em termos do potencial, consiste em encontrar o domínio  $\omega$  tal que a solução  $\varphi_\omega$  da Eq. (5.8) verifique a Eq. (5.9) para uma determinada constante real  $\bar{p}$ .

### 5.1.1 A formulação variacional

Sob certas condições as configurações de equilíbrio podem ser calculadas como os pontos estacionários em relação ao domínio do seguinte funcional energia total:

$$E(\omega) = -\frac{1}{2\mu_0} \int_{\Omega} \|\nabla\varphi_\omega\|^2 d\Omega + \sigma P(\omega), \quad (5.10)$$

sujeito à restrição de igualdade na área de  $\omega$ :

$$\int_{\omega} d\Omega = S_0. \quad (5.11)$$

Na Eq. (5.10),  $\varphi_\omega$  é a solução da Eq. (5.8) e  $P(\omega)$  é o perímetro de  $\omega$ , ou seja, a longitude de  $\Gamma = \partial\omega$  quando  $\partial\omega$  é suficientemente regular (por exemplo de classe  $C^1$ ):

$$P(\omega) = \int_{\Gamma} d\Gamma. \quad (5.12)$$

A formulação variacional do problema direto consiste em encontrar o domínio  $\omega$  como um ponto estacionário da energia total dada pela Eq. (5.10), sujeito a restrição de igualdade dada pela Eq. (5.11). Como  $\varphi_\omega$  é solução da Eq. (5.8), para provar a equivalência do problema direto com esta formulação variacional é suficiente mostrar que a relação de equilíbrio é automaticamente assegurada para todos os pontos estacionários.

Nesta seção são derivadas as condições necessárias para que o domínio  $\omega$  seja um ponto estacionário da energia total dada pela Eq. (5.10), sujeito à restrição dada pela Eq. (5.11).

Seja  $L$  a função lagrangiana definida em  $\mathcal{O}(\omega) \times \mathbb{R}$  por:

$$L(\omega, \bar{p}) = E(\omega) - \bar{p}(S(\omega) - S_0), \quad (5.13)$$

onde  $S(\omega) = \int_{\omega} d\Omega$  e  $\bar{p}$  aqui aparece como sendo o multiplicador de Lagrange associado à restrição de igualdade. Então, de forma análoga à Eqs. (2.10), a condição de otimalidade de primeira ordem é [38, Teorema 3.23]:

$$L'(\omega, \bar{p})(V) = 0, \quad \forall V \in W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2). \quad (5.14)$$

onde  $L'(\omega, \bar{p})(V)$  é a derivada em relação à forma de  $L$  avaliada na direção  $V$ .

Este tipo de condição de otimalidade aparece usualmente em problemas de hidrostática e outros problemas considerando fluidos. Ver, por exemplo, o trabalho [74] onde uma grande classe de problemas de equilíbrio é considerada.

O teorema seguinte mostra a expressão que pode ser utilizada para calcular  $L'(\omega, \bar{p})(V)$ .

**Teorema 5.1** *Seja  $\Omega$  o complemento do domínio compacto de interior não vazio  $\omega$  em  $\mathbb{R}^2$ . Assuma que  $\Gamma = \partial\omega = \partial\Omega$  é de classe  $C^2$ . Seja  $V \in W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2)$  com suporte compacto e  $\|V\|_{W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2)} < 1$ . Seja  $j_0$  uma função quadrado integrável de  $\Omega$  em  $\mathbb{R}$  com suporte compacto em  $\Omega$ .*

Então, existe uma única solução  $\varphi_{\omega_V} \in C^1(\overline{\Omega}_V)$  (ver [75, 76]) de:

$$\begin{aligned} -\Delta\varphi_{\omega_V} &= \mu_0 j_0 & \text{em } \Omega_V, \\ \varphi_{\omega_V} &= 0 & \text{em } \partial\Omega_V, \\ \varphi_{\omega_V}(x) &= O(1) & \text{quando } \|x\| \rightarrow \infty. \end{aligned} \quad (5.15)$$

e a derivada com relação à forma da função lagrangiana  $L$  é dada por:

$$L'(\omega, \bar{p})(V) = \int_{\Gamma} \left( \frac{1}{2\mu_0} \|\nabla\varphi_{\omega}\|^2 + \sigma\mathcal{C} - \bar{p} \right) (V \cdot \mathbf{n}) \, d\Gamma, \quad (5.16)$$

onde  $\mathbf{n}$  é a normal unitária de  $\Gamma$  orientada em direção a  $\Omega$ ,  $\mathcal{C}$  é a curvatura de  $\Gamma$  (vista desde o metal líquido) e  $\varphi_{\omega}$  é a solução da Eq. (5.8).

Prova: Ver na Seção 2.4 ou nas referências [38, 76–78].  $\square$

Este problema é muito similar a outros considerados por vários autores. Mais detalhes sobre o modelo matemático e sobre as hipóteses simplificadoras requeridas podem ser encontrados nas referências: [76, 78–92].

## 5.2 O problema direto

No problema direto a distribuição da densidade de corrente  $j_0$  é considerada conhecida, e deve-se achar o domínio  $\omega$  que verifique a Eq. (5.11) e seja tal que a função  $\varphi$  que satisfaz o sistema da Eq. (5.8) verifique também a equação de equilíbrio na forma variacional:

$$\int_{\Gamma} \left( \frac{1}{2\mu_0} \|\nabla\varphi\|^2 + \sigma\mathcal{C} - \bar{p} \right) (V \cdot \mathbf{n}) \, d\Gamma = 0 \quad \forall V \text{ in } W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2). \quad (5.17)$$

Para isto é considerada a função  $\varphi_1$ , solução particular da Eq. (5.8) em  $\mathbb{R}^2$ :

$$\varphi_1(x) = -\frac{\mu_0}{2\pi} \int_{\mathbb{R}^2} \ln \|x - y\| j_0(y) \, d\Omega. \quad (5.18)$$

Esta função é solução do problema:

$$-\Delta\varphi_1(x) = \mu_0 j_0 \quad \text{em } \mathbb{R}^2, \quad (5.19)$$

$$\varphi_1(x) = O(1) \quad \text{quando } \|x\| \rightarrow \infty. \quad (5.20)$$

Então, a função  $\varphi$  pode ser calculada como:

$$\varphi(x) = v(x) + \varphi_1(x), \quad (5.21)$$

onde a função  $v$  é solução do seguinte problema exterior:

$$\begin{aligned} -\Delta v(x) &= 0 && \text{em } \Omega, \\ v(x) &= -\varphi_1(x) && \text{em } \Gamma, \\ v(x) &= O(1) && \text{quando } \|x\| \rightarrow \infty. \end{aligned} \quad (5.22)$$

Seguindo as idéias apresentadas em [75], uma representação integral da solução da Eq. (5.22) é dada por:

$$v(x) = -\frac{1}{2\pi} \int_{\Gamma} q(y) \ln \|x - y\| d\Gamma + c, \quad (5.23)$$

onde a constante  $c$  é o valor no infinito da função  $v$  e a função  $q \in H^{-1/2}(\Gamma)$  deve verificar:

$$\int_{\Gamma} q(x) d\Gamma = 0. \quad (5.24)$$

Restam as condições de contorno em  $\Gamma$  que aqui são impostas em uma formulação fraca. A função  $q$  deve verificar:

$$\begin{aligned} -\frac{1}{2\pi} \int_{\Gamma} g(x) \int_{\Gamma} q(y) \ln \|x - y\| d\Gamma d\Gamma + \\ + c \int_{\Gamma} g(x) d\Gamma = - \int_{\Gamma} \varphi_1(x) g(x) d\Gamma \quad \forall g \in H^{-1/2}(\Gamma). \end{aligned} \quad (5.25)$$

Finalmente, a norma do gradiente da função  $\varphi$  presente na Eq. (5.17) pode ser calculada como:

$$\begin{aligned} \|\nabla \varphi\| &= \left\| \frac{\partial \varphi}{\partial \mathbf{n}} \right\|, \\ &= \left\| \frac{\partial \varphi_1}{\partial \mathbf{n}} + \frac{\partial v}{\partial \mathbf{n}} \right\|, \end{aligned} \quad (5.26)$$

onde a primeira igualdade provém do fato de  $\varphi$  ser constante em  $\Gamma$ . A derivada normal de  $\varphi_1$  pode ser obtida da Eq. (5.18):

$$\frac{\partial \varphi_1}{\partial \mathbf{n}_x}(x) = -\frac{\mu_0}{2\pi} \int_{\mathbb{R}^2} \frac{\partial}{\partial \mathbf{n}_x} \ln \|x - y\| j_0(y) d\Omega. \quad (5.27)$$

A seguinte expressão pode ser empregada para a função  $v$ :

$$\frac{\partial v}{\partial \mathbf{n}_x}(x) = -\frac{1}{2\pi} \int_{\Gamma} q(y) \frac{\partial}{\partial \mathbf{n}_x} \ln \|x - y\| d\Gamma + \frac{1}{2} q(x) \quad \forall x \in \Gamma, \quad (5.28)$$

onde a integral da Eq. (5.28) deve ser entendida no sentido do valor principal de Cauchy.

Em resumo, o problema direto é resolvido encontrando o domínio  $\omega$ , a função  $q$  e as constantes reais  $c$  e  $\bar{p}$  que verificam as seguintes equações:

- As equações de estado:

$$-\frac{1}{2\pi} \int_{\Gamma} g(x) \int_{\Gamma} q(y) \ln \|x - y\| d\Gamma d\Gamma + c \int_{\Gamma} g(x) d\Gamma = - \int_{\Gamma} \varphi_1(x) g(x) d\Gamma \quad \forall g \in H^{-1/2}(\Gamma), \quad (5.29)$$

$$\int_{\Gamma} q(x) d\Gamma = 0. \quad (5.30)$$

- A restrição de igualdade na área de  $\omega$ :

$$\int_{\omega} d\Omega = S_0. \quad (5.31)$$

- A equação de equilíbrio:

$$\int_{\Gamma} \left( \frac{1}{2\mu_0} \|\nabla\varphi\|^2 + \sigma\mathcal{C} - \bar{p} \right) (V \cdot \mathbf{n}) d\Gamma = 0 \quad \forall V \text{ in } W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2). \quad (5.32)$$

Nas equações anteriores, as funções  $\varphi_1$ ,  $\varphi$  e  $v$  são dadas pelas Eqs. (5.18), (5.21) e (5.23). A norma  $\|\nabla\varphi\|$  da Eq. (5.32) pode ser calculada utilizando as Eqs. (5.26)-(5.28).

## Discretização do contorno

Para o domínio é considerada uma aproximação definida por um contorno  $\Gamma^h$  linear por partes, ou seja,  $\Gamma^h$  é a união de  $n_e$  elementos lineares  $\ell^j$  em  $\mathbb{R}^2$ ,  $1 \leq j \leq n_e$ . Os nós do contorno são aqui denominados  $x^i$ ,  $1 \leq i \leq n_e$ .

Uma transformação paramétrica do contorno  $\Gamma^h$  é definida associando a cada nó  $x^i$  do contorno uma direção  $\hat{V}^i \in \mathbb{R}^2$ . Desta forma, é construído um campo de velocidades linear por partes  $V^i : \Gamma^h \rightarrow \mathbb{R}^2$  tal que  $V^i(x^k) = \delta_{ik} \hat{V}^i$ . O suporte de  $V^i$  é, então, igual à união dos elementos que possuem o nó  $x^i$ . A transformação paramétrica  $T_{\mathbf{u}}$  é dada como:

$$T_{\mathbf{u}}(x) = x + V_{\mathbf{u}}(x), \quad (5.33)$$

com  $V_{\mathbf{u}}$  definida como:

$$V_{\mathbf{u}}(x) = \sum_{i=1}^{n_e} \mathbf{u}_i V^i(x), \quad (5.34)$$

onde  $\mathbf{u}^T = (\mathbf{u}_1, \dots, \mathbf{u}_{n_e}) \in \mathbb{R}^{n_e}$  é o vetor das incógnitas que determinam a evolução do contorno. Portanto, o contorno atualizado  $\Gamma_{\mathbf{u}}$  é dado por:

$$\Gamma_{\mathbf{u}} = \{X \mid X = x + V_{\mathbf{u}}(x); x \in \Gamma^h\}. \quad (5.35)$$

Esta representação tem a vantagem de definir um único grau de liberdade por cada nó. No que segue  $\omega_{\mathbf{u}}$  e  $\Omega_{\mathbf{u}}$  são os domínios interior e exterior relacionados ao contorno  $\Gamma_{\mathbf{u}}$  com o objetivo de destacar a dependência em relação ao vetor das incógnitas  $\mathbf{u}$ .

### Discretização da equação de estado

Para os cálculos numéricos se considera a aproximação constante por partes  $q_h(x)$  de  $q(x)$  da seguinte forma:

$$q_h(x) = \sum_{j=1}^{n_e} q_j e_j(x), \quad (5.36)$$

onde  $e_j(x) = 1$  se  $x \in \ell_j$  ou zero no caso contrário.

Substituindo a função  $g$  na Eq. (5.29) por  $e_i$ , com  $1 \leq i \leq n_e$ , a formulação fraca do problema de valores de contorno dado pelas Eqs. (5.29) e (5.30) fica:

$$\mathbf{A}(\mathbf{u})\mathbf{q} = \mathbf{b}(\mathbf{u}), \quad (5.37)$$

onde o vetor  $\mathbf{q}^T = (q_1, \dots, q_{n_e}, c) \in \mathbb{R}^{n_e+1}$ , os coeficientes  $\mathbf{A}_{ij}$  da matriz simétrica  $\mathbf{A}$  são:

$$\mathbf{A}_{ij}(\mathbf{u}) = -\frac{1}{2\pi} \int_{\ell_i} \int_{\ell_j} \ln \|x - y\| \, d\Gamma \, d\Gamma, \quad 1 \leq i, j \leq n_e, \quad (5.38)$$

$$\mathbf{A}_{ij}(\mathbf{u}) = \int_{\ell_j} d\Gamma, \quad i = n_e + 1, \text{ e } 1 \leq j \leq n_e, \quad (5.39)$$

e os componentes  $\mathbf{b}_i$  do vetor  $\mathbf{b}$  são:

$$\mathbf{b}_i(\mathbf{u}) = - \int_{\ell_i} \varphi_1(x) \, d\Gamma, \quad 1 \leq i \leq n_e, \quad (5.40)$$

$$\mathbf{b}_i(\mathbf{u}) = 0, \quad i = n_e + 1. \quad (5.41)$$

### Equação de equilíbrio

Projetando a equação de equilíbrio no espaço de dimensão finita gerado por  $V_i$ ,  $1 \leq i \leq n_e$ , a versão discreta da equação de equilíbrio é a seguinte:

$$DE_i(\mathbf{u}, \mathbf{q}, \bar{p}) = \int_{\Gamma_{\mathbf{u}}} \left( \frac{1}{2\mu_0} \|\nabla\varphi\|^2 - \bar{p} \right) (V^i \cdot \mathbf{n}) \, d\Gamma + \mathcal{C}^i \cdot \hat{V}^i, \quad (5.42)$$

onde  $\mathcal{C}^i$  é uma aproximação da curvatura média no ponto  $x^i$ :

$$\mathcal{C}^i = \left( \frac{(x^i - x^{i-1})}{\|x^i - x^{i-1}\|} - \frac{(x^{i+1} - x^i)}{\|x^{i+1} - x^i\|} \right). \quad (5.43)$$

O gradiente  $\nabla\varphi$  na equação anterior pode ser obtido a partir das Eqs. (5.26)-(5.28).

## Restrição de área

Chamando  $(x_1^j, y_1^j)$  e  $(x_2^j, y_2^j)$  às coordenadas do primeiro e do segundo nós do elemento  $\ell_j$ , a área  $S(\mathbf{u})$  encerrada pelo contorno  $\Gamma_{\mathbf{u}}$  pode ser calculada na forma:

$$S(\mathbf{u}) = \frac{1}{2} \sum_{i=1}^{n_e} (y_2^i x_1^i - y_1^i x_2^i). \quad (5.44)$$

### 5.2.1 Algoritmo de resolução

As equações a serem satisfeitas podem ser expressadas como:

$$\mathbf{G}(\mathbf{u}, \mathbf{q}, \bar{p}) = \begin{pmatrix} \mathbf{A}(\mathbf{u})\mathbf{q} - \mathbf{b}(\mathbf{u}) \\ S(\mathbf{u}) - S_0 \\ DE(\mathbf{u}, \mathbf{q}, \bar{p}) \end{pmatrix} = 0. \quad (5.45)$$

O algoritmo utilizado para resolver o problema direto é similar ao algoritmo de Newton, o pseudocódigo é o seguinte:

#### Algoritmo

Parâmetros do algoritmo: Tol, MaxIt, PasoMax.

Dados:  $k = 0$ ,  $\mathbf{u}_0 = 0$ ,  $\mathbf{q}_0 = 0$ ,  $\bar{p}_0 = 0$ ,  $I$ ,  $\{\alpha_i, 1 \leq i \leq n_c\}$ ,  $S_0$ ,  $\sigma$ ,  $\mu_0$ , as posições dos nós  $\{x^i, 1 \leq i \leq n_e\}$  e as posições dos indutores  $\{x_c^i, 1 \leq i \leq n_c\}$ .

**Passo 1:** Ajuste da malha.

**Passo 2:** Cálculo das direções  $\hat{V}_i$  de movimento dos nós.

**Passo 3:** Cálculo da função  $\mathbf{G}(\mathbf{u}_k, \mathbf{q}_k, \bar{p}_k)$ .

**Passo 4:** Critério de parada: Se  $(\|\mathbf{G}(\mathbf{u}_k, \mathbf{q}_k, \bar{p}_k)\| < \text{Tol})$  ou  $(k \geq \text{MaxIt})$  pare.

**Passo 5:** Cálculo do gradiente  $\nabla \mathbf{G}(\mathbf{u}_k, \mathbf{q}_k, \bar{p}_k)$ .

**Passo 6:** Cálculo da direção de busca  $\mathbf{d}_k = (\mathbf{d}_k^{\mathbf{u}}, \mathbf{d}_k^{\mathbf{q}}, d_k^{\bar{p}})$  como a solução do sistema linear:

$$\nabla \mathbf{G}(\mathbf{u}_k, \mathbf{q}_k, \bar{p}_k) \mathbf{d}_k = -\mathbf{G}(\mathbf{u}_k, \mathbf{q}_k, \bar{p}_k).$$

**Passo 7:** Busca linear.

7.1: Defina  $t_k = \min(1, \text{PasoMax} / \max(|\mathbf{d}_k^{\mathbf{u}}|))$ .

7.2: Defina  $\mathbf{u}_{k+1} = t_k \mathbf{d}_k^{\mathbf{u}}$ ,  $\mathbf{q}_{k+1} = \mathbf{q}_k + t_k \mathbf{d}_k^{\mathbf{q}}$ ,  $\bar{p}_k = \bar{p}_k + t_k d_k^{\bar{p}}$ .

7.3: Até  $\|\mathbf{G}(\mathbf{u}_{k+1}, \mathbf{q}_{k+1}, \bar{p}_{k+1})\| \leq \|\mathbf{G}(\mathbf{u}_k, \mathbf{q}_k, \bar{p}_k)\|$ , defina  $t_k = 0.5t_k$  e retorne a 7.2.

**Passo 8:** Atualize as posições dos nós substituindo  $x^i$  por  $x^i + u_i \hat{V}^i$ , defina  $k = k + 1$  e retorne ao Passo 1.

## 5.2.2 Exemplos

Nesta seção apresentam-se seis exemplos que são resolvidos utilizando o algoritmo de otimização descrito na Seção 5.2.1. Para cada exemplo é mostrada a forma inicial e final do metal líquido. Todas as variáveis de estado são definidas com valores iniciais iguais a zero.

Para todos os problemas serão considerados indutores de seção transversal de área desprezível. Portanto, a distribuição da densidade de corrente elétrica  $j_0$  pode ser expressada na forma:

$$j_0 = I \sum_{i=1}^{n_c} \alpha_i \delta_{x_i}, \quad (5.46)$$

onde  $I$  é uma dada intensidade de corrente,  $\delta_{x_i}$ ,  $1 \leq i \leq n_c$ , são deltas de Dirac nos pontos  $x_i$  no plano, e  $\alpha_i$  são coeficientes adimensionados.

Note que por tratar-se de corrente alterna o sinal dos coeficientes  $\alpha_i$  somente indica quais deles estão em fase e quais deles correspondem a fase oposta. Em todas as figuras apresentadas, a posição dos indutores de corrente positiva será indicada por sinais de adição, em quanto que a posição dos indutores de corrente negativa será indicada por círculos.

### Exemplo 1

Para este exemplo consideraram-se os seguintes valores para os parâmetros:  $I = 0.1$ ,  $\alpha_i = \pm 1$  com o sinal dado pela Figura 5.2,  $\sigma = 1.0 \times 10^{-4}$  e  $S_0 = \pi$ .

O resultado do algoritmo de análise é apresentado pela Figura 5.2 e a evolução de  $\|G\|$  ao longo das iterações é mostrada pela Figura 5.3.

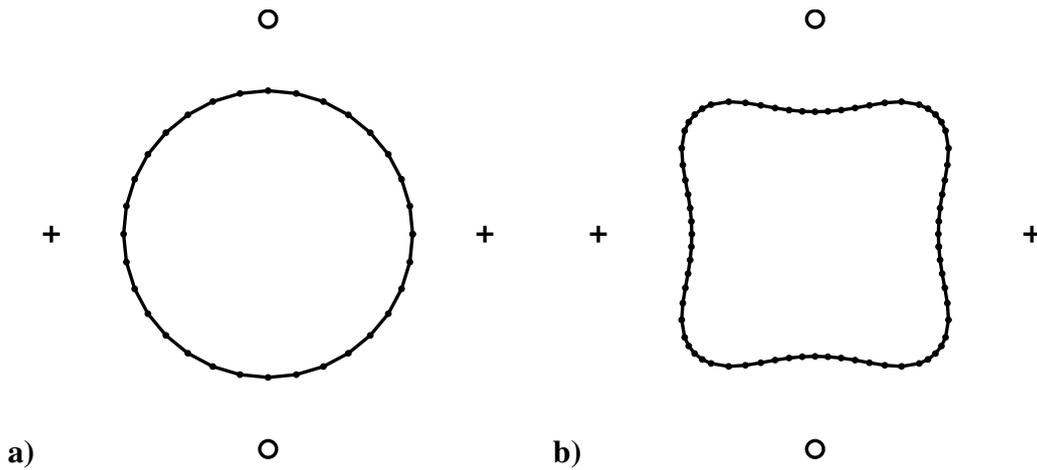


Figura 5.2: Exemplo 1, **a)** forma inicial, **b)** forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

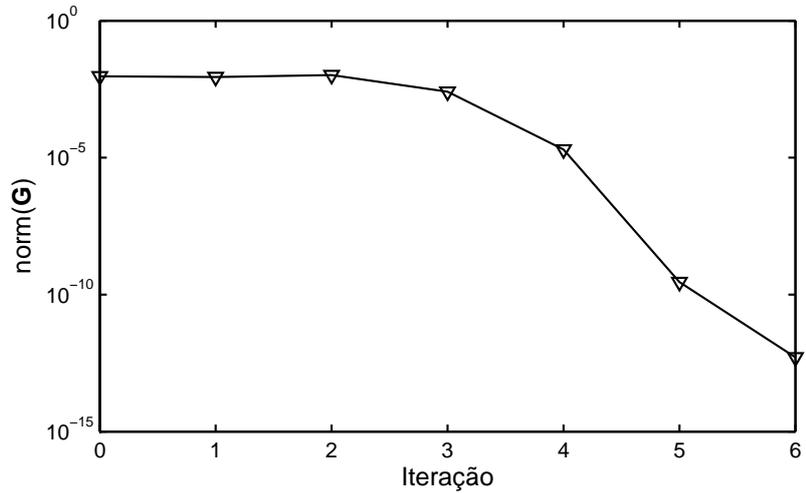


Figura 5.3: Exemplo 1, evolução de  $\|G\|$ .

**Exemplo 2**

Neste exemplo os parâmetros são:  $I = 0.2$ ,  $\alpha_i = \pm 1$  com o sinal dado pela Figura 5.4,  $\sigma = 1.0 \times 10^{-4}$  e  $S_0 = \pi$ .

O resultado do algoritmo de análise é apresentado pela Figura 5.4 e a evolução de  $\|G\|$  ao longo das iterações é mostrada pela Figura 5.5.

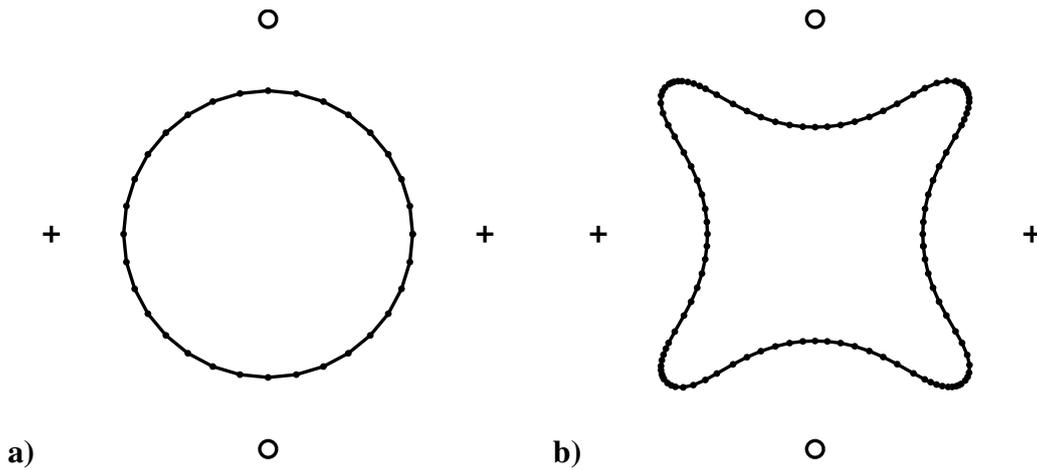


Figura 5.4: Exemplo 2, **a)** forma inicial, **b)** forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

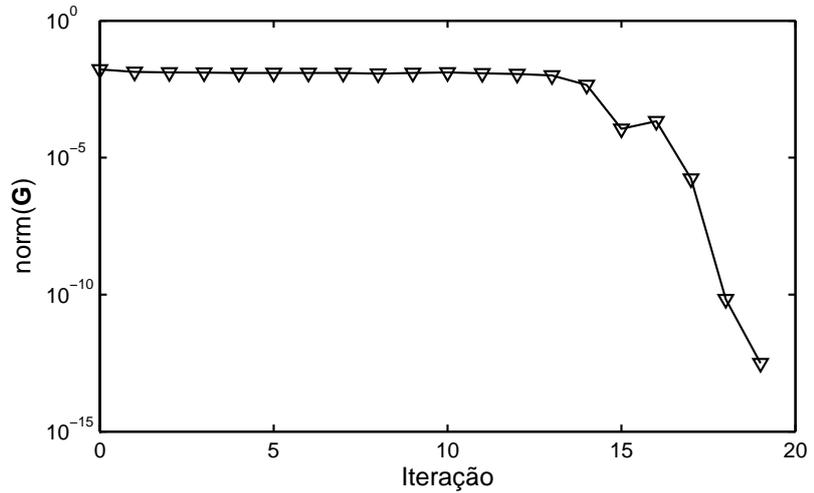


Figura 5.5: Exemplo 2, evolução de  $\|G\|$ .

**Exemplo 3**

Neste exemplo os parâmetros são:  $I = 1.0$ ,  $\alpha_i = \pm 4$  para os indutores centrais,  $\alpha_i = \pm 2$  para os extremos, com o sinal dado pela Figura 5.6,  $\sigma = 0.5 \times 10^{-1}$  e  $S_0 = 2.0$ .

O resultado do algoritmo de análise é apresentado pela Figura 5.6 e a evolução de  $\|G\|$  ao longo das iterações é mostrada pela Figura 5.7.

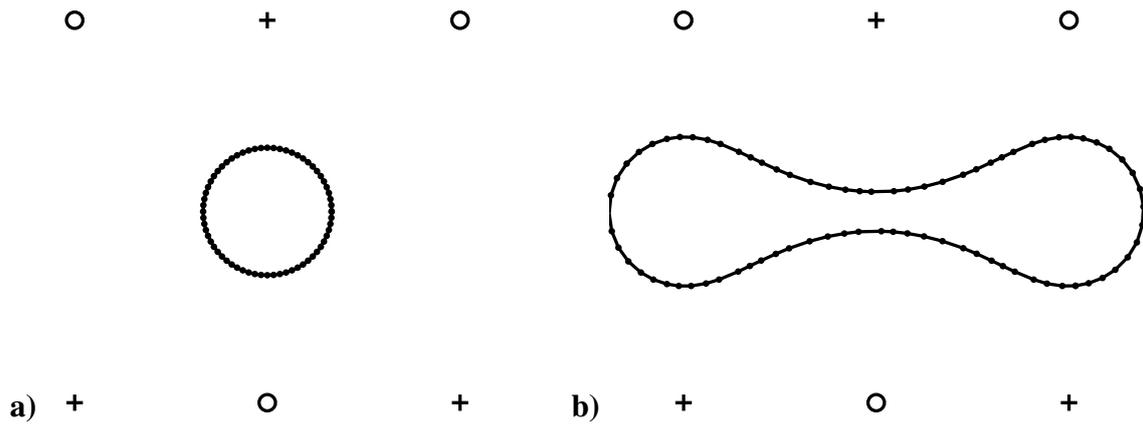


Figura 5.6: Exemplo 3, **a)** forma inicial, **b)** forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

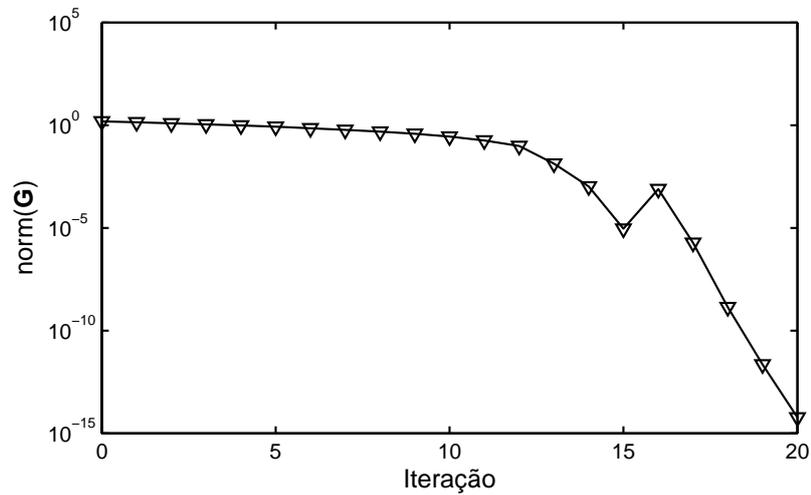


Figura 5.7: Exemplo 3, evolução de  $\|\mathbf{G}\|$ .

#### Exemplo 4

Neste exemplo os parâmetros são:  $I = 0.075$ ,  $\alpha_i = \pm 1$  com o sinal dado pela Figura 5.8,  $\sigma = 1.0 \times 10^{-4}$  e  $S_0 = 4.98$ .

O resultado do algoritmo de análise é apresentado pela Figura 5.8 e a evolução de  $\|\mathbf{G}\|$  ao longo das iterações é mostrada pela Figura 5.9.

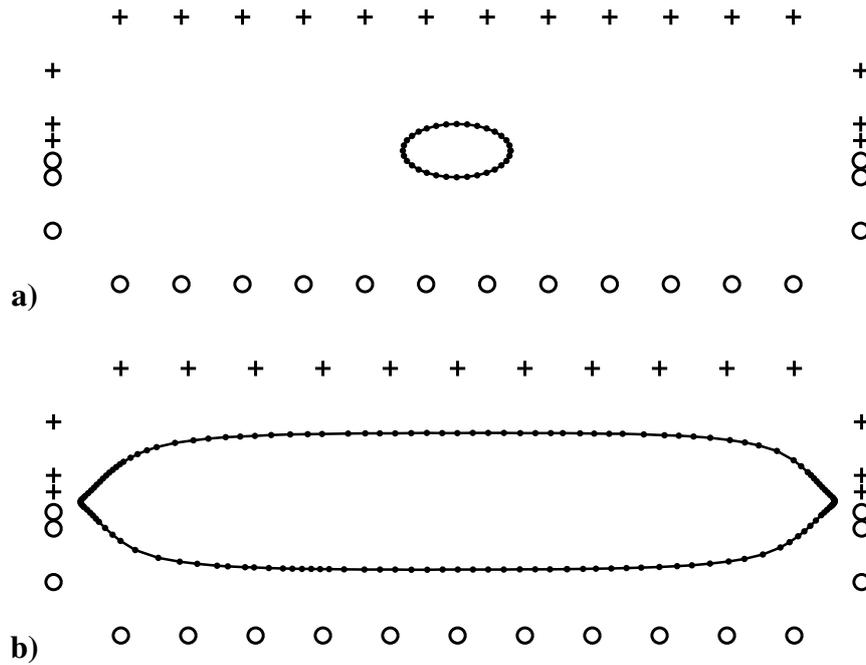


Figura 5.8: Exemplo 4, **a)** forma inicial, **b)** forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

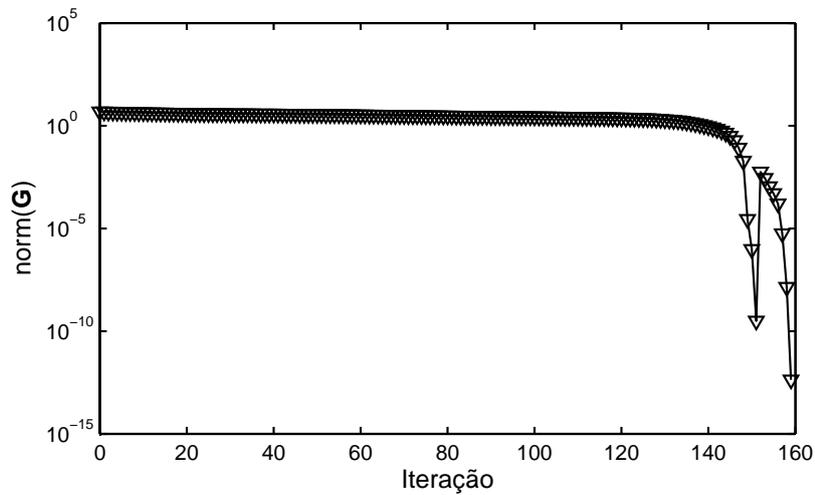


Figura 5.9: Exemplo 4, evolução de  $\|G\|$ .

### Exemplo 5

Neste exemplo os parâmetros são:  $I = 0.3$ ,  $\alpha_i = \pm 1$  com o sinal dado pela Figura 5.10,  $\sigma = 1.0 \times 10^{-4}$  e  $S_0 = 30.0$ .

O resultado do algoritmo de análise é apresentado pela Figura 5.10 e a evolução de  $\|G\|$  ao longo das iterações é mostrada pela Figura 5.11.

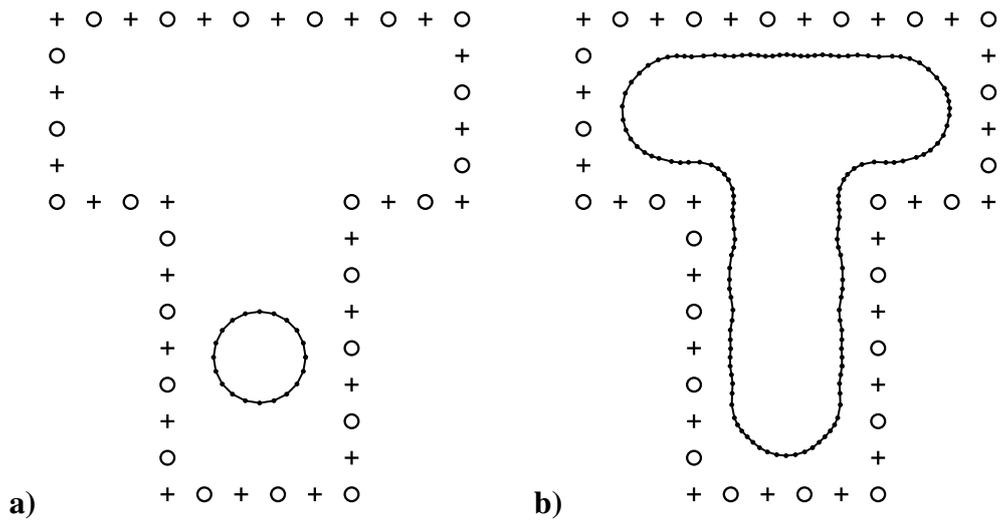


Figura 5.10: Exemplo 5, **a)** forma inicial, **b)** forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

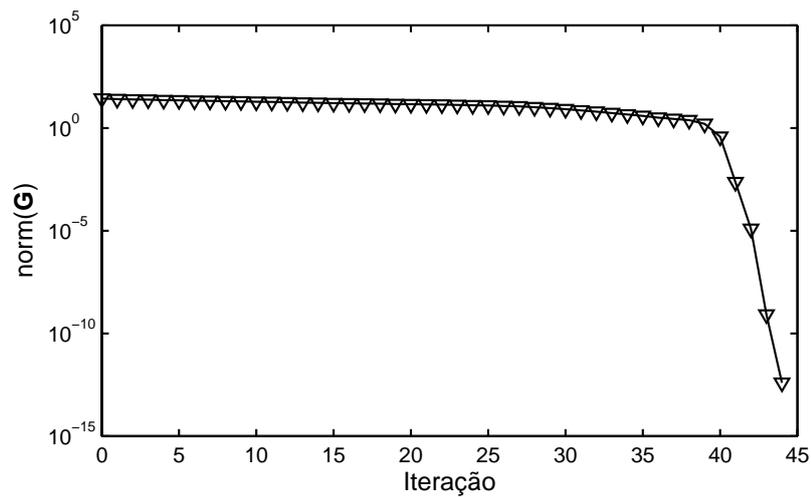


Figura 5.11: Exemplo 5, evolução de  $\|G\|$ .

### Exemplo 6

Neste exemplo os parâmetros são:  $I = 0.03$ ,  $\alpha_i = \pm 1$  com o sinal dado pela Figura 5.12,  $\sigma = 1.0 \times 10^{-4}$  e  $S_0 = 30.0$ .

O resultado do algoritmo de análise é apresentado pela Figura 5.12 e a evolução de  $\|G\|$  ao longo das iterações é mostrada pela Figura 5.13.

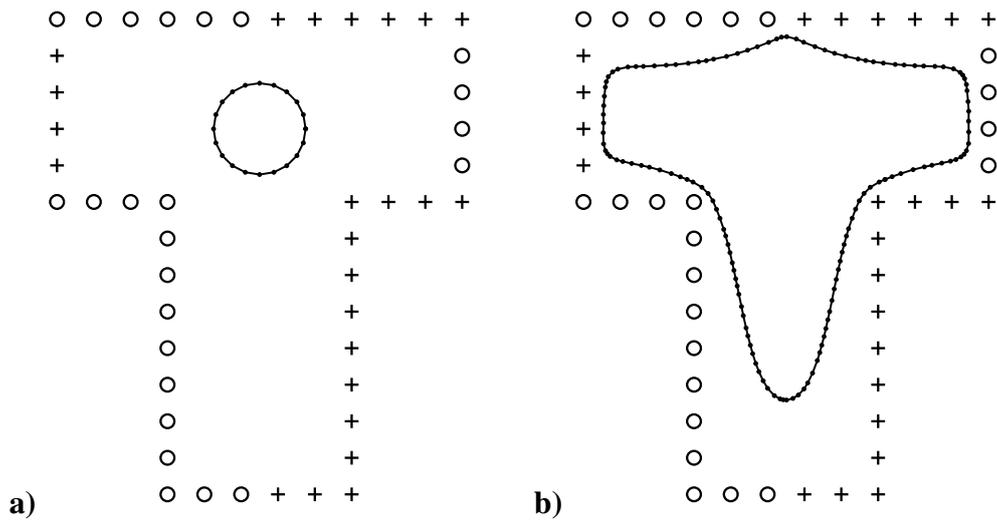


Figura 5.12: Exemplo 6, **a)** forma inicial, **b)** forma final. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

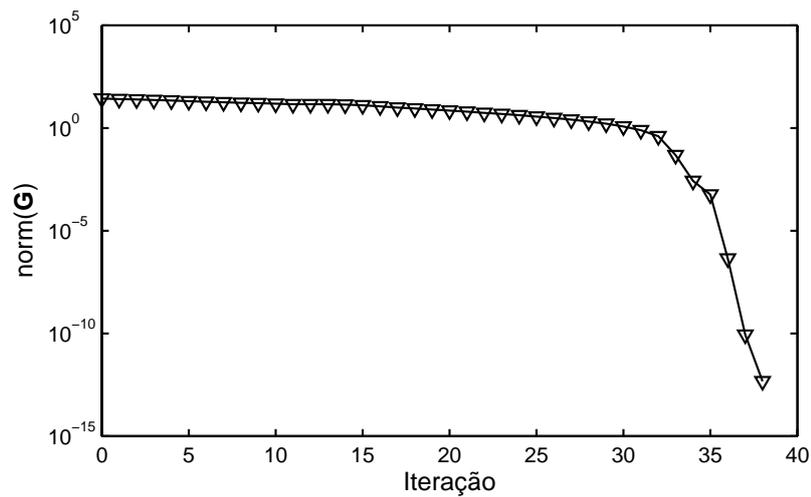


Figura 5.13: Exemplo 6, evolução de  $\|G\|$ .

### 5.3 O problema inverso

No problema inverso procura-se uma distribuição de corrente elétrica ao redor da coluna do metal líquido de forma tal que a seção transversal adquira uma dada forma objetivo.

**Nota:** Em alguns casos é possível encontrar uma distribuição de densidade de corrente elétrica de forma tal que o domínio objetivo  $\omega^*$  fique exatamente em equilíbrio. Este

assunto tem sido estudado e existem alguns poucos trabalhos sobre a existência de tais soluções. Ver, por exemplo, nas referências [76, 93, 94].

No caso bidimensional, assumindo  $\omega^*$  simplesmente conexo e de contorno dado por uma única curva de Jordan  $\Gamma$ , Henrot e Pierre [76] mostraram que uma solução das Eqs. (5.1)-(5.4) e (5.6), com  $j_0$  com suporte compacto em  $\Omega$ , pode ser encontrada para qualquer  $\bar{p}$  que satisfaça:

$$\bar{p} \geq \sigma \max_{x \in \Gamma} \mathcal{C}(x), \quad (5.47)$$

Isto é, assumindo que  $j_0$  tem suporte compacto em  $\Omega$ , e escolhendo  $\bar{p}$  satisfazendo a Eq. (5.47), então, existe  $\mathbf{B}$  que satisfaz as Eqs. (5.1)-(5.4) e (5.6) se e somente se:

- (i)  $\Gamma$  é uma curva analítica.
- (ii) Se  $\bar{p}$  é escolhido satisfazendo a igualdade na Eq. (5.47), então o máximo global da curvatura deve ser atingido em um número par de pontos de  $\Gamma$ .

Além disso, o campo magnético está bem determinado em uma vizinhança de  $\omega$  (unicidade local).

A Eq. (5.5) é também obtida se  $\bar{p}$  é escolhido satisfazendo a igualdade em (5.47). Uma densidade de corrente elétrica concentrada sobre uma curva em  $\Omega$  pode sempre ser encontrada. Ver na referência [76].

O procedimento aqui proposto é diferente dos procedimentos utilizados anteriormente por ser baseado na formulação do problema inverso como um problema de otimização. Assim, seja  $\omega^*$  a forma objetivo, deseja-se encontrar  $j_0$  solução do seguinte problema:

$$\min_{j_0} d(\omega, \omega^*), \quad (5.48)$$

onde  $d(\omega, \omega^*)$  corresponde a alguma noção de “distancia” entre os domínios  $\omega$  e  $\omega^*$ , com  $\omega$  pertencente ao conjunto de domínios admissíveis  $\mathcal{O}(\omega^*)$  e, além disso, em equilíbrio sob a ação do campo magnético induzido pela densidade de corrente  $j_0$ . Em outras palavras,  $\omega$  verifica a restrição de área dada pela Eq. (5.11) e o potencial  $\varphi_\omega$ , solução da Eq. (5.8), verifica a condição de equilíbrio da Eq. (5.49) para uma dada constante real  $\bar{p}$ :

$$\int_{\Gamma} \left( \frac{1}{2\mu_0} \|\nabla \varphi_\omega\|^2 + \sigma \mathcal{C} - \bar{p} \right) (V \cdot \mathbf{n}) \, d\Gamma = 0 \quad \forall W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2). \quad (5.49)$$

### 5.3.1 Duas abordagens para o problema inverso

Duas abordagens não equivalentes para encontrar uma solução aproximada do problema inverso da Eq. (5.48) são propostas: A primeira considera uma deformação do domínio  $\omega^*$  definida pelo mapeamento seguinte:

$$T_Z(x) = (Id + Z)(x), \quad \forall x \in \mathbb{R}^2, \quad (5.50)$$

onde  $Z$  é um campo vetorial regular com suporte compacto em  $\mathbb{R}^2$ . Então, definindo:

$$\begin{aligned} \omega_Z &= T_Z(\omega^*), \\ \Gamma_Z &= T_Z(\Gamma^*). \end{aligned}$$

A primeira formulação do problema inverso é:

$$\begin{aligned} &\min_{j_0, Z} \|Z\|_{L^2(\Gamma^*)}^2, \\ &\text{sujeito a:} \\ &\omega_Z \text{ é equilibrado sob } j_0. \end{aligned} \quad (5.51)$$

A segunda formulação do problema inverso é definida considerando uma função de folga  $p(x) : \Gamma \rightarrow \mathbb{R}$  de forma tal que a equação de equilíbrio seja verificada para o domínio objetivo:

$$\int_{\Gamma^*} \left( \frac{1}{2\mu_0} \|\nabla \varphi_\omega\|^2 + \sigma \mathcal{C} - \bar{p} + p \right) (V \cdot \mathbf{n}) \, d\Gamma = 0 \quad \forall V \text{ in } W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2). \quad (5.52)$$

A função  $p$  pode ser fisicamente interpretada como sendo uma pressão adicional atuando na interface. Dado  $j_0$  e  $\omega^*$ ,  $p$  é a pressão que equilibra a ação conjunta do campo magnético e a tensão superficial. A segunda formulação do problema inverso é uma abordagem indireta que tenta minimizar a norma  $L^2(\Gamma^*)$  da função  $p$ , sujeito à condição de equilíbrio modificada:

$$\begin{aligned} &\min_{j_0, p} \|p\|_{L^2(\Gamma^*)}^2, \\ &\text{sujeito a:} \\ &\omega^* \text{ é equilibrado sob a ação de } j_0 \text{ e } p. \end{aligned} \quad (5.53)$$

O fato de não ter variáveis de forma torna a solução numérica desta segunda formulação bem mais simples quando comparada com a primeira formulação. Em particular, a

análise de sensibilidade é consideravelmente simplificada e pode ser feita com um custo computacional muito menor. Se a função  $p$  resultar nula na solução do problema da Eq. (5.53), a distribuição da densidade de corrente elétrica  $j_0$  será também solução da primeira formulação com o domínio equilibrado  $\omega_Z$  exatamente igual ao domínio objetivo. No caso geral, a função  $p$  não se anulará na solução e, portanto, sem considerar o efeito de  $p$ , o domínio objetivo  $\omega^*$  não estará em equilíbrio sob a ação do campo magnético. No entanto, a norma de  $p$  foi minimizada, portanto o domínio em equilíbrio será uma aproximação do domínio objetivo. Por outro lado, desde que a solução da formulação da Eq. (5.53) possa ser calculada com um custo computacional menor, a distribuição  $j_0$  pode ser empregada como solução inicial para a formulação da Eq. (5.51).

### 5.3.2 Formulação SAND

Para achar soluções aproximadas para os problemas dados pelas Eqs. (5.51) e (5.53) é utilizada uma formulação SAND. As variáveis de estado  $\bar{p}$ ,  $c$  e  $q$  são incorporadas como incógnitas do problema de otimização e as equações de estado e equilíbrio são incorporadas como restrições de igualdade. Desta forma, o problema de otimização correspondente ao problema inverso da Eq. (5.51) fica:

$$\min_{j_0, Z, \bar{p}, c, q} \|Z\|_{L^2(\Gamma^*)}^2, \quad (5.54)$$

sujeito à restrição de área:

$$\int_{\omega_Z} d\Omega = S_0, \quad (5.55)$$

às equações de estado:

$$\begin{aligned} -\frac{1}{2\pi} \int_{\Gamma_Z} g(x) \int_{\Gamma_Z} q(y) \ln \|x - y\| d\Gamma d\Gamma + \\ + c \int_{\Gamma_Z} g(x) d\Gamma = - \int_{\Gamma_Z} \varphi_1(x) g(x) d\Gamma \quad \forall g \in H^{-1/2}(\Gamma_Z), \end{aligned} \quad (5.56)$$

$$\int_{\Gamma_Z} q(y) d\Gamma = 0, \quad (5.57)$$

e à equação de equilíbrio:

$$\int_{\Gamma_Z} \left( \frac{1}{2\mu_0} \|\nabla\varphi\|^2 + \sigma C - \bar{p} \right) (V \cdot \mathbf{n}) d\Gamma = 0 \quad \forall V \text{ in } W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2), \quad (5.58)$$

onde  $\varphi_1$ ,  $\varphi$ , e  $v$  são dadas pelas Eqs. (5.18), (5.21) e (5.23).

No caso da formulação da Eq. (5.53), o problema de otimização fica:

$$\min_{j_0, \bar{p}, \bar{c}, q} \|p\|_{L^2(\Gamma^*)}^2, \quad (5.59)$$

sujeito às equações de estado:

$$\begin{aligned} -\frac{1}{2\pi} \int_{\Gamma^*} g(x) \int_{\Gamma^*} q(y) \ln \|x - y\| d\Gamma d\Gamma + \\ + c \int_{\Gamma^*} g(x) d\Gamma = - \int_{\Gamma^*} \varphi_1(x) g(x) d\Gamma \quad \forall g \in H^{-1/2}(\Gamma^*), \end{aligned} \quad (5.60)$$

$$\int_{\Gamma^*} q(y) d\Gamma = 0, \quad (5.61)$$

e à equação de equilíbrio:

$$\int_{\Gamma^*} \left( \frac{1}{2\mu_0} \|\nabla\varphi\|^2 + \sigma\mathcal{C} - \bar{p} + p \right) (V \cdot \mathbf{n}) d\Gamma = 0 \quad \forall V \text{ in } W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2). \quad (5.62)$$

## 5.4 Otimização da posição dos indutores

Nesse caso é assumido que os indutores têm seção transversal de área desprezível. A distribuição da densidade de corrente elétrica  $j_0$  é assumida da forma:

$$j_0 = I \sum_{i=1}^{n_c} \alpha_i \delta_{x_i}, \quad (5.63)$$

onde  $I$  é uma dada intensidade de corrente,  $\delta_{x_i}$ ,  $1 \leq i \leq n_c$ , são deltas de Dirac nos pontos  $x_i$  no plano, e  $\alpha_i$  são coeficientes adimensionados. O problema inverso consiste em determinar as posições no plano dos pontos  $x_i$ .

### 5.4.1 Problema discreto

Na versão discreta do problema o campo  $Z$  é tratado na mesma forma que o campo  $V$ . Para isso, a cada nó  $x^i$  do contorno objetivo  $\Gamma^*$  é associada uma direção  $\hat{Z}^i \in \mathbb{R}^2$ . Desta forma é construído um campo de velocidades linear por partes  $Z^i : \Gamma^* \rightarrow \mathbb{R}^2$  tal que  $Z^i(x^k) = \delta_{ik} \hat{Z}^i$ . O suporte de  $Z^i$  é, então, igual à união dos elementos que possuem o nó  $x^i$ .  $Z(\mathbf{z})$  é então definida como:

$$Z(\mathbf{z})(x) = \sum_{i=1}^{n_e} \mathbf{z}_i Z^i(x), \quad (5.64)$$

onde  $\mathbf{z}^T = (\mathbf{z}_1, \dots, \mathbf{z}_{n_e}) \in \mathbb{R}^{n_e}$  é o vetor das incógnitas que determinam a evolução do contorno. O contorno atualizado  $\Gamma_{\mathbf{z}}$  é dado por:

$$\Gamma_{\mathbf{z}} = \{X \mid X = x + Z(\mathbf{z})(x); x \in \Gamma^*\} . \quad (5.65)$$

### Primeira formulação

A função objetivo do primeiro problema inverso pode ser calculada como:

$$\|Z(\mathbf{z})\|_{L^2(\Gamma^*)}^2 = \frac{1}{3} \sum_{j=1}^{n_e} [Z(\mathbf{z})_1^j \cdot Z(\mathbf{z})_1^j + Z(\mathbf{z})_1^j \cdot Z(\mathbf{z})_2^j + Z(\mathbf{z})_2^j \cdot Z(\mathbf{z})_2^j], \quad (5.66)$$

onde  $Z(\mathbf{z})_1^j$  e  $Z(\mathbf{z})_2^j$  são os deslocamentos do primeiro e do segundo nós do elemento  $\ell_j$ .

Chamando  $\mathbf{u}_c$  ao vetor das variáveis que definem as posições dos pontos  $x_i$ , com  $1 \leq i \leq n_c$ , a versão discreta da primeira formulação do problema inverso fica:

$$\min_{\mathbf{u}_c, \mathbf{z}, \mathbf{q}, \bar{p}} \|Z(\mathbf{z})\|_{L^2(\Gamma^*)}^2 ,$$

sujeito a:

$$\mathbf{A}(\mathbf{z})\mathbf{q} - \mathbf{b}(\mathbf{u}_c, \mathbf{z}) = 0 , \quad (5.67)$$

$$S(\mathbf{z}) - S_0 = 0 ,$$

$$DE(\mathbf{u}_c, \mathbf{z}, \mathbf{q}, \bar{p}) = 0 .$$

### Segunda formulação

Na versão discreta  $p$  é considerada linear nos elementos. Então, tem-se:

$$\|p\|_{L^2(\Gamma^*)}^2 = \frac{1}{3} \sum_{j=1}^{n_e} [p_1^j \cdot p_1^j + p_1^j \cdot p_2^j + p_2^j \cdot p_2^j], \quad (5.68)$$

onde  $p_1^j$  e  $p_2^j$  são os valores nodais de  $p$  no elemento  $j$ .

Chamando  $\mathbf{p}$  ao vetor dos valores nodais da função  $p$ , a equação de equilíbrio é definida neste caso como:

$$DF_i(\mathbf{u}_c, \mathbf{q}, \bar{p}, \mathbf{p}) = \int_{\Gamma^*} \left( \frac{1}{2\mu_0} \|\nabla\varphi\|^2 - \bar{p} + p \right) (V^i \cdot \mathbf{n}) d\Gamma + \mathcal{C}^i \cdot \hat{V}^i . \quad (5.69)$$

Com esta definição, a versão discreta da segunda formulação do problema inverso fica:

$$\min_{\mathbf{u}_c, \mathbf{q}, \bar{p}, \mathbf{p}} \|p\|_{L^2(\Gamma^*)}^2 ,$$

sujeito a:

$$\mathbf{A}\mathbf{q} - \mathbf{b}(\mathbf{u}_c) = 0 , \quad (5.70)$$

$$DF(\mathbf{u}_c, \mathbf{q}, \bar{p}, \mathbf{p}) = 0 .$$

## 5.4.2 Exemplos

Seis exemplos são analisados para ilustrar o comportamento das formulações propostas para o problema inverso. Os valores iniciais das variáveis de estado  $\mathbf{q}$  e  $\bar{p}$ , as variáveis de forma  $\mathbf{z}$  da primeira formulação e a pressão  $\mathbf{p}$  da segunda são definidos iguais a zero para todos os exemplos.

Para a solução dos problemas de otimização das Eqs. (5.67) e (5.70) o algoritmo FAIPA [95] foi utilizado. Algumas restrições de caixa nas variáveis relacionadas às posições dos indutores foram consideradas para evitar que os condutores entrem dentro do metal líquido ou se dirijam ao infinito. Estas restrições de caixa são facilmente incluídas e eficientemente tratadas pelo FAIPA.

### Exemplo 1

No primeiro exemplo são considerados quatro indutores e uma forma objetivo de área  $S_0 = \pi$ . A intensidade da corrente é  $I = 0.1$ , a tensão superficial é  $\sigma = 1.0 \times 10^{-4}$  e os coeficientes adimensionados  $\alpha_i$  tem valor absoluto igual a 1.0 com o sinal dado pela Figura 5.14.

Para fins de validação, a forma objetivo considerada neste exemplo é a solução do problema direto de superfície livre para posições dos indutores conhecidas. Após a resolução do problema direto a forma final obtida é considerada como forma objetivo para o problema inverso e as posições dos indutores são modificadas. Desta forma, as duas formulações para o problema inverso têm soluções conhecidas com valor nulo para a função objetivo.

A Figura 5.14.a apresenta as posições iniciais dos indutores e a forma objetivo. As posições finais dos indutores são mostradas pela Figura 5.14.b. Neste exemplo as duas formulações fornecem as mesmas posições para os indutores, em coincidência com a solução conhecida.

As Figuras 5.15 e 5.16 apresentam a evolução das funções objetivo durante o processo iterativo. Os valores das funções objetivo das duas formulações são zero no início devido à escolha dos valores iniciais dos vetores  $\mathbf{z}$  e  $\mathbf{p}$ . Por causa das restrições de igualdade, que não são verificadas no início, nas primeiras iterações tem-se um aumento do valor da função objetivo. Após algumas iterações, logo após atingir um valor máximo, a função objetivo decresce novamente. No final do processo iterativo as restrições de igualdade são

verificadas e a função objetivo converge para o valor ótimo.

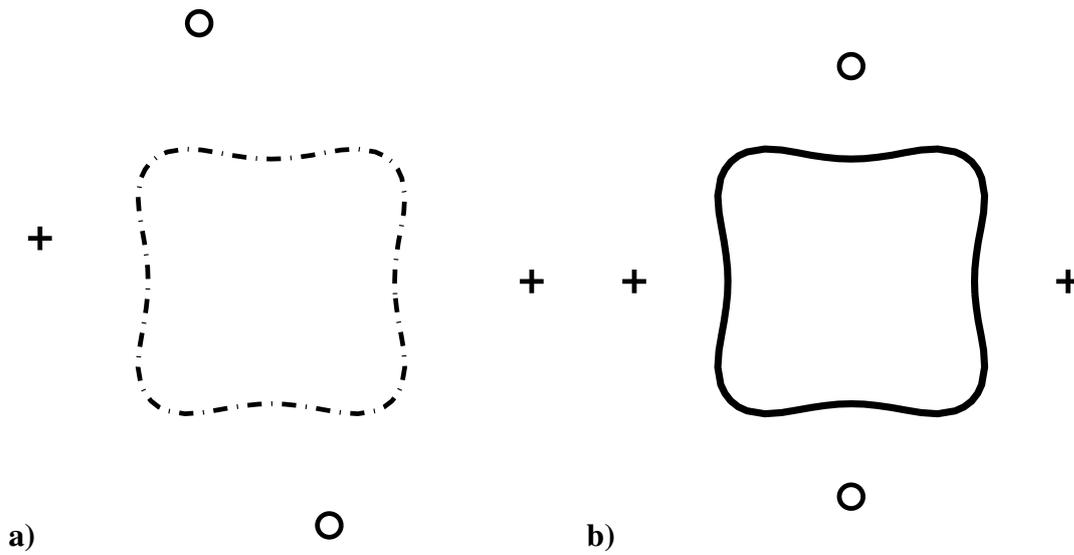


Figura 5.14: Exemplo 1, posição dos indutores, a) distribuição inicial, b) distribuição final. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

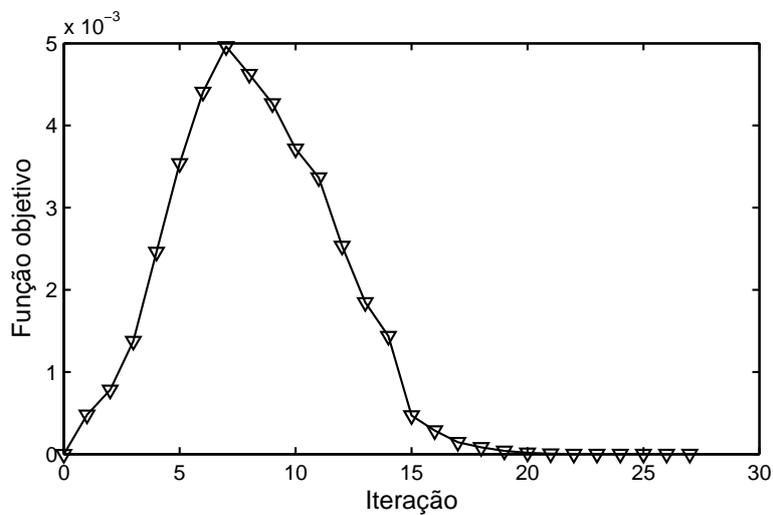


Figura 5.15: Exemplo 1, evolução da função objetivo da primeira formulação.

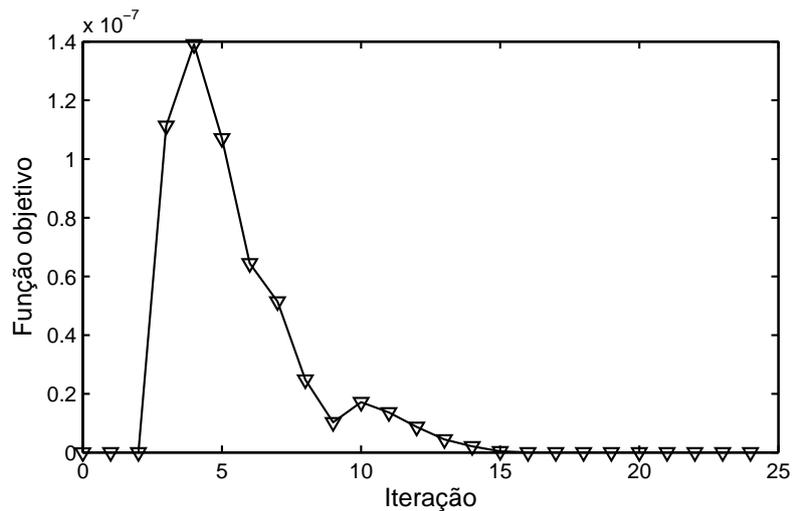


Figura 5.16: Exemplo 1, evolução da função objetivo da segunda formulação.

## Exemplo 2

No segundo exemplo a intensidade de corrente é modificada para o valor  $I = 0.04$ . Este valor é menor que o usado no primeiro exemplo. Então, os indutores vão se posicionar mais perto da superfície do metal líquido que no Exemplo 1. Neste caso o algoritmo de otimização não encontra uma solução que anule a função objetivo. Para a primeira formulação a forma final em equilíbrio não é exatamente igual à forma objetivo e na segunda formulação a forma objetivo não é equilibrada. Isto significa que, neste caso, a forma objetivo não é alcançável para a distribuição de corrente dada pelos quatro indutores.

A Figura 5.17.a apresenta a forma obtida e as posições dos indutores usando a primeira formulação. A Figura 5.17.b mostra os mesmos resultados para a segunda formulação. A forma em equilíbrio na Figura 5.17.b é a obtida após a solução do problema direto de superfície livre para a configuração de indutores obtida. As Figuras 5.18 e 5.19 apresentam a evolução das funções objetivo no processo de otimização.

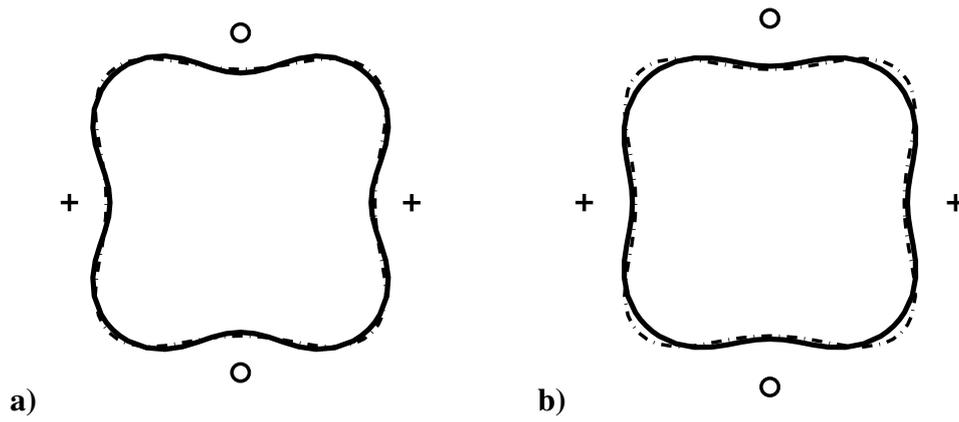


Figura 5.17: Exemplo 2, configuração de indutores e forma final, **a)** primeira formulação, **b)** segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

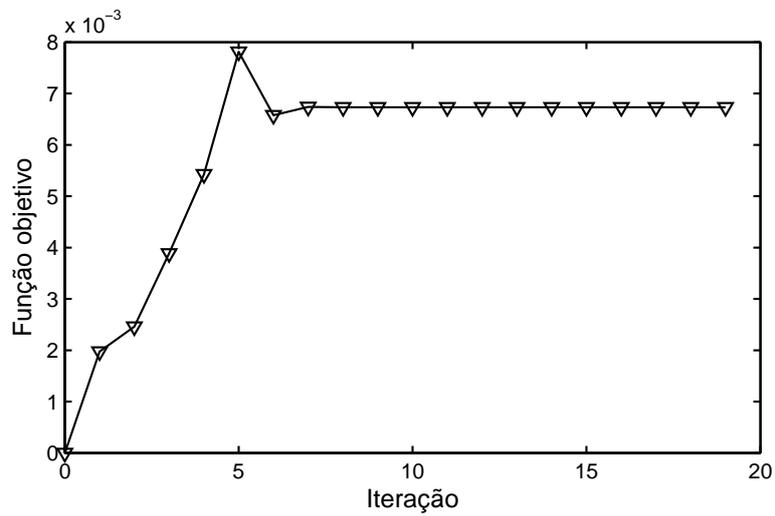


Figura 5.18: Exemplo 2, evolução da função objetivo da primeira formulação.

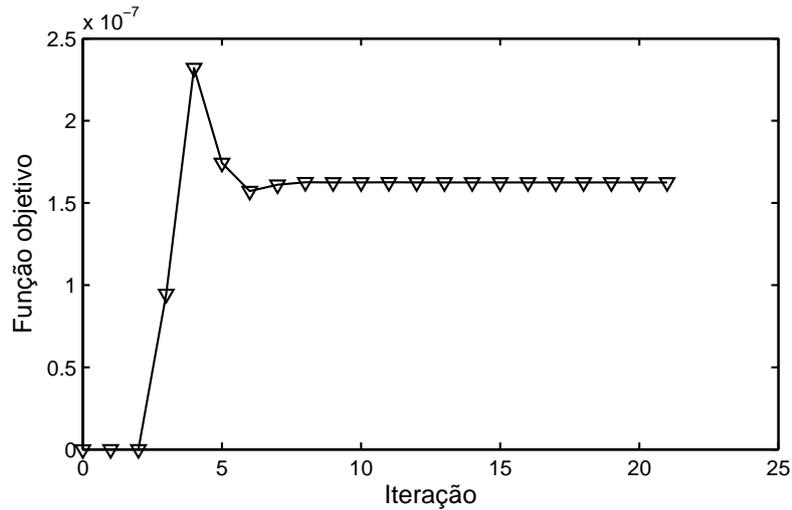


Figura 5.19: Exemplo 2, evolução da função objetivo da segunda formulação.

### Exemplo 3

Este exemplo é similar ao Exemplo 1, a forma objetivo é solução do problema direto para a mesma configuração de indutores mas para o valor da intensidade  $I = 0.2$ . Como no Exemplo 1 as posições dos indutores reproduzem exatamente as posições conhecidas. A Figura 5.20 apresenta a configuração inicial dos indutores, a forma objetivo e os resultados para as duas formulações. As Figuras 5.21 e 5.22 mostram a evolução das funções objetivo no processo de otimização.

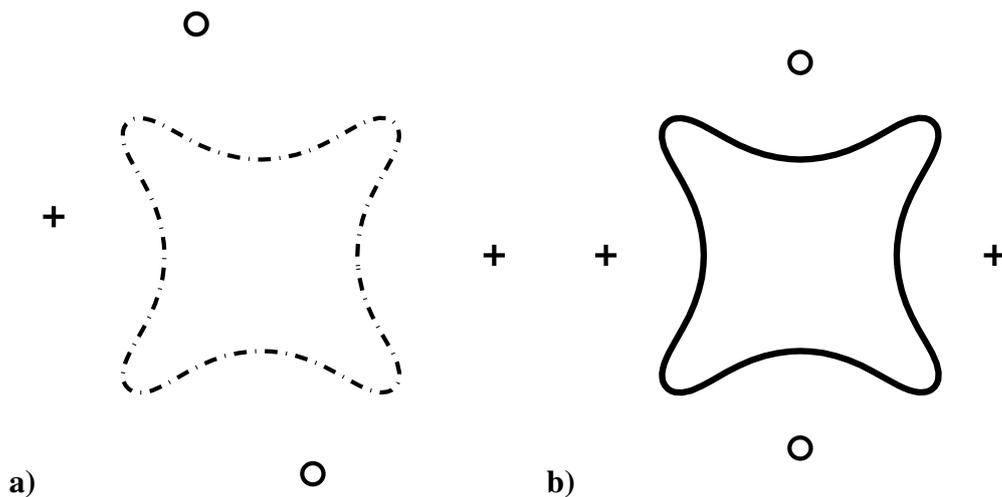


Figura 5.20: Exemplo 3, posição dos indutores, **a)** distribuição inicial, **b)** distribuição final. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

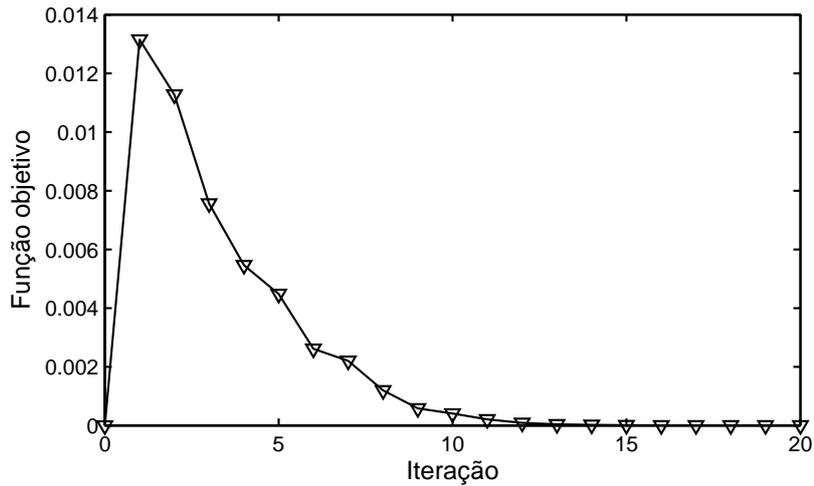


Figura 5.21: Exemplo 3, evolução da função objetivo da primeira formulação.

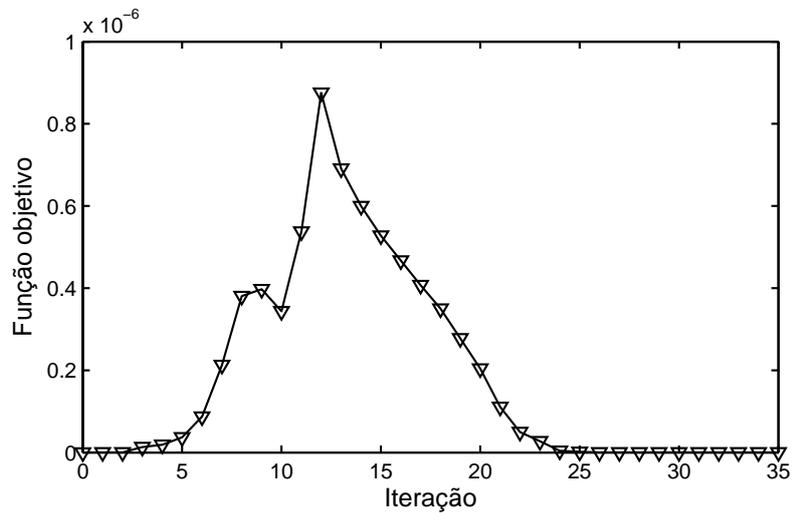


Figura 5.22: Exemplo 3, evolução da função objetivo da segunda formulação.

#### Exemplo 4

Este exemplo é similar ao anterior com a única modificação do valor da intensidade que agora é  $I = 0.1$ . Para este valor os indutores ficam mais próximos do metal líquido que no exemplo anterior. Como no Exemplo 1, a primeira formulação não obtém exatamente a forma objetivo e a segunda não obtém exatamente o equilíbrio para a forma objetivo.

A Figura 5.23a apresenta os resultados obtidos para este exemplo. As Figuras 5.24 e 5.25 mostram a evolução das funções objetivo no processo de otimização.

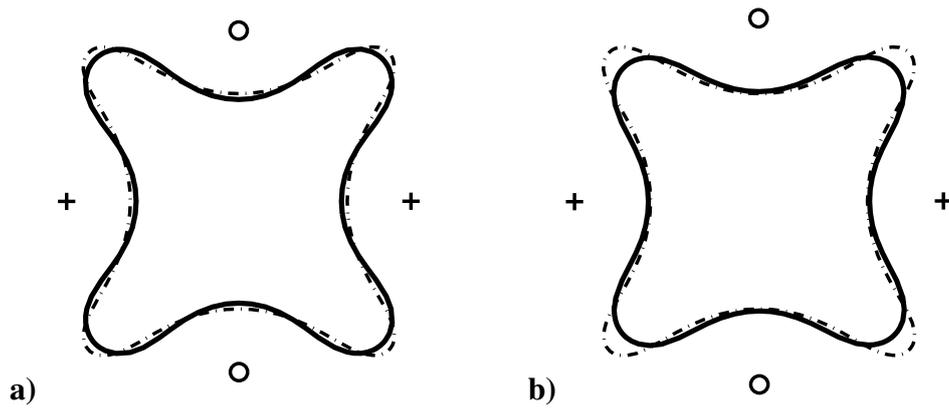


Figura 5.23: Exemplo 4, configuração de indutores e forma final, **a)** primeira formulação, **b)** segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

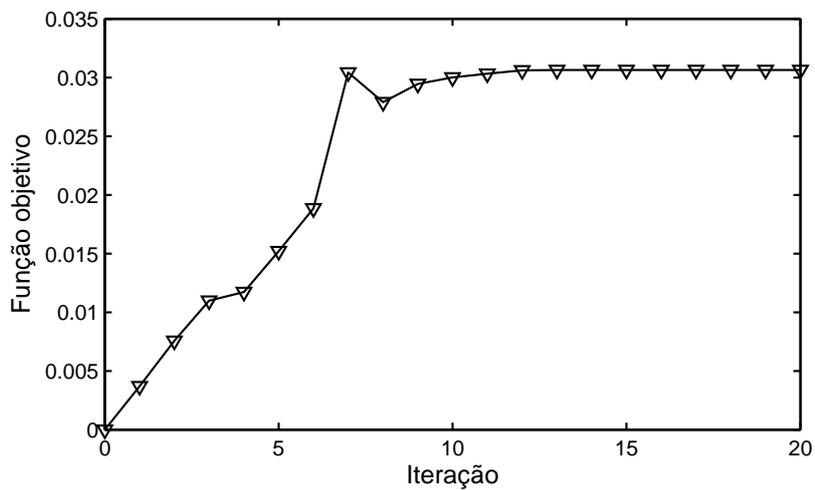


Figura 5.24: Exemplo 4, evolução da função objetivo da primeira formulação.

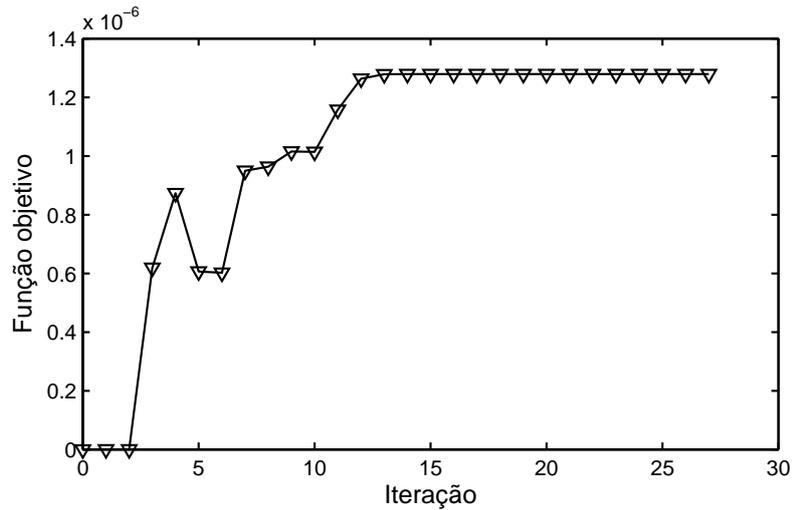


Figura 5.25: Exemplo 4, evolução da função objetivo da segunda formulação.

### Exemplo 5

Neste exemplo é procurada uma distribuição de indutores conveniente para construir uma barra de seção retangular de área  $S_0 = 4.98$ . Trinta e seis indutores são considerados, a intensidade é  $I = 0.075$  e a tensão superficial é  $\sigma = 1.0 \times 10^{-4}$ . Os coeficientes  $\alpha_i = \pm 1$  com o sinal dado pela Figura 5.26.

A Figura 5.27 apresenta as posições dos indutores obtidas para as duas formulações do problema inverso. A forma em equilíbrio representada na Figura 5.27.b é a solução do problema direto de superfície livre para a configuração de indutores obtida. As Figuras 5.28 e 5.29 mostram a evolução das funções objetivo no processo de otimização.

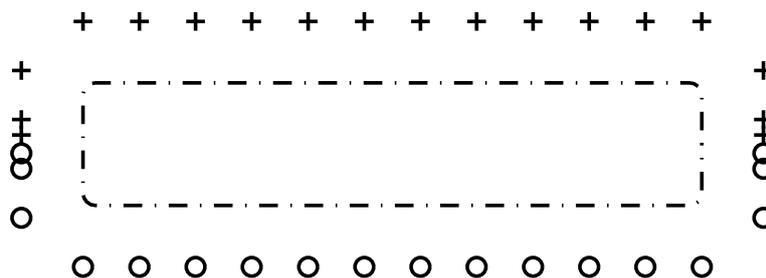


Figura 5.26: Exemplo 5, forma objetivo e configuração inicial de indutores. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

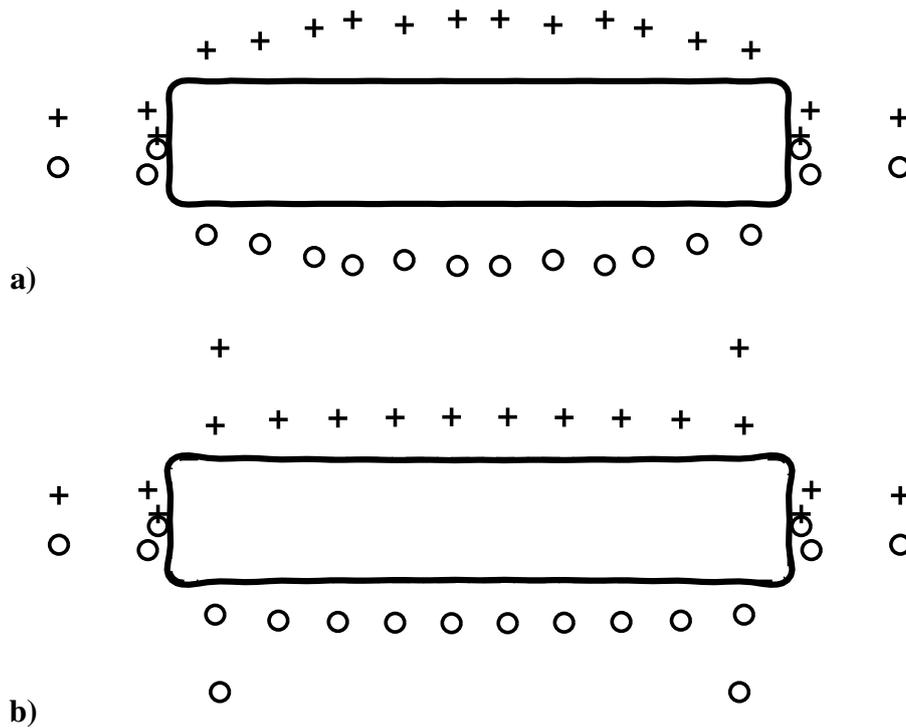


Figura 5.27: Exemplo 5, configuração final de indutores e forma final, **a)** primeira formulação, **b)** segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

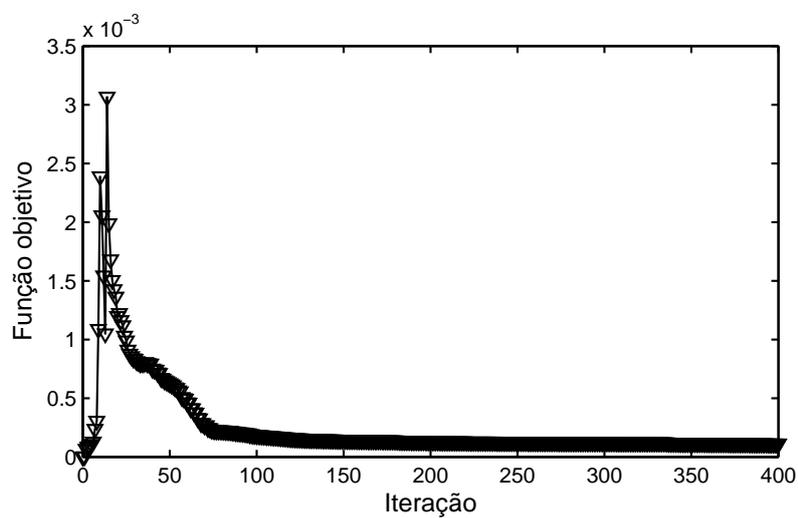


Figura 5.28: Exemplo 5, evolução da função objetivo da primeira formulação.

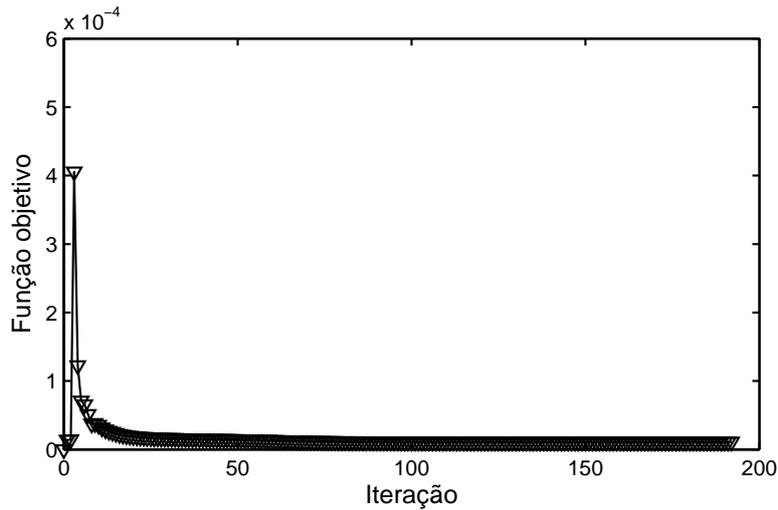


Figura 5.29: Exemplo 5, evolução da função objetivo da segunda formulação.

### Exemplo 6

Este último exemplo considera quarenta e oito indutores e uma forma objetivo de área  $S_0 = 30.0$  mostrada pela Figura 5.30. A intensidade da corrente é  $I = 0.03$ , a tensão superficial é  $\sigma = 1.0 \times 10^{-4}$  e os coeficientes adimensionados  $\alpha_i = \pm 1$ , com o sinal dado pela Figura 5.30.

A Figura 5.31 apresenta as posições obtidas para os indutores empregando as duas formulações do problema inverso. As Figuras 5.32 e 5.33 mostram a evolução das funções objetivo no processo de otimização.

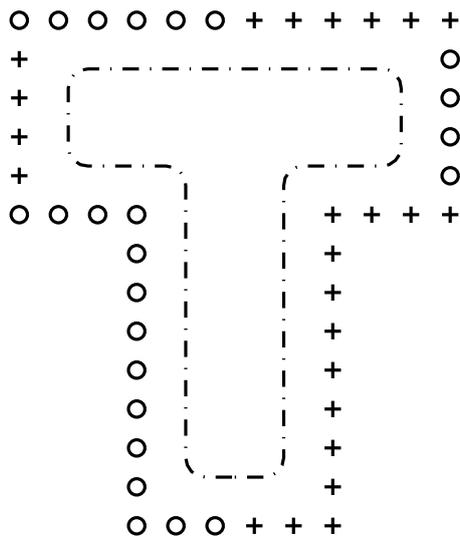


Figura 5.30: Exemplo 6, forma objetivo e configuração inicial de indutores. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

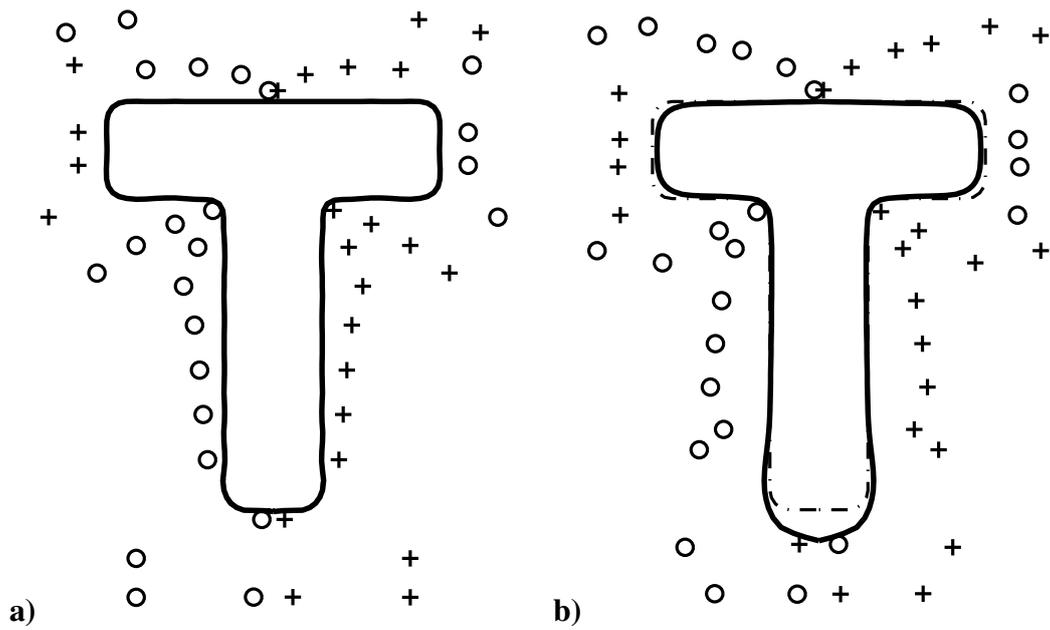


Figura 5.31: Exemplo 6, configuração final de indutores e forma final, **a)** primeira formulação, **b)** segunda formulação. Linha sólida: Forma em equilíbrio. Linha tracejada: Forma objetivo. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

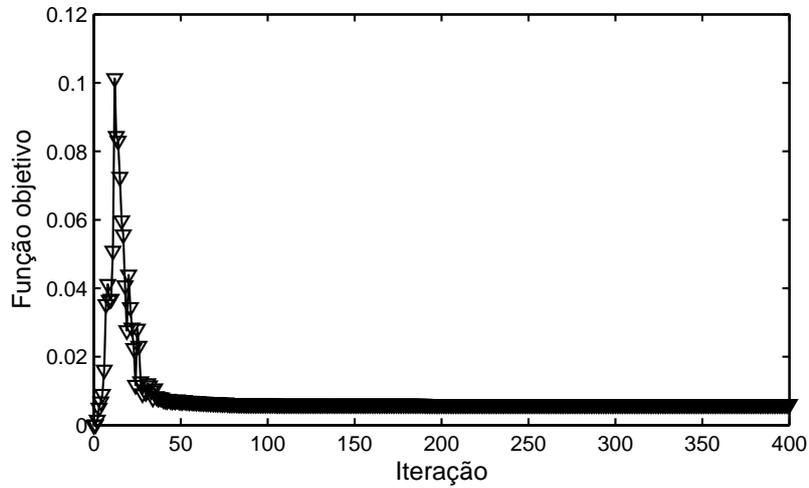


Figura 5.32: Exemplo 6, evolução da função objetivo da primeira formulação.

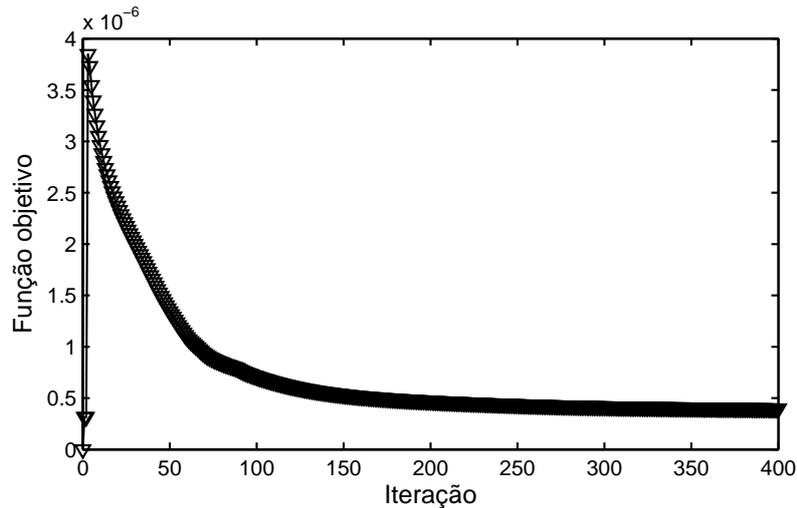


Figura 5.33: Exemplo 6, evolução da função objetivo da segunda formulação.

### Resumo dos resultados

A Tabela 5.1 resume os resultados obtidos para os exemplos considerados. Para cada exemplo é dado o número de nós da aproximação de elementos lineares do contorno objetivo  $\Gamma^*$ , o número de indutores, o número de iterações feito pelo algoritmo de otimização, e o valor final da função objetivo. A Tabela 5.1 mostra que o número total de iterações é similar para as duas formulações em todos os exemplos salvo no Exemplo 6 onde o número de iterações requerido pela primeira formulação foi menor ao requerido pela segunda. Para fins de comparação, o valor da função objetivo da primeira formulação é dado na tabela para as duas soluções. Na tabela pode ser observado que a primeira formulação obtém valores menores da função objetivo em todos os exemplos. A vantagem principal da segunda formulação é a de encontrar soluções similares às encontradas com a primeira formulação mas com um custo computacional menor, principalmente por causa da ausência de variáveis de forma. Portanto, a segunda formulação é indicada para encontrar uma solução inicial que pode, se necessário, ser melhorada em um processo posterior utilizando esta solução como ponto inicial para a primeira formulação.

Tabela 5.1: Resumo dos resultados obtidos para os exemplos de otimização da posição dos indutores.

Exemplos	Nós Indut.		Primeira formulação		Segunda formulação	
			IFA	OFA	IFA	OFA*
Ex1	72	4	27	2.403e-13	24	1.785e-17
Ex2	72	4	19	6.745e-03	21	2.261e-02
Ex3	120	4	20	2.181e-13	35	2.889e-15
Ex4	120	4	20	3.020e-02	27	8.831e-02
Ex5	136	36	400	1.103e-04	192	3.206e-03
Ex6	152	48	400	6.192e-03	400	1.713e+00

IFA: Iterações realizadas pelo FAIPA.

OFA: Função objetivo alcançada pelo FAIPA.

(\*): Resultado para a função objetivo da primeira formulação.

## 5.5 Otimização da forma dos indutores

Neste caso a área da seção transversal dos indutores não é considerada desprezível. Assim, as seções transversais dos indutores são representadas pelos domínios  $\Theta_i$ , como mostrado pela Figura 5.34.

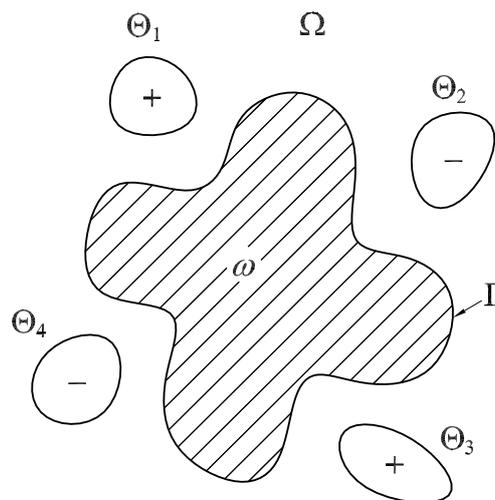


Figura 5.34: Problema de otimização de forma dos indutores.

A densidade de corrente elétrica  $j_0$  é assumida da forma:

$$j_0 = I \sum_{i=1}^{n_c} \alpha_i \chi_{\Theta_i}, \quad (5.71)$$

onde  $I$  é uma dada intensidade de corrente,  $\Theta_i$ ,  $1 \leq i \leq n_c$ , são domínios em  $\mathbb{R}^2$ ,  $\chi_{\Theta_i}$  são suas funções características, e  $\alpha_i$ ,  $1 \leq i \leq n_c$ , são coeficientes adimensionados. O problema inverso consiste em determinar os domínios  $\Theta_i$ .

Note que a Eq. (5.71) assume que a densidade de corrente elétrica é uniforme em cada região  $\Theta_i$ . Indutores formados por múltiplos filamentos isolados uns dos outros permitem o uso da Eq. (5.71) como uma boa aproximação, ver [96]. Esta técnica permite também a fabricação de indutores de formas específicas.

Neste caso a expressão para a função  $\varphi_1$  é:

$$\varphi_1(x) = -\frac{\mu_0 I}{2\pi} \sum_{i=1}^{n_c} \alpha_i \int_{\Theta_i} \ln \|x - y\| d\Omega_y. \quad (5.72)$$

Ela pode ser calculada como uma soma de integrais nos contornos  $\Gamma_i$  dos domínios  $\Theta_i$ . Considere o campo vetorial  $\mathbf{w} : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida como:

$$\mathbf{w}(x, y) = (1/4)(1 - 2 \ln \|x - y\|)(x - y). \quad (5.73)$$

A divergência de  $\mathbf{w}$  é  $\text{div}_y \mathbf{w} = \ln \|x - y\|$ . Portanto, a Eq. (5.72) fica:

$$\varphi_1(x) = -\frac{\mu_0 I}{2\pi} \sum_{i=1}^{n_c} \alpha_i \int_{\Gamma_i} \mathbf{w} \cdot \mathbf{n} d\Gamma_y. \quad (5.74)$$

### 5.5.1 Indutores

A seção transversal dos indutores corresponderá a uma das formas paramétricas representadas na Figura 5.35. A contribuição de cada indutor à função  $\varphi_1$  é calculada pela Eq. (5.74). O contorno  $\Gamma_i$  de cada indutor é dividido em elementos lineares para o cálculo numérico da integral. O vetor de todos os parâmetros de forma dos indutores é denotado  $\mathbf{u}_c$ . Note que a condição dada pela Eq. (5.5) não é satisfeita de forma automática neste caso. Para uma densidade de corrente dada pela Eq. (5.71), a condição dada pela Eq. (5.5) deve ser satisfeita explicitamente impondo:

$$\sum_{i=1}^{n_c} \alpha_i \int_{\Theta_i} d\Omega = 0.$$

Chamando  $S_{\Theta}(\mathbf{u}_c) = \sum_{i=1}^{n_c} \alpha_i \int_{\Theta_i} d\Omega$ , tem-se a restrição de igualdade:

$$S_{\Theta}(\mathbf{u}_c) = 0. \quad (5.75)$$

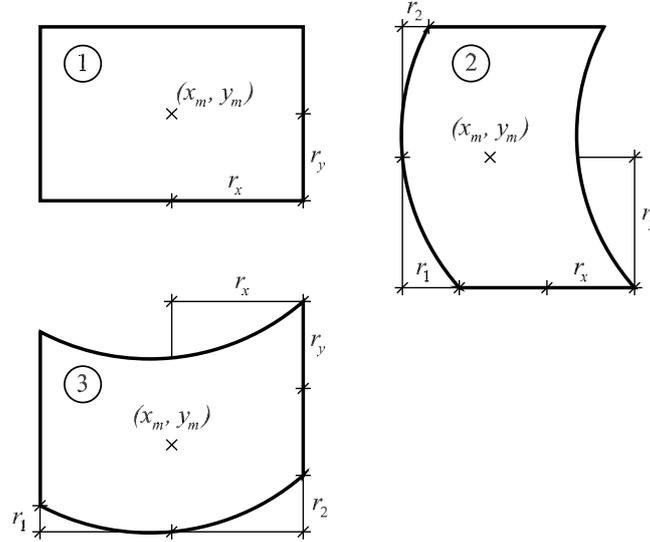


Figura 5.35: Forma paramétrica da seção transversal dos indutores. **1.** indutor retangular de quatro parâmetros; **2.** indutor de seis parâmetros com lados verticais parabólicos; **3.** indutor de seis parâmetros com lados horizontais parabólicos.

### 5.5.2 Restrições geométricas

Nesta seção é descrito um procedimento para evitar a superposição entre os domínios ocupados pelo metal líquido e pelos indutores. O mesmo consiste em considerar as seguintes restrições de desigualdade:

$$\psi(x_j) \leq \psi_0, \quad \forall x_j \in \mathcal{X}, \quad (5.76)$$

onde  $\mathcal{X}$  é um conjunto de pontos escolhidos no contorno dos indutores. A função escalar  $\psi$  se anula no interior do metal líquido e é negativa no exterior. Portanto, escolhendo um valor negativo para o parâmetro  $\psi_0$ , a Eq. (5.76) obriga aos pontos  $x_j$  a permanecerem no exterior do metal líquido como mostrado na Figura 5.36.

A função  $\psi$  proposta é definida pela solução do problema:

$$\begin{aligned} \Delta\psi(x) &= 0 && \text{in } \Omega^*, \\ \psi(x) &= 0 && \text{on } \Gamma^*, \\ \int_{\Gamma^*} \nabla\psi(x) \cdot \mathbf{n} \, d\Gamma &= -1. \end{aligned} \quad (5.77)$$

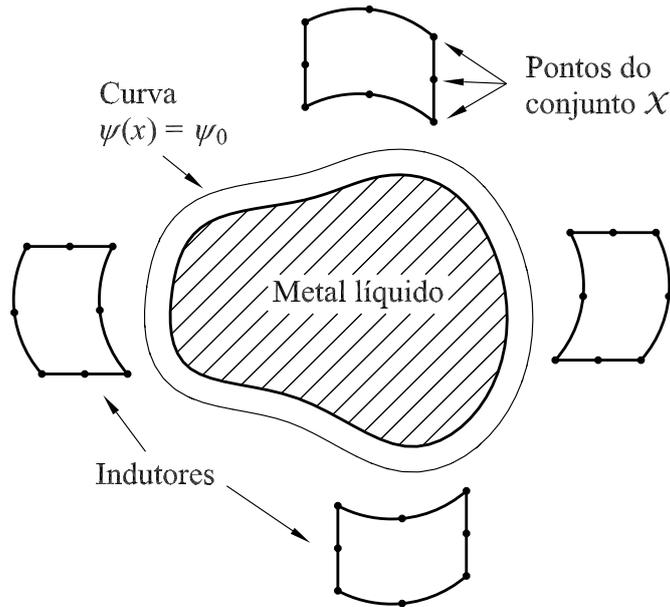


Figura 5.36: Restrições Geométricas.

De forma similar à função  $v$  descrita na Seção 5.2,  $\psi$  pode ser calculada como:

$$\psi(x) = -\frac{1}{2\pi} \int_{\Gamma} q(y) \ln \|x - y\| d\Gamma + c, \quad (5.78)$$

onde  $q$  deve satisfazer:

$$\int_{\Gamma} q(y) d\Gamma = -1. \quad (5.79)$$

De igual forma que na Seção 5.2, uma solução aproximada para  $q$  e  $c$  pode ser obtida resolvendo um sistema linear similar ao da Eq. (5.37). A aproximação numérica da função  $\psi$  é obtida empregando a Eq. (5.78).

O valor de  $\psi_0$  é definido escolhendo um ponto no exterior do metal líquido e calculando o valor de  $\psi$  nesse ponto. Nas Figuras 5.37 e 5.38 é mostrado o gráfico da função  $\psi$  para dois formas objetivo diferentes.

Definindo  $\psi_j(\mathbf{u}_c) = \psi(x_j(\mathbf{u}_c)) - \psi_0$ , todas as restrições geométricas são expressadas na forma:

$$\psi(\mathbf{u}_c) \leq 0. \quad (5.80)$$

Note que a função  $\psi$  foi definida para o domínio objetivo  $\Omega^*$ . Portanto, no sentido exato, as restrições dadas pela Eq. (5.76) são suficientes somente para a segunda formulação do problema inverso. No caso da primeira formulação  $\psi$  deveria ser definida para

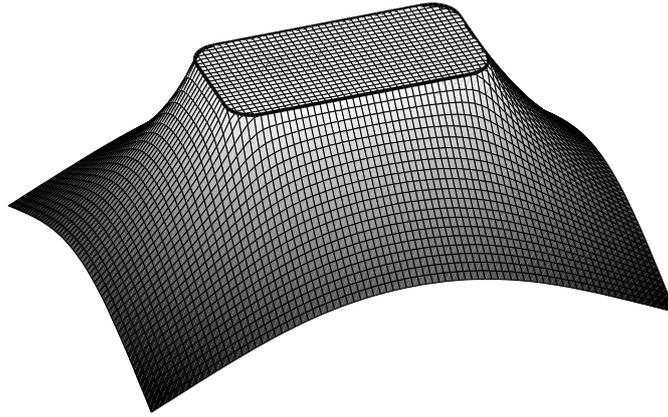


Figura 5.37: Função  $\psi$  para uma forma objetivo aproximadamente retangular.

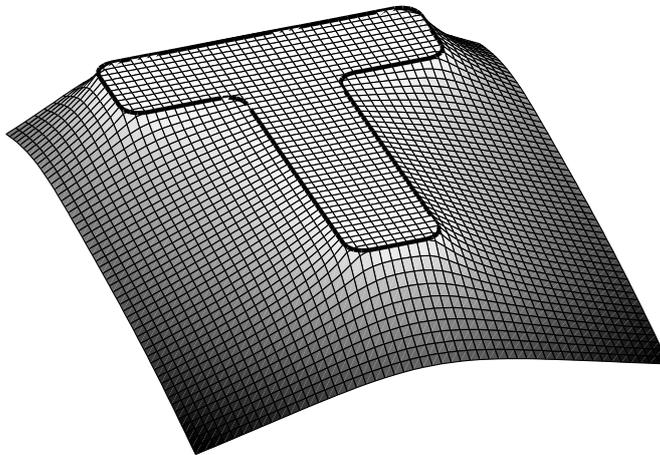


Figura 5.38: Função  $\psi$  para uma forma objetivo similar à letra “T”.

o domínio  $\Omega_Z$ . Nesse caso existe um custo computacional adicional associado ao cálculo de  $q$  e  $c$  cada vez que o domínio  $\Omega_Z$  é atualizado. Contudo, para todos os exemplos apresentados na Seção 5.5.4 foi observado que a função  $\psi$  definida para o domínio fixo  $\Omega^*$  é suficiente para evitar a superposição dos domínios. A razão é que o domínio ocupado pelo metal líquido mantém-se muito próximo do domínio objetivo ao longo de todo o processo de otimização. Somente os indutores experimentam grandes variações de forma, porém, eles são mantidos afastados da forma objetivo de forma eficaz pelas restrições dadas pela Eq. (5.76).

### 5.5.3 Problema discreto

A discretização do domínio ocupado pelo metal líquido é feita da mesma forma que na Seção 5.4.1. Portanto, a diferença do modelo atual com relação ao descrito na Se-

ção 5.4.1 consiste unicamente na definição da geometria dos indutores, a restrição na intensidade total e a restrições geométricas de não superposição dos domínios ocupados pelos indutores e pelo metal líquido.

### Primeira formulação

Neste caso, a versão discreta da primeira formulação do problema inverso fica:

$$\begin{aligned}
 & \min_{\mathbf{u}_c, \mathbf{z}, \mathbf{q}, \bar{p}} \|Z(\mathbf{z})\|_{L^2(\Gamma^*)}^2, \\
 & \text{sujeito a:} \\
 & \mathbf{A}(\mathbf{z})\mathbf{q} - \mathbf{b}(\mathbf{u}_c, \mathbf{z}) = 0, \\
 & S(\mathbf{z}) - S_0 = 0, \\
 & S_{\Theta}(\mathbf{u}_c) = 0, \\
 & DE(\mathbf{u}_c, \mathbf{z}, \mathbf{q}, \bar{p}) = 0, \\
 & \psi(\mathbf{u}_c) \leq 0.
 \end{aligned} \tag{5.81}$$

### Segunda formulação

Para a segunda formulação a versão discreta do problema inverso fica:

$$\begin{aligned}
 & \min_{\mathbf{u}_c, \mathbf{q}, \bar{p}, \mathbf{p}} \|p\|_{L^2(\Gamma^*)}^2, \\
 & \text{sujeito a:} \\
 & \mathbf{A}\mathbf{q} - \mathbf{b}(\mathbf{u}_c) = 0, \\
 & S_{\Theta}(\mathbf{u}_c) = 0, \\
 & DF(\mathbf{u}_c, \mathbf{q}, \bar{p}, \mathbf{p}) = 0, \\
 & \psi(\mathbf{u}_c) \leq 0.
 \end{aligned} \tag{5.82}$$

## 5.5.4 Exemplos

Vários exemplos são considerados para ilustrar o comportamento das formulações propostas para o problema inverso. A forma objetivo, a tensão superficial  $\sigma$ , a intensidade  $I$  e os coeficientes adimensionados  $\alpha_i$  são dados. Os parâmetros, incluindo os valores  $\psi_0$  das restrições geométricas, são iguais para todos os exemplos e para as duas formulações do problema inverso. Os valores iniciais das variáveis de estado  $\mathbf{q}$  e  $\bar{p}$ , as variáveis de

forma  $z$  da primeira formulação e as pressões  $\mathbf{p}$  da segunda são definidas nulas para todos os exemplos.

Os algoritmos FAIPA e FDIPA-H são utilizados para resolver os problemas de otimização. Como critério de parada o ponto atual é aceito se a norma das restrições de igualdade é menor que  $1.0 \times 10^{-8}$ , e se um dos valores seguintes, a norma do gradiente da função lagrangiana ou a norma da direção  $\mathbf{d}_x^\alpha$ , é menor que  $1.0 \times 10^{-5}$ .

Para cada exemplo é mostrada a posição inicial, a forma objetivo e a forma dos indutores obtida pelo algoritmo de otimização. A evolução da função objetivo ao longo das iterações é mostrada em escala logarítmica para melhor visualização das diferentes curvas.

### **Exemplo 1**

A forma objetivo deste exemplo é a solução de um problema direto considerando quatro intensidades concentradas de valor  $I = 0.1$ , com o sinal dado pela Figura 5.39. Ver o Exemplo 1 da Seção 5.4.2.

Quatro indutores de tipo 1, segundo a Figura 5.35, são considerados para o problema inverso. A superfície da forma objetivo tem área  $S_0$  igual a  $\pi$ , a tensão superficial  $\sigma$  é  $1.0 \times 10^{-4}$ , a intensidade  $I$  é igual a 0.1, e os coeficientes adimensionados  $\alpha_p$  têm valor absoluto igual a 4.0 com o sinal dado pela Figura 5.40. Duas configurações para a posição inicial dos indutores são consideradas como mostrado pela Figura 5.40.

As soluções calculadas coincidem para as duas formulações do problema inverso e para as duas configurações iniciais. A forma em equilíbrio é praticamente a mesma em todos os casos, coincidindo perfeitamente com a forma objetivo, sem nenhuma restrição geométrica ativa na solução. A Figura 5.41 apresenta os indutores obtidos e algumas curvas de nível da função de fluxo magnético  $\varphi$ . As Figuras 5.42 e 5.43 mostram a evolução das funções objetivo durante as iterações do algoritmo.

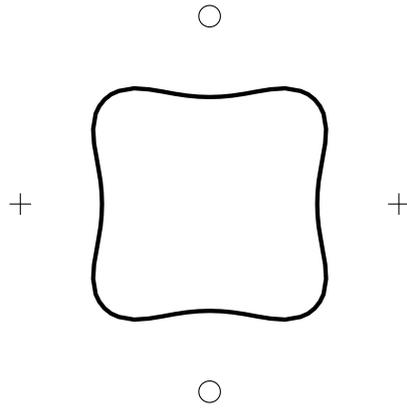


Figura 5.39: Exemplo 1, forma objetivo em equilíbrio com intensidades concentradas. Linha sólida: Forma em equilíbrio. Sinal de adição: Corrente positiva. Círculo: Corrente negativa.

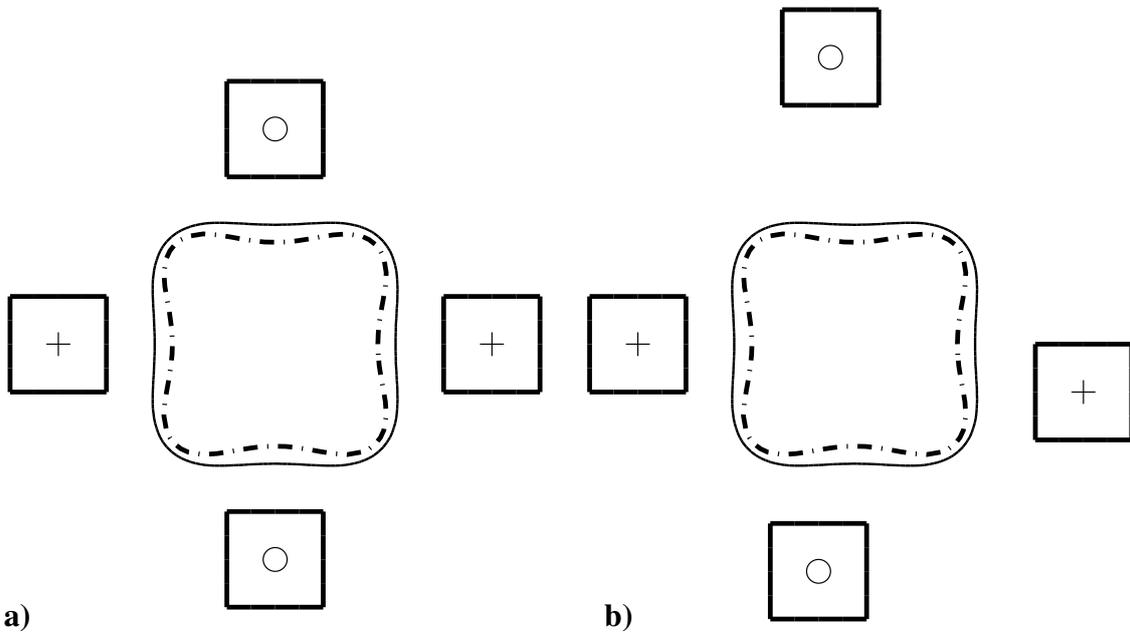


Figura 5.40: Exemplo 1, configuração inicial e restrições geométricas dos exemplos Ex1a e Ex1b, **a)** exemplo Ex1a. **b)** exemplo Ex1b. Linha sólida: Curva  $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa.

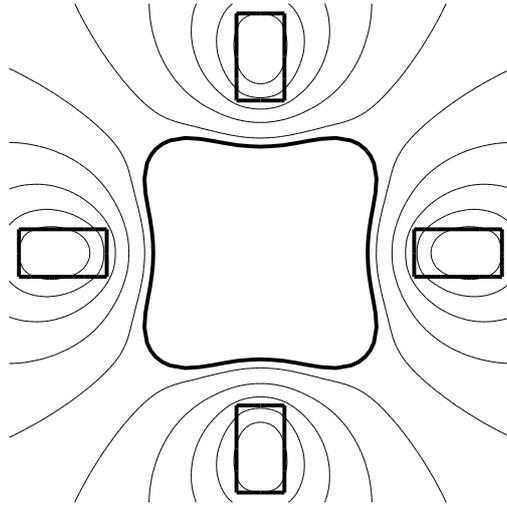


Figura 5.41: Solução do Exemplo 1, ambas as formulações, forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

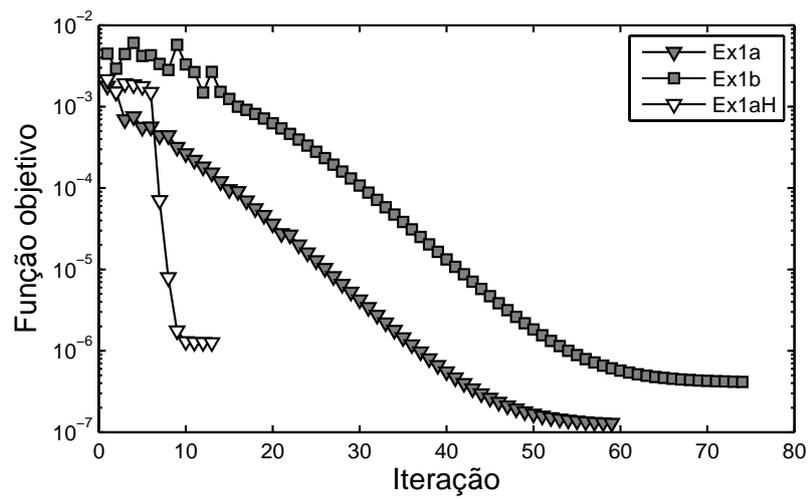


Figura 5.42: Exemplo 1, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

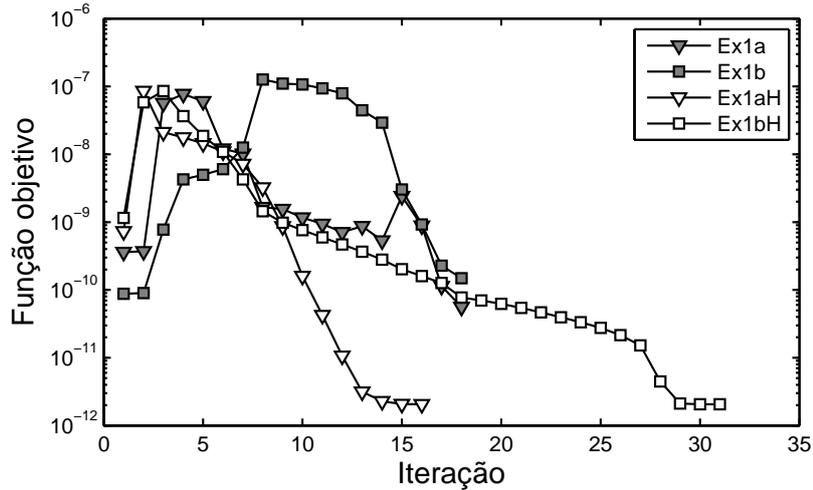


Figura 5.43: Exemplo 1, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

## Exemplo 2

Neste exemplo a forma objetivo é o quadrado arredondado mostrado pela Figura 5.44. Quatro indutores de tipo 1 segundo a Figura 5.35 são considerados. A área da forma objetivo  $S_0$  é igual a 3.86 e a tensão superficial  $\sigma$  é igual a  $1.0 \times 10^{-4}$ . A intensidade  $I$  é igual a 0.1 e os coeficientes adimensionados  $\alpha_p$  têm valor absoluto igual a 4.0 com o sinal dado pela Figura 5.44. Quatro valores diferentes para o parâmetro  $\psi_0$  das restrições geométricas são considerados, estes valores geram quatro problemas diferentes chamados Ex2a a Ex2d, como mostrado pela Figura 5.44. O exemplo Ex2a corresponde à restrição dada pela curva mais próxima à forma objetivo enquanto que o exemplo Ex2d corresponde à curva mais afastada.

Para a primeira formulação, a mesma configuração de indutores foi obtida para todos os casos como mostrado pela Figura 5.45.a; a forma em equilíbrio e algumas curvas de nível da função de fluxo magnético  $\varphi$  são apresentadas pela Figura 5.45.b. Para a primeira formulação, as restrições geométricas não resultaram ativas em nenhum dos exemplos. Para a segunda formulação, a configuração final de indutores depende do valor do parâmetro  $\psi_0$ . Para o maior valor, a Figura 5.46.a apresenta os indutores obtidos e a Figura 5.46.b mostra a forma em equilíbrio e algumas curvas de nível da função de fluxo magnético  $\varphi$ . A Figura 5.47 mostra o mesmo para o menor valor de  $\psi_0$ . Ao contrário da primeira formulação, a segunda posiciona os indutores tão próximos quanto possível

ao metal líquido. Nos quatro casos as restrições geométricas resultam ativas na solução. Embora a localização dos indutores é bastante diferente usando uma ou outra formulação, o valor ótimo da função objetivo da primeira formulação é praticamente o mesmo para todos os resultados obtidos. A evolução das funções objetivo durante o processo iterativo é mostrada pelas Figuras 5.48 e 5.49.

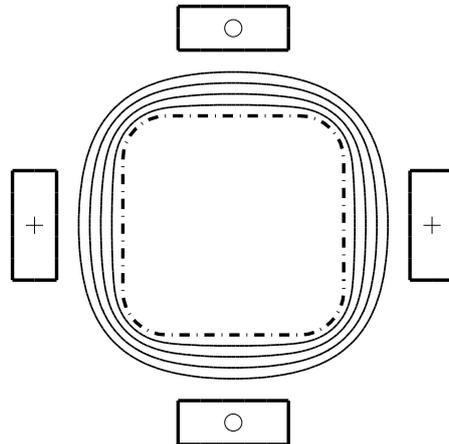


Figura 5.44: Exemplo 2, configuração inicial e restrições geométricas dos exemplos Ex2a a Ex2d. Linha sólida: Curva  $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa.

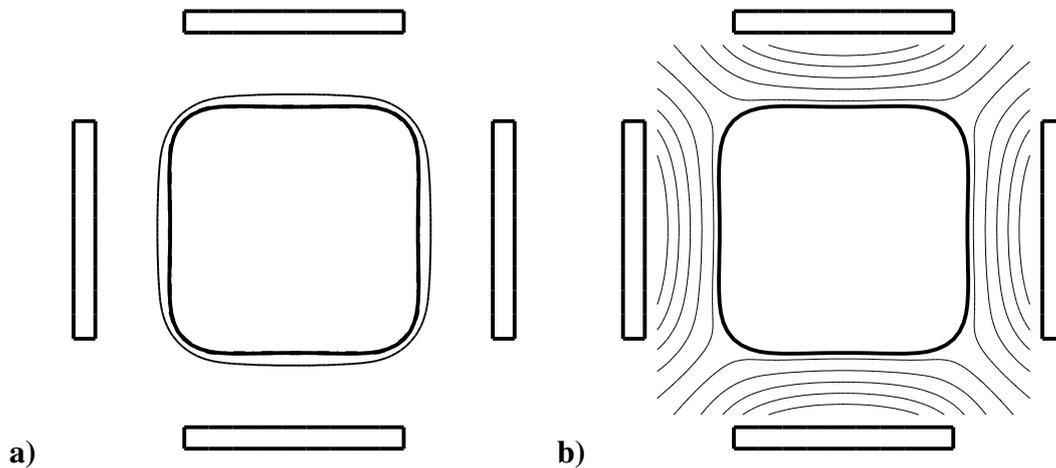


Figura 5.45: Solução do exemplo Ex2a, primeira formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

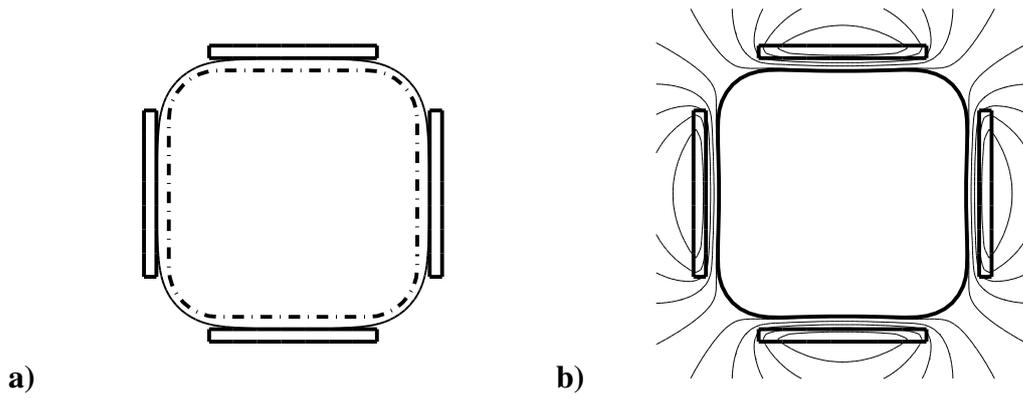


Figura 5.46: Solução do exemplo Ex2a, segunda formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

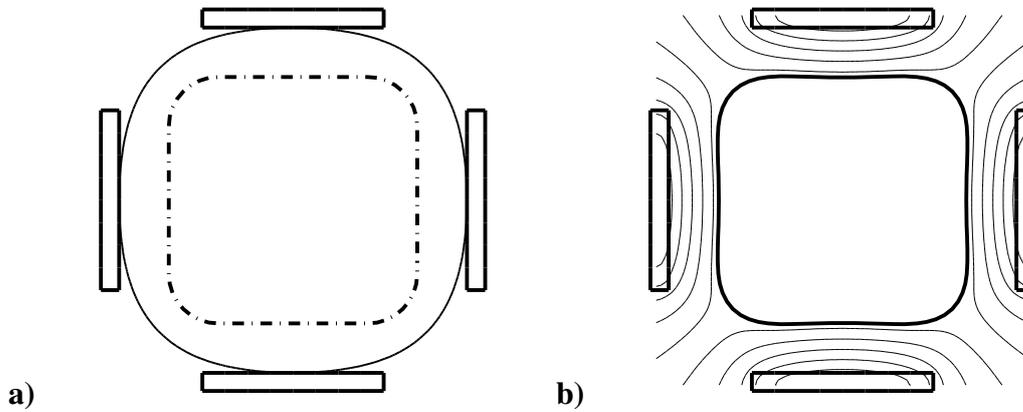


Figura 5.47: Solução do exemplo Ex2d, segunda formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

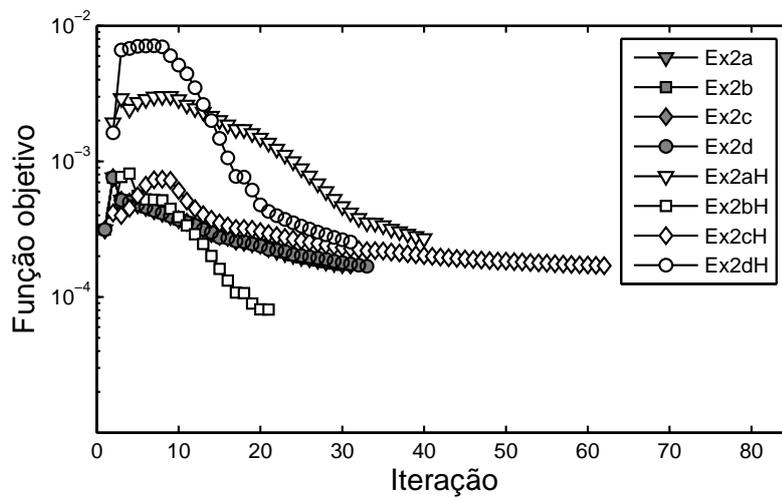


Figura 5.48: Exemplo 2, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

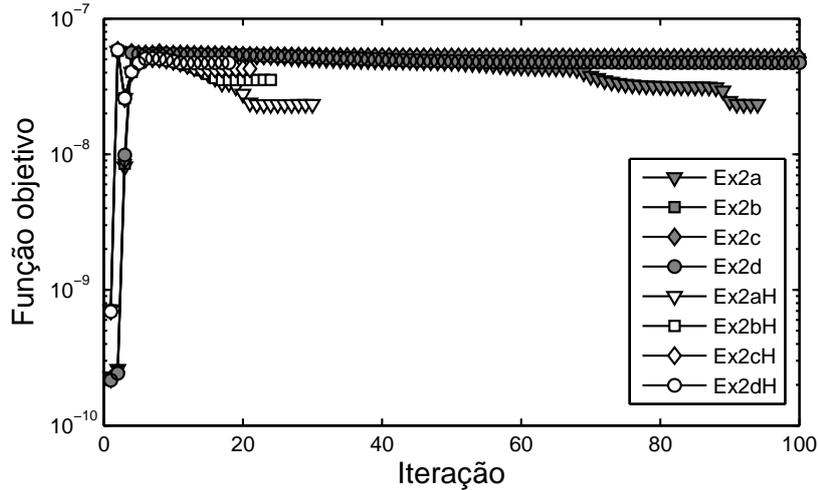


Figura 5.49: Exemplo 2, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

### Exemplo 3

A forma objetivo deste exemplo é o retângulo arredondado mostrado pela Figura 5.50. Para o problema inverso são considerados oito indutores de tipo 1 segundo a Figura 5.35. A forma objetivo tem área  $S_0$  igual a 7.86, a tensão superficial  $\sigma$  é igual a  $1.0 \times 10^{-4}$ , a intensidade  $I$  é igual a 0.1, e os coeficientes adimensionados  $\alpha_p$  têm valor absoluto igual a 4.0 com o sinal dado pela Figura 5.50. As soluções para dois valores do parâmetro  $\psi_0$ , exemplos Ex3a e Ex3b, são comparadas.

Neste exemplo, a maior diferença entre a solução de uma e outra formulação é o tamanho dos indutores localizados à esquerda e à direita do metal líquido. Este tamanho depende fortemente do valor do parâmetro  $\psi_0$  no caso da primeira formulação, fato que não acontece quando é empregada a segunda formulação. Para o maior valor de  $\psi_0$  e para a primeira formulação, a Figura 5.51.a mostra os indutores obtidos e a Figura 5.51.b apresenta a forma em equilíbrio e algumas curvas de nível da função de fluxo magnético  $\varphi$ . As Figuras 5.52.a e 5.52.b mostram o mesmo para a segunda formulação. As Figuras 5.53 e 5.54 apresentam a evolução das funções objetivo durante o processo iterativo.

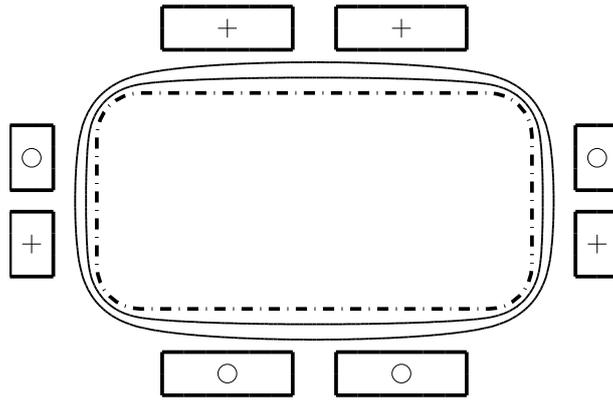


Figura 5.50: Exemplo 3, configuração inicial e restrições geométricas para os exemplos Ex3a e Ex3b. Linha sólida: Curva  $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa.

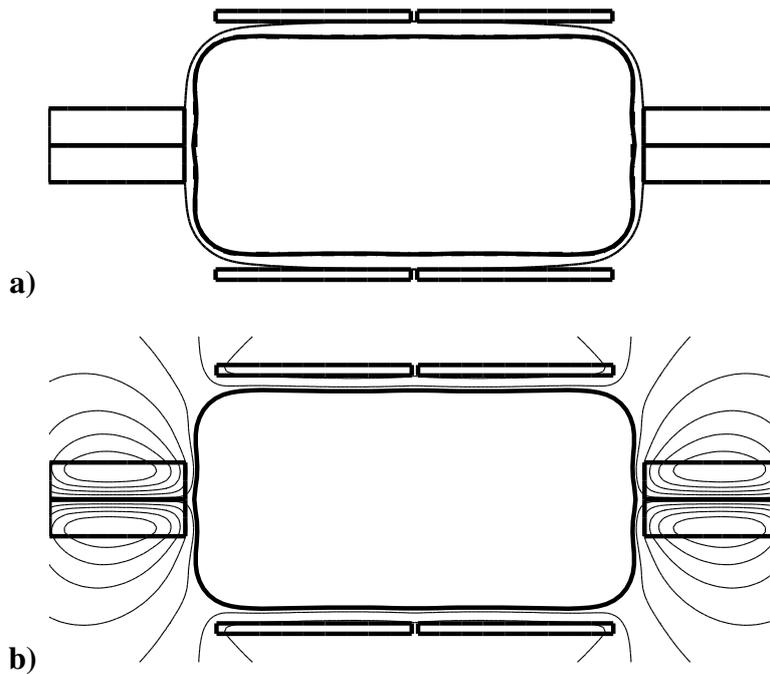


Figura 5.51: Solução do exemplo Ex3a, primeira formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

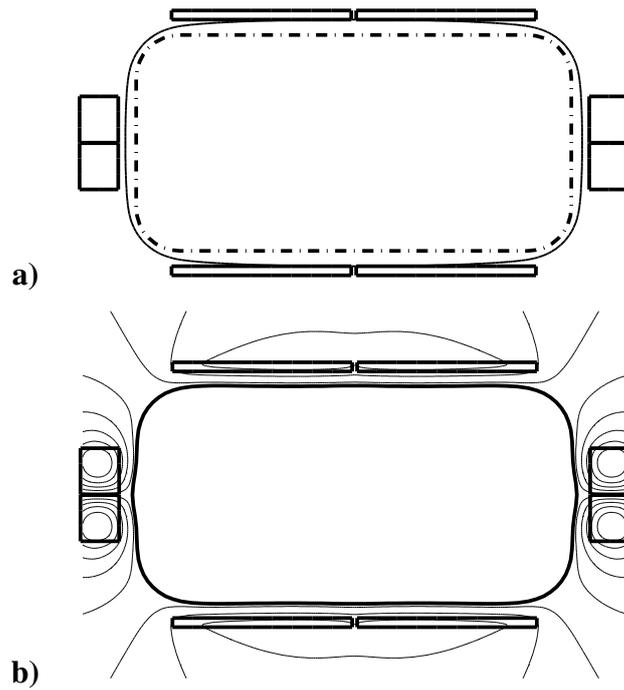


Figura 5.52: Solução do exemplo Ex3b, segunda formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

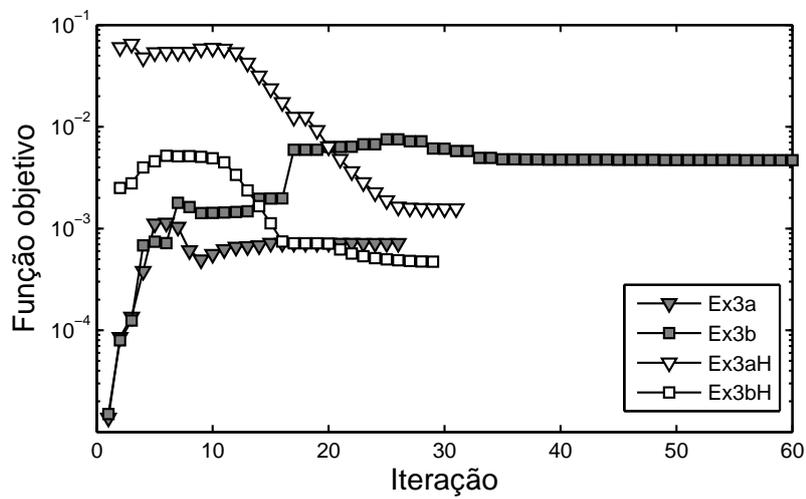


Figura 5.53: Exemplo 3, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

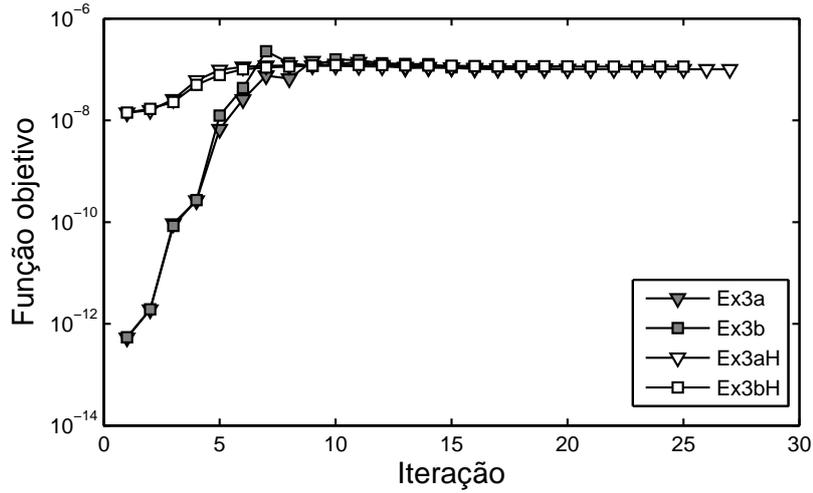


Figura 5.54: Exemplo 3, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

#### Exemplo 4

Este exemplo tem, como única diferença em relação com o anterior, o sinal dos coeficientes  $\alpha_p$  como mostrado pela Figura 5.55. Para o maior valor do parâmetro  $\psi_0$  e para a primeira formulação, a Figura 5.56.a apresenta os indutores obtidos e a Figura 5.56.b mostra a forma em equilíbrio e algumas curvas de nível da função de fluxo magnético  $\varphi$  na solução do problema. As Figuras 5.57.a e 5.57.b apresentam o mesmo para a segunda formulação. A evolução das funções objetivo durante o processo iterativo é mostrada pelas Figuras 5.58 e 5.59.

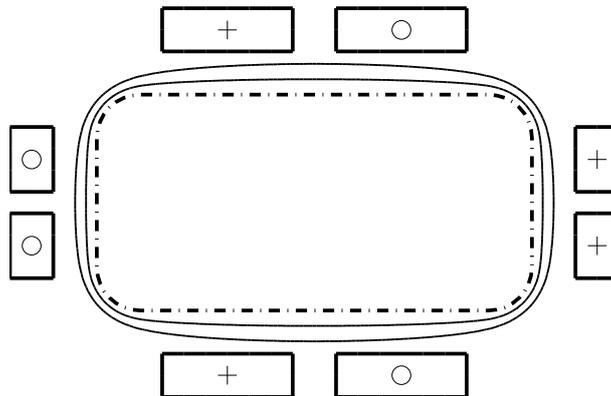


Figura 5.55: Exemplo 4, configuração inicial e restrições geométricas para os exemplos Ex4a e Ex4b. Linha sólida: Curva  $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa.

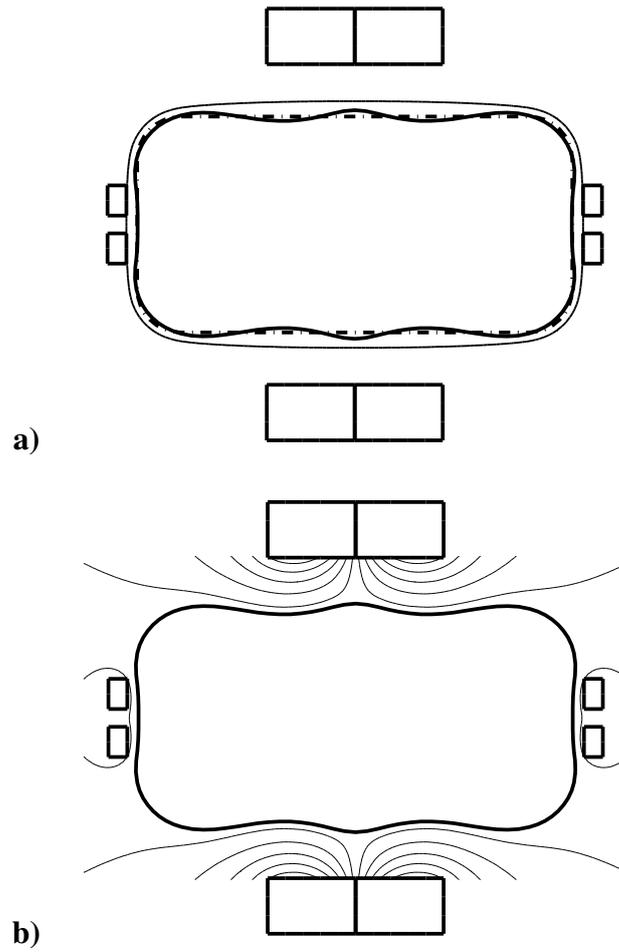


Figura 5.56: Solução do exemplo Ex4a, primeira formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

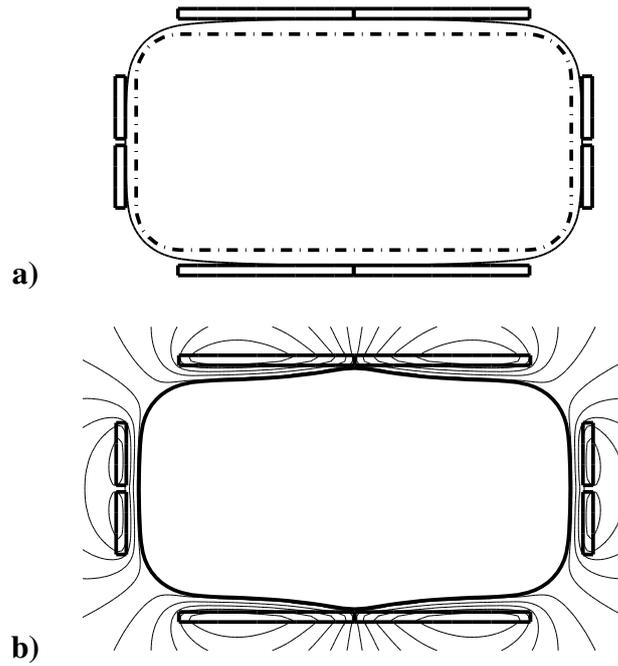


Figura 5.57: Solução do exemplo Ex4a, segunda formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

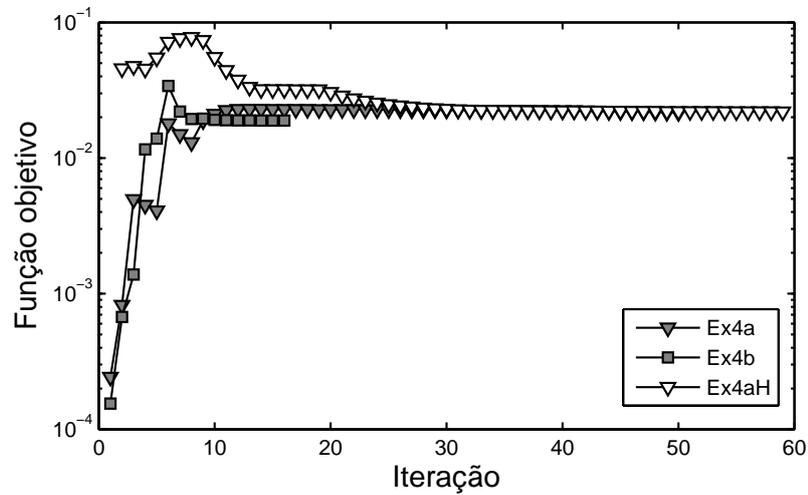


Figura 5.58: Exemplo 4, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

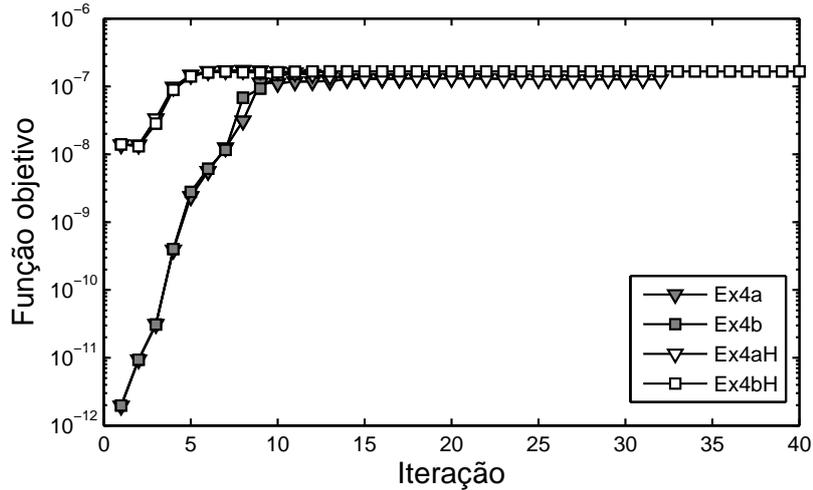


Figura 5.59: Exemplo 4, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

### Exemplo 5

A forma objetivo deste exemplo é o retângulo arredondado mostrado pela Figura 5.60. Oito indutores são considerados para o problema inverso. Os indutores nos lados superior e inferior são de tipo 3 segundo a Figura 5.35, enquanto que os localizados sobre os lados esquerdo e direito são de tipo 1. A forma objetivo tem área  $S_0$  igual a 4.99 e a tensão superficial  $\sigma$  é igual a  $1.0 \times 10^{-4}$ . A intensidade  $I$  é igual a 0.1 e os coeficientes adimensionados  $\alpha_p$  têm valor absoluto igual a 4.0 com o sinal dado pela Figura 5.60.

A Figura 5.61.a apresenta os indutores obtidos e a Figura 5.61.b mostra a forma em equilíbrio e algumas curvas de nível da função de fluxo magnético  $\varphi$  na solução do problema. As Figuras 5.62.a e 5.62.b mostram o mesmo para a segunda formulação. A evolução das funções objetivo durante o processo iterativo é apresentada pelas Figuras 5.63 e 5.64.

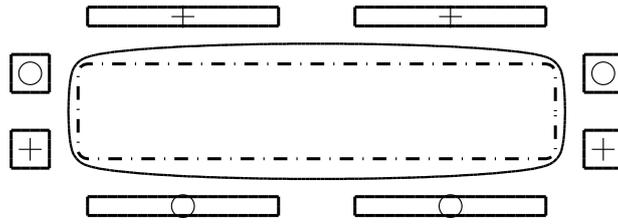


Figura 5.60: Exemplo 5, configuração inicial. Linha sólida: Curva  $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa.

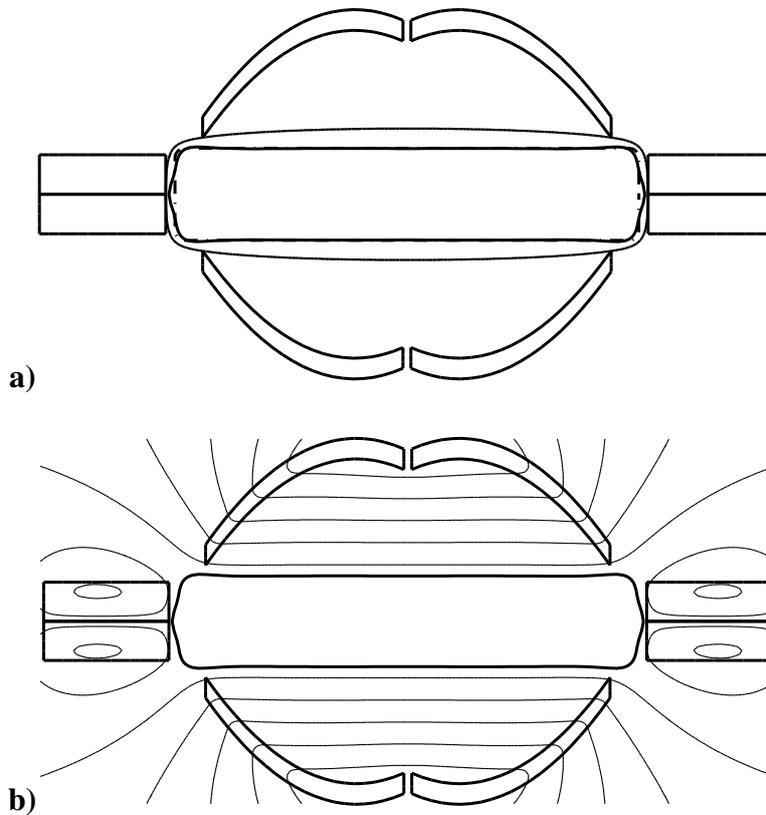


Figura 5.61: Solução do exemplo Ex5, primeira formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

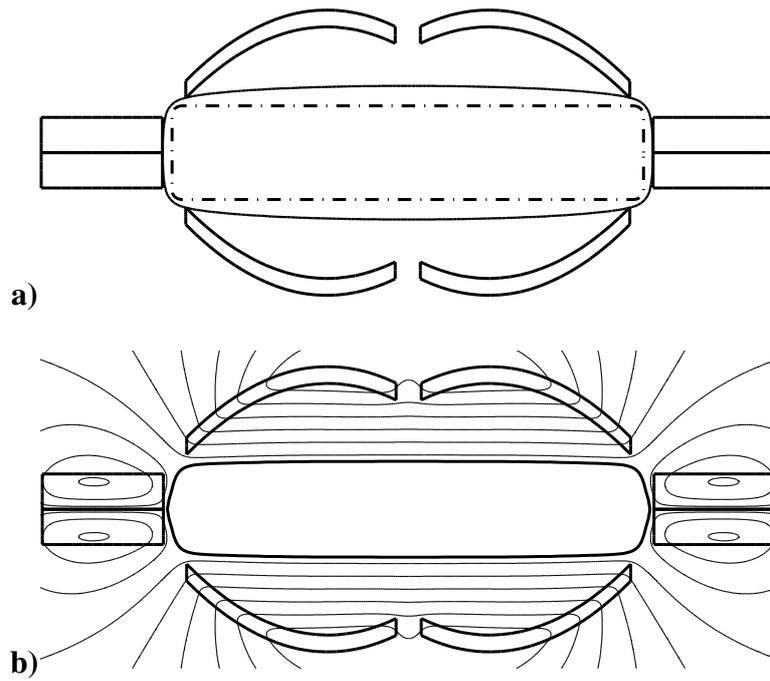


Figura 5.62: Solução do exemplo Ex5, segunda formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

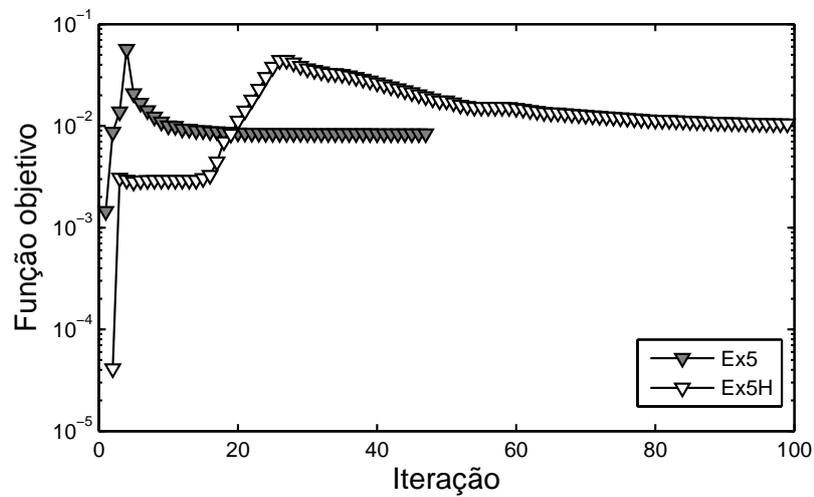


Figura 5.63: Exemplo 5, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

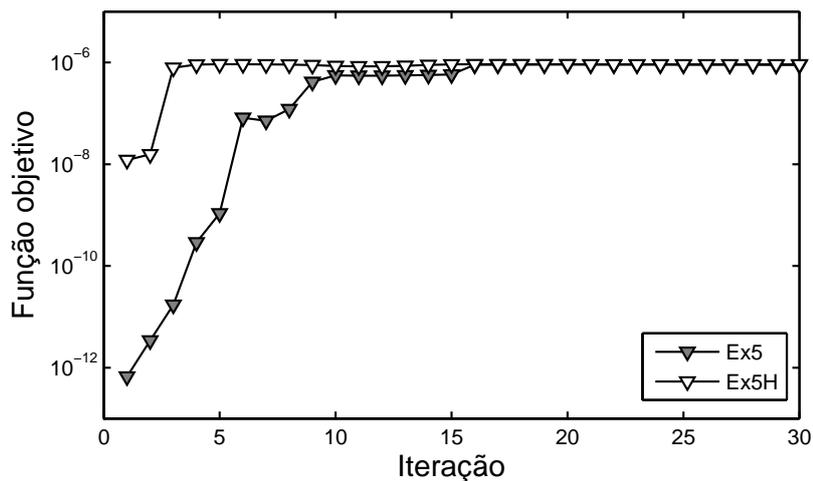


Figura 5.64: Exemplo 5, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

### Exemplo 6

Este exemplo é similar ao anterior com a diferença do sinal dos coeficientes  $\alpha_p$  que, neste caso, são mostrados pela Figura 5.65. Para o maior valor do parâmetro  $\psi_0$  e para a primeira formulação, a Figura 5.66.a mostra os indutores obtidos e a Figura 5.66.b apresenta a forma de equilíbrio e algumas curvas de nível da função de fluxo magnético  $\varphi$ . As Figuras 5.67.a e 5.67.b mostram o mesmo para a segunda formulação. Como as figuras mostram, as soluções obtidas para as duas formulações são muito diferentes neste caso. Isto mostra que os resultados da segunda formulação devem ser analisados com cuidado, visto que a sua solução pode ser bem diferente da solução da primeira formulação. A evolução das funções objetivo durante o processo iterativo é apresentada pelas Figuras 5.68 e 5.69.

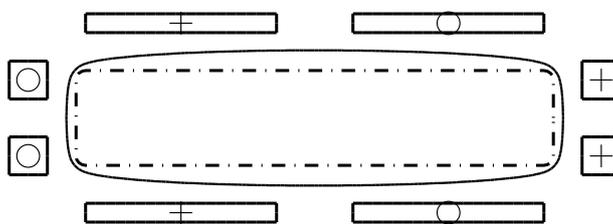


Figura 5.65: Exemplo 6, configuração inicial. Linha sólida: Curva  $\psi(x) = \psi_0$ . Linha tracejada: Forma objetivo. Sinal de adição: Indutor de corrente positiva. Círculo: Indutor de corrente negativa.

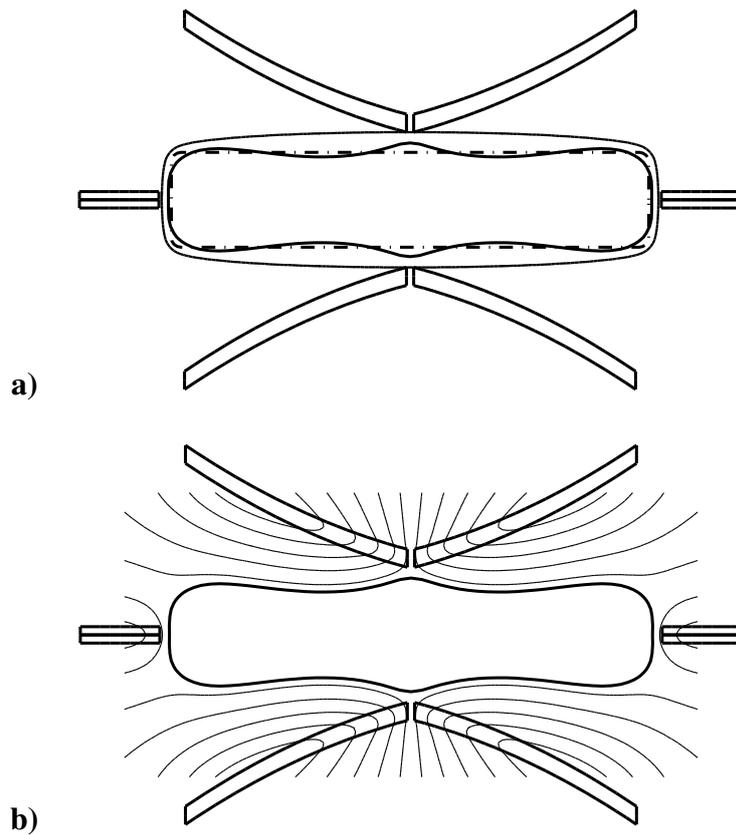


Figura 5.66: Solução do exemplo Ex6, primeira formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

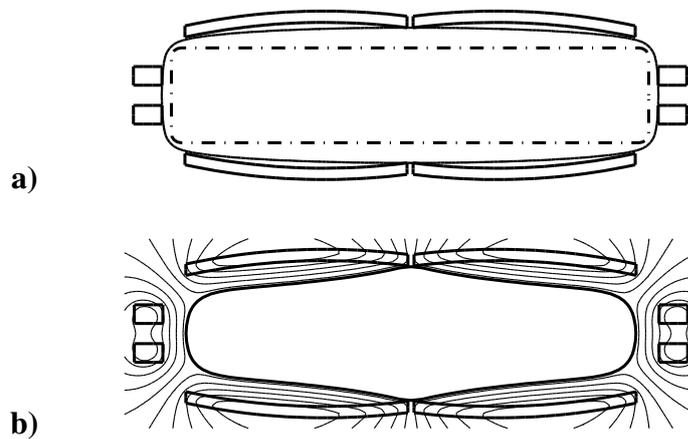


Figura 5.67: Solução do exemplo Ex6, segunda formulação, **a)** solução e restrição geométrica, **b)** forma em equilíbrio e curvas de nível da função de fluxo magnético  $\varphi$ .

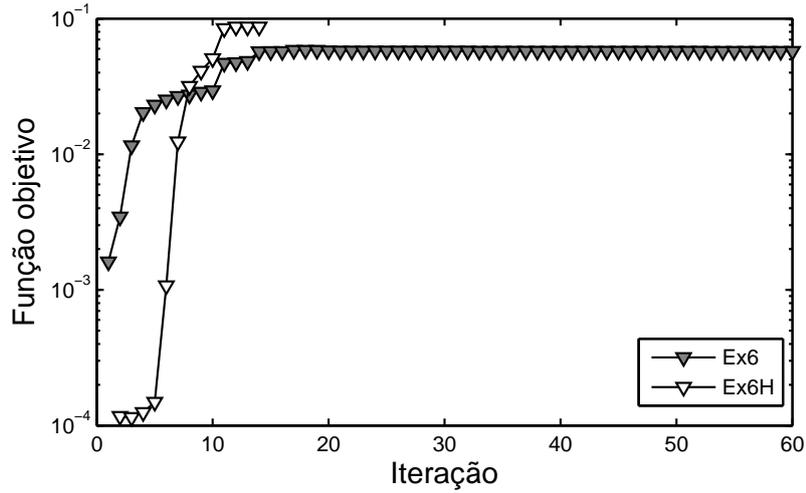


Figura 5.68: Exemplo 6, evolução da função objetivo da primeira formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

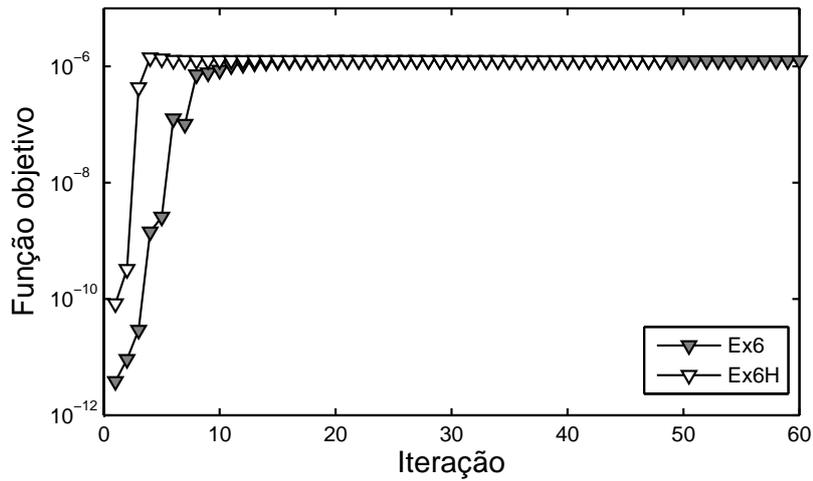


Figura 5.69: Exemplo 6, evolução da função objetivo da segunda formulação. Cinza: Algoritmo FAIPA. Branco: Algoritmo FDIPA-H.

### Resumo dos resultados

A tabela 5.2 resume a informação sobre os exemplos estudados. O número de nós usado para a aproximação por elementos lineares do contorno  $\Gamma^*$  da forma objetivo e o número de indutores são dados. Para cada formulação, o número de iterações feito pelos algoritmos FAIPA e FDIPA-H são indicados assim como o valor final da função objetivo.

Na tabela 5.2 pode-se ver que o algoritmo FAIPA foi o melhor no caso de utilizar a primeira formulação do problema inverso. O algoritmo FDIPA-H falha em dois dos

problemas considerados. No caso da Segunda formulação o FDIPA-H é claramente o melhor considerando o menor número de iterações feito para vários dos problemas onde o FAIPA realiza um número de iterações relativamente grande.

Tabela 5.2: Resumo dos resultados obtidos para os exemplos de otimização de forma dos indutores.

Exemplos			Primeira formulação				Segunda formulação			
Nós	Indut.		IFA	OFA	IFH	OFH	IFA	OFA	IFH	OFH
Ex1a	72	4	59	1.293e-07	13	1.263e-06	18	5.578e-11	16	2.062e-12
Ex1b	72	4	74	4.156e-07	-	ERRO	18	1.488e-10	31	2.062e-12
Ex2a	80	4	30	1.733e-04	40	2.703e-04	94	2.339e-08	30	2.338e-08
Ex2b	80	4	30	1.728e-04	21	1.619e-04	332	3.535e-08	24	3.535e-08
Ex2c	80	4	31	1.721e-04	62	1.701e-04	322	4.268e-08	21	4.268e-08
Ex2d	80	4	33	1.688e-04	31	2.542e-04	103	4.731e-08	18	4.731e-08
Ex3a	120	8	26	7.169e-04	31	1.567e-03	21	1.008e-07	27	1.009e-07
Ex3b	120	8	400	4.539e-03	29	4.715e-03	20	1.156e-07	25	1.143e-07
Ex4a	120	8	50	2.170e-02	59	2.197e-02	17	1.282e-07	32	1.273e-07
Ex4b	120	8	16	1.888e-02	-	ERRO	14	1.668e-07	77	1.663e-07
Ex5	136	8	47	8.390e-03	99	1.051e-02	30	8.990e-07	30	9.185e-07
Ex6	136	8	400	5.690e-02	14	8.691e-02	329	1.261e-06	48	1.264e-06

IFA: Iterações realizadas pelo FAIPA, OFA: Função objetivo alcançada pelo FAIPA.

IFH: Iterações realizadas pelo FDIPA-H, OFH: Função objetivo alcançada pelo FDIPA-H.

# Capítulo 6

## Conclusões

Neste trabalho foi desenvolvido um método para otimização em Engenharia Mecânica. Este método foi baseado na utilização do MEC para a análise do problema tratado, na formulação SAND do problema de otimização e na utilização de um Algoritmo de Ponto Interior baseado na iteração de Newton para resolver o problema discreto. Como foi mostrado, por causa das características do MEC, da formulação SAND e do Algoritmo de Ponto Interior, o método proposto é especialmente indicado para problemas de otimização de grande porte em Engenharia Mecânica, nos casos onde a equação de estado seja não linear ou esteja definida sobre um domínio infinito. Para ilustrar o desempenho do método de otimização proposto foram estudados os problemas de otimização de forma em elasticidade linear e o problema inverso em EMC.

No Capítulo 3 foi apresentado um algoritmo de ponto interior e direções viáveis, aqui chamado FDIPA-H, que baseia-se no algoritmo de otimização FDIPA. O algoritmo FDIPA-H permite utilizar a matriz hessiana do problema de otimização no lugar da matriz quasi-Newton utilizada pelo algoritmo FDIPA. Por esta razão, o algoritmo FDIPA-H é conveniente para os problemas de grande porte onde seja possível calcular eficientemente a matriz hessiana. Como mostrado no Capítulo 4, nesta importante categoria podem ser citados os problemas de otimização em Engenharia Mecânica, nos quais a formulação SAND seja utilizada.

O resultado teórico mais importante obtido para o algoritmo FDIPA-H é o Teorema 3.50 do Capítulo 3, que prova que toda subsequência convergente da sequência gerada pelo algoritmo converge para um ponto de Karush-Kuhn-Tucker do problema de otimização.

A utilização do algoritmo FDIPA-H na solução dos problemas teste da coleção de Hock e Schittkowski mostrou que a implementação realizada do algoritmo proposto é, para este extenso conjunto de problemas de pequeno tamanho, tão robusta quanto as implementações quasi-Newton largamente utilizadas dos algoritmos FDIPA e FAIPA. A solução dos problemas de grande porte da coleção CUTE mostrou que o algoritmo FDIPA-H é promissor, visto que a implementação realizada apresentou-se muito competitiva quando comparada com dois dos mais conhecidos programas para otimização não linear, o SQP KNITRO e o Primal-dual LOQO.

Existem ainda alguns assuntos de pesquisa que deverão ser considerados no futuro para melhorar a robustez e a eficiência do algoritmo. Com relação à robustez, será importante estudar as técnicas de busca em arco, como a utilizada no algoritmo FAIPA, as técnicas de ajuste de escala das restrições, assim como as técnicas de redução não monótona da função potencial. Será também relevante estudar uma formulação não viável do algoritmo, como a implementada no algoritmo LOQO [48], para melhorar o desempenho nos problemas que não possuam um ponto inicial viável. Por outro lado, será importante realizar estudos teóricos visando esclarecer o comportamento assintótico do algoritmo.

No Capítulo 4 mostrou-se, através do estudo do problema de otimização de forma em elasticidade linear, que as matrizes de derivadas primeiras e segundas das funções da formulação SAND, que surgem da aplicação do MEC, podem ser calculadas eficientemente, ou seja, podem ser calculadas com um custo computacional inferior ao custo de resolução dos sistemas lineares do algoritmo de otimização. Este resultado é fundamental para a aplicabilidade do método proposto. É importante salientar que nesse estudo foram consideradas as funções que definem três problemas de otimização diferentes, entre elas as que definem as tensões e deslocamentos no contorno do domínio do problema, assim como alguns funcionais definidos como integrais no contorno, como no caso da complacência. Assim, o método proposto para o cálculo das derivadas apresenta-se suficientemente geral como para ser utilizado em diversos problemas de otimização estrutural.

Por último, uma aplicação do método proposto ao problema EMC inverso foi apresentada no Capítulo 5. Neste problema um conjunto de indutores elétricos gera um campo magnético que exerce pressão sobre a superfície de uma coluna de metal líquido. O problema consiste em encontrar a configuração de indutores de forma tal que a seção transversal da coluna adquira uma forma predeterminada. No problema EMC inverso

destaca-se a natureza não linear das equações de equilíbrio, assim como a necessidade de considerar domínios infinitos. Estas características tornam muito difícil a solução dos problemas de otimização relativos ao problema inverso em EMC. Porém, utilizando o método proposto foi possível resolver os problemas considerados.

Outras contribuições mais específicas ao problema EMC foram apresentadas no Capítulo 5. Em primeiro lugar, duas formulações foram propostas para a solução do problema inverso, isto é, o problema de projetar os indutores elétricos. Na primeira formulação é minimizada a o deslocamento associado à transformação que vincula a forma objetivo com a mais próxima forma em equilíbrio. A segunda formulação minimiza a pressão fictícia necessária para equilibrar a pressão magnética e a tensão superficial na forma objetivo.

Os exemplos apresentados permitem observar que as duas formulações são efetivas para projetar os indutores, porém, elas não são equivalentes. A primeira formulação permitiu encontrar configurações de indutores que fornecem formas em equilíbrio mais próximas das formas objetivo para quase todos os exemplos. A vantagem principal da segunda formulação é a de encontrar soluções similares às encontradas com a primeira formulação mas com um custo computacional menor, principalmente por causa da ausência de variáveis de forma. Portanto, a segunda formulação pode ser utilizada para encontrar uma solução inicial que pode, se necessário, ser melhorada em um processo posterior utilizando esta solução como ponto inicial para a primeira formulação.

Em segundo lugar, foi apresentada uma técnica geral para satisfazer as restrições geométricas de não superposição que devem ser consideradas no problema.

Em trabalhos futuros seria interessante estudar a utilização de técnicas de otimização topológica para determinar configurações iniciais convenientes para o problema de otimização de forma dos indutores. Estas técnicas foram utilizadas com sucesso no problema de otimização de forma em elasticidade linear. No caso do problema de otimização de forma em EMC alguns exemplos considerados sugerem que a otimização topológica pode ser igualmente importante que no problema de elasticidade linear, visto que, na otimização de forma, a solução obtida pelo algoritmo de otimização depende fortemente da configuração inicial de indutores.

# Referências Bibliográficas

- [1] KARMAKAR, N., “A new polynomial-time algorithm for linear programming”, *Combinatorica*, v. 4, n. 4, pp. 373–395, 1984.
- [2] MACULAN, N., COSTA FAMPA, M. H., *Otimização linear*. Editora Universidade de Brasília: Brasília, 2006.
- [3] GOULD, N. I. M., “Some reflections on the current state of active-set and interior-point methods for constrained optimization”, *SIAG/OPT Views-and-News*, v. 14, n. 1, pp. 2–7, 2003.
- [4] BENSON, H. Y., SHANNO, D. F., VANDERBEI, R. J., “A comparative study of large-scale nonlinear optimization algorithms”, In: *High performance algorithms and software for nonlinear optimization (Erice, 2001)*, v. 82, pp. 95–127, *Appl. Optim.*, Kluwer Acad. Publ.: Norwell, MA, 2003.
- [5] DOLAN, E. D., MOREÉ, J. J., *Benchmarking optimization software with COPS 3.3*, Tech. Rep. ANL/MCS-TM-273, Argonne National Laboratory, Argonne, Illinois, USA, 2004.
- [6] MORALES, J. L., NOCEDAL, J., WALTZ, R. A., et al., “Assessing the potential of interior methods for nonlinear optimization”, In: *Large-scale PDE-constrained optimization (Santa Fe, NM, 2001)*, v. 30, pp. 167–183, *Lect. Notes Comput. Sci. Eng.*, Springer: Berlin, 2003.
- [7] TAPIA, R. A., *On the fundamental role of interior-point methodology in constrained optimization*, Tech. Rep. TR97-09, Department of Computational and Applied Mathematics, Rice University, 1997.

- [8] NOCEDAL, J., WRIGHT, S. J., *Numerical optimization. Springer Series in Operations Research*, Springer-Verlag: New York, 1999.
- [9] BAZARAA, M. S., SHERALI, H. D., SHETTY, C. M., *Nonlinear programming*. 3rd ed. Wiley-Interscience [John Wiley & Sons]: Hoboken, NJ, 2006.
- [10] ZANGWILL, W. I., MOND, B., *Nonlinear programming: a unified approach*. Prentice-Hall Inc.: Englewood Cliffs, N.J., 1969.
- [11] LUENBERGER, D. G., *Linear and nonlinear programming*. 2nd ed. Kluwer Academic Publishers: Boston, MA, 2003.
- [12] KELLEY, C. T., *Solving nonlinear equations with Newton's method. Fundamentals of Algorithms*, Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, 2003.
- [13] DENNIS, JR., J. E., SCHNABEL, R. B., *Numerical methods for unconstrained optimization and nonlinear equations*. v. 16. *Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, 1996.
- [14] HERSKOVITS, J., "Feasible direction interior-point technique for nonlinear optimization", *J. Optim. Theory Appl.*, v. 99, n. 1, pp. 121–146, 1998.
- [15] BREBBIA, C. A., TELLES, J. C. F., WROBEL, L. C., *Boundary Element Technique: Theory and Applications in Engineering*. Springer-Verlag, 1984.
- [16] BREBBIA, C. A., DOMINGUEZ, J., *Boundary elements: an introductory course*. 2nd ed. Computational Mechanics Publications: Southampton, 1992.
- [17] BEER, G., WATSON, J. O., *Introduction to Finite and Boundary Element Methods for Engineers*. Wiley, 1992.
- [18] PARÍS, F., CAÑAS, J., *Boundary element method. Oxford Science Publications*, The Clarendon Press Oxford University Press: New York, 1997.
- [19] GURTIN, M. E., *An introduction to continuum mechanics*. v. 158. *Mathematics in Science and Engineering*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers]: New York, 1981.

- [20] ŠILHAVÝ, M., *The mechanics and thermodynamics of continuous media. Texts and Monographs in Physics*, Springer-Verlag: Berlin, 1997.
- [21] ZHANG, X., ZHANG, X., “Exact integration for stress evaluation in the boundary element analysis of two-dimensional elastostatics”, *Eng. Anal. Boundary Elem.*, v. 28, n. 9, pp. 997–1004, 2004.
- [22] FEDOTOV, V. P., SPEVAK, L., “One approach to the derivation of exact integration formulae in the boundary element method”, *Eng. Anal. Boundary Elem.*, v. 32, n. 10, pp. 883–888, 2008.
- [23] HORN, R. A., JOHNSON, C. R., *Matrix analysis*. Cambridge University Press: Cambridge, 1990.
- [24] GOLUB, G. H., VAN LOAN, C. F., *Matrix computations*. 3rd ed. *Johns Hopkins Studies in the Mathematical Sciences*, Johns Hopkins University Press: Baltimore, MD, 1996.
- [25] HAFTKA, R. T., GRANDHI, R. V., “Structural shape optimization—a survey”, *Comput. Methods Appl. Mech. Engrg.*, v. 57, n. 1, pp. 91–106, 1986.
- [26] SAMAREH, J. A., *A survey of shape parameterization techniques*, Tech. Rep. NASA-CP-1999-209136, NASA Langley Research Center, Hampton, VA 23681, 1999.
- [27] CHOI, K. K., KIM, N. H., *Structural Sensitivity Analysis and Optimization, 1 and 2*. Springer: Berlin, Heidelberg, New York, 2004.
- [28] BENDSØE, M. P., *Optimization of structural topology, shape, and material*. Springer-Verlag: Berlin, 1995.
- [29] HAFTKA, R. T., “Simultaneous analysis and design”, *AIAA Journal*, v. 23, n. 7, pp. 1099–1103, 1985.
- [30] HAFTKA, R. T., KAMAT, M. P., “Simultaneous nonlinear structural analysis and design”, *Comput. Mech.*, v. 4, n. 6, pp. 409–416, 1989.
- [31] ARORA, J. S., WANG, Q., “Review of formulations for structural and mechanical system optimization”, *Struct. Multidiscip. Optim.*, v. 30, n. 4, pp. 251–272, 2005.

- [32] ARORA, J. S., WANG, Q., *Optimization of Large-Scale Structural Systems Using Sparse SAND Formulations*, Tech. rep., Optimal Design Lab/CCAD, College of Engineering/4110 SC, The University of Iowa, Iowa City, IA 52242, 2004.
- [33] GILL, P. E., MURRAY, W., SAUNDERS, M. A., “SNOPT: an SQP algorithm for large-scale constrained optimization”, *SIAM J. Optim.*, v. 12, n. 4, pp. 979–1006 (electronic), 2002.
- [34] CANELAS, A., *Técnicas de ponto interior para sistemas de equações e otimização não linear*, Master’s Thesis, Universidade federal de Rio de Janeiro, COPPE, Rio de Janeiro, Brasil, 2005.
- [35] BEN-TAL, A., NEMIROVSKI, A., “Robust truss topology design via semidefinite programming”, *SIAM J. Optim.*, v. 7, n. 4, pp. 991–1016, 1997.
- [36] CANELAS, A., HERSKOVITS, J., TELLES, J. C. F., “Shape Optimization using the Boundary Element Method and a SAND Interior Point Algorithm for Constrained Optimization”, *Comput. & Structures*, v. 86, n. 13–14, pp. 1517–1526, 2007.
- [37] CANELAS, A., HERSKOVITS, J., TELLES, J. C. F., “Shape optimization using the BEM and SAND formulations”. In: *Proceedings of COBEM 2007, 19th International Congress of Mechanical Engineering*, 2007.
- [38] ALLAIRE, G., *Conception Optimale de Structures*. v. 58. *Mathématiques and Applications*, Springer-Verlag: Berlin, 2007.
- [39] MURAT, S., SIMON, J., *Sur le contrôle par un domaine géométrique*, Tech. Rep. 76015, Laboratoire d’Analyse Numérique, Université de Paris, 1976.
- [40] SIMON, J., “Differentiation with respect to the domain in boundary value problems”, *Numer. Funct. Anal. Optim.*, v. 2, n. 7-8, pp. 649–687 (1981), 1980.
- [41] SOKOŁOWSKI, J., ZOLÉSIO, J.-P., *Introduction to shape optimization*. v. 16. *Springer Series in Computational Mathematics*, Springer-Verlag: Berlin, 1992.
- [42] NOVOTNY, A. A., FEIJÓO, R. A., TAROCO, E., “Análise de Sensibilidade Contínua”, Notas de Aula do curso “Análise de Sensibilidade” dictado no Laboratório nacional de computação Científica, 2007.

- [43] HIGHAM, N. J., *Accuracy and stability of numerical algorithms*. 2nd ed. Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, 2002.
- [44] BUNCH, J. R., PARLETT, B. N., “Direct methods for solving symmetric indefinite systems of linear equations”, *SIAM J. Numer. Anal.*, v. 8, pp. 639–655, 1971.
- [45] BUNCH, J. R., KAUFMAN, L., “Some stable methods for calculating inertia and solving symmetric linear systems”, *Math. Comp.*, v. 31, n. 137, pp. 163–179, 1977.
- [46] ASHCRAFT, C., GRIMES, R. G., LEWIS, J. G., “Accurate symmetric indefinite linear equation solvers”, *SIAM J. Matrix Anal. Appl.*, v. 20, n. 2, pp. 513–561 (electronic), 1999.
- [47] GOLDFELD, S. M., QUANDT, R. E., TROTTER, H. F., “Maximization by quadratic hill-climbing”, *Econometrica*, v. 34, pp. 541–551, 1966.
- [48] VANDERBEI, R. J., SHANNO, D. F., “An interior-point algorithm for nonconvex nonlinear programming”, *Comput. Optim. Appl.*, v. 13, n. 1-3, pp. 231–252, 1999.
- [49] HOCK, W., SCHITTKOWSKI, K., *Test examples for nonlinear programming codes*. v. 187. *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag: Berlin, 1981.
- [50] BONGARTZ, I., CALAMAI, P. H., CONN, A. R., et al., “CUTE: Constrained and Unconstrained Testing Environment”, *ACM Trans. Math. Software*, v. 21, n. 1, pp. 123–160, 1995.
- [51] DUFF, I. S., “MA57—a code for the solution of sparse symmetric definite and indefinite systems”, *ACM Trans. Math. Software*, v. 30, n. 2, pp. 118–144, 2004.
- [52] DUFF, I. S., “Sparse system solution and the HSL library”, In: *Some topics in industrial and applied mathematics*, v. 8, pp. 78–94, *Ser. Contemp. Appl. Math. CAM*, Higher Ed. Press: Beijing, 2007.
- [53] MARATOS, N., *Exact penalty function algorithms for finite dimensional optimization problems with equality and inequality constraints*, Ph.D. Thesis, Imperial College of Science and Technology, London, 1978.

- [54] BYRD, R. H., NOCEDAL, J., WALTZ, R. A., “KNITRO: An integrated package for nonlinear optimization”, In: *Large-scale nonlinear optimization*, v. 83, pp. 35–59, *Nonconvex Optim. Appl.*, Springer: New York, 2006.
- [55] PARVIZIAN, J., FENNER, R. T., “Shape optimization by the boundary element method: a comparison between mathematical programming and normal movement approaches”, *Eng. Anal. Boundary Elem.*, v. 19, n. 2, pp. 137–145, 1997.
- [56] TAFRESHI, A., “Shape design sensitivity analysis of 2D anisotropic structures using the boundary element method”, *Eng. Anal. Boundary Elem.*, v. 26, n. 3, pp. 237–251, 2002.
- [57] BURCZYŃSKI, T., “Applications of BEM in sensitivity analysis and optimization”, *Comput. Mech.*, v. 13, n. 1-2, pp. 29–44, 1993.
- [58] WESSEL, C., CISILINO, A., SENSALÉ, B., “Structural shape optimisation using boundary elements and the biological growth method”, *Struct. Multidiscip. Optim.*, v. 28, n. 2-3, pp. 221–227, 2004.
- [59] CERROLAZA, M., ANNICCHIARICO, W., M., M., “Optimization of 2D boundary element models using  $\beta$ -splines and genetic algorithms”, *Eng. Anal. Boundary Elem.*, v. 24, n. 5, pp. 427–440, 2000.
- [60] BIALECKI, R. A., BURCZYŃSKI, T., DLUGOSZ, A., et al., “Evolutionary shape optimization of thermoelastic bodies exchanging heat by convection and radiation”, *Comput. Methods Appl. Mech. Engrg.*, v. 194, n. 17, pp. 1839–1859, 2005.
- [61] KITA, E., TANIE, H., “Shape optimization of continuum structures by genetic algorithm and boundary element method”, *Eng. Anal. Boundary Elem.*, v. 19, n. 2, pp. 129–136, 1997.
- [62] BARRA, L. P., *Sensitivity analysis—an application of the boundary element method to shape optimization (in Portuguese)*, Master’s Thesis, Federal University of Rio de Janeiro, COPPE, Rio de Janeiro, Brazil, 1990.
- [63] MARTIN, T. J., DULIKRAVICH, G. S., “An implicit and explicit BEM sensitivity approach for thermo-structural optimization”, *Eng. Anal. Boundary Elem.*, v. 28, n. 3, pp. 257–265, 2004.

- [64] SFANTOS, G. K., ALIABADI, M. H., “A boundary element sensitivity formulation for contact problems using the implicit differentiation method”, *Eng. Anal. Boundary Elem.*, v. 30, n. 1, pp. 22–30, 2006.
- [65] MERIÇ, R. A., “Differential and integral sensitivity formulations and shape optimization by BEM”, *Eng. Anal. Boundary Elem.*, v. 15, n. 2, pp. 181–188, 1995.
- [66] BRAIBANT, V., FLEURY, C., “Shape optimal design using B-Splines”, *Comput. Methods Appl. Mech. Engrg.*, v. 44, pp. 247–267, 1984.
- [67] SIGMUND, O., “A 99 line topology optimization code written in Matlab”, *Struct. Multidiscip. Optim.*, v. 21, n. 2, pp. 120–127, 2001.
- [68] “TopGUIT”, Disponível em: <http://www.geocities.com/chinakkong/research/topgui/>.
- [69] HERSKOVITS, J., DIAS, G. P., MOTA SOARES, C. M., “A full-stress technique for structural shape optimization”, *Appl. Math. Comput. Sci.*, v. 6, n. 2, pp. 303–319, 1996.
- [70] ZHIQIANG, C., FEI, J., XINGGUO, Z., et al., “Microstructures and mechanical characteristics of electromagnetic casting and direct-chill casting 2024 aluminum alloys”, *Mater. Sci. Eng., A*, v. 327, n. 2, pp. 133–137, 2002.
- [71] FU, H. Z., SHEN, J., L., L., et al., “Electromagnetic shaping and solidification control of Ni-base superalloys under vacuum”, *J. Mater. Process. Technol.*, v. 148, n. 1, pp. 25–29, 2004.
- [72] CANELAS, A., ROCHE, J. R., HERSKOVITS, J., “The inverse electromagnetic shaping problem.” *Struct. Multidiscip. Optim.*, v. 38, n. 4, pp. 389–403, 2009.
- [73] CANELAS, A., ROCHE, J. R., HERSKOVITS, J., “Inductor shape optimization for electromagnetic casting.” *Struct. Multidiscip. Optim.*, doi: 10.1007/s00158-009-0386-0, 2009.
- [74] SHERCLIFF, J. A., “Magnetic Shaping of Molten Metal Columns”, *Proc. R. Soc. Lond. A*, v. 375, pp. 455–473, April 1981.
- [75] KRESS, R., *Linear integral equations*. 2nd ed., v. 82. *Applied Mathematical Sciences*, Springer-Verlag: New York, 1999.

- [76] HENROT, A., PIERRE, M., “Un problème inverse en formage de métaux liquides”, *Modélisation mathématique et analyse numérique*, v. 23, n. 1, pp. 155–177, 1989.
- [77] PIERRE, M., ROCHE, J.-R., “Computation of free surfaces in the electromagnetic shaping of liquid metals by optimization algorithms”, *European J. Mech. B Fluids*, v. 10, n. 5, pp. 489–500, 1991.
- [78] NOVRUZI, A., ROCHE, J. R., *Second Order Derivatives, Newton Method, Application to Shape Optimization*, Tech. Rep. RR-2555, INRIA, 1995.
- [79] BRANCHER, J.-P., SÉRO-GUILLAUME, O. E., “Étude de la déformation d’un liquide magnétique”, *Arch. Rational Mech. Anal.*, v. 90, n. 1, pp. 57–85, 1985.
- [80] COULAUD, O., HENROT, A., “Numerical approximation of a free boundary problem arising in electromagnetic shaping”, *SIAM J. Numer. Anal.*, v. 31, n. 4, pp. 1109–1127, 1994.
- [81] GAGNOUD, A., ETAY, J., GARNIER, M., “Le problème de frontière libre en lévitation électromagnétique”, *J. Méc. Théor. Appl.*, v. 5, n. 6, pp. 911–934, 1986.
- [82] HENROT, A., BRANCHER, J.-P., PIERRE, M., “Existence of equilibria in electromagnetic casting”. In: *Proceedings of the Fifth International Symposium on Numerical Methods in Engineering, Vol. 1, 2 (Lausanne, 1989)*, pp. 221–228, Comput. Mech.: Southampton, 1989.
- [83] MOFFATT, H. K., “Magnetostatic equilibria and analogous Euler flows of arbitrarily complex topology. I. Fundamentals”, *J. Fluid Mech.*, v. 159, pp. 359–378, 1985.
- [84] NOVRUZI, A., *Contribution en Optimization de Formes et Applications*, Ph.D. Thesis, Université Henri Poincaré, Nancy 1, 1997.
- [85] NOVRUZI, A., ROCHE, J. R., “Newton’s method in shape optimisation: a three-dimensional case”, *BIT*, v. 40, n. 1, pp. 102–120, 2000.
- [86] PIERRE, M., ROCHE, J.-R., “Computation of free surfaces in the electromagnetic shaping of liquid metals by optimization algorithms”, *European J. Mech. B Fluids*, v. 10, n. 5, pp. 489–500, 1991.

- [87] PIERRE, M., ROCHE, J.-R., “Numerical simulation of tridimensional electromagnetic shaping of liquid metals”, *Numer. Math.*, v. 65, n. 2, pp. 203–217, 1993.
- [88] ROCHE, J. R., “Algorithmes numériques en optimization de formes et électromagnétisme”, Mémoire d’Habilitation à Diriger des Recherches, 1996.
- [89] ROCHE, J. R., “Gradient of the discretized energy method and discretized continuous gradient in electromagnetic shaping simulation”, *Appl. Math. Comput. Sci.*, v. 7, n. 3, pp. 545–565, 1997.
- [90] SÉRO-GUILLAUME, O. E., ZOUAOU, D., BERNARDIN, D., et al., “The shape of a magnetic liquid drop”, *J. Fluid Mech.*, v. 241, pp. 215–232, 1992.
- [91] SNEYD, A. D., MOFFATT, H. K., “Fluid dynamical aspects of the levitation-melting process”, *J. Fluid Mech.*, v. 117, pp. 45–70, 1982.
- [92] ZOUAOU, D., SÉRO-GUILLAUME, O. E., BRANCHER, J.-P., “Equilibrium of a magnetic liquid drop: variational approach and computation”, *Magnit. Gidrodinamika*, v. 26, n. 4, pp. 32–35, 150, 1990.
- [93] FELICI, T. P., BRANCHER, J.-P., “The inverse shaping problem”, *European J. Mech. B Fluids*, v. 10, n. 5, pp. 501–512, 1991.
- [94] PIERRE, M., ROUY, E., “A tridimensional inverse shaping problem”, *Comm. Partial Differential Equations*, v. 21, n. 7-8, pp. 1279–1305, 1996.
- [95] HERSKOVITS, J., MAPPA, P., GOULART, E., et al., “Mathematical programming models and algorithms for engineering design optimization”, *Comput. Methods Appl. Mech. Engrg.*, v. 194, n. 30-33, pp. 3244–3268, 2005.
- [96] SULLIVAN, C. R., “Optimal Choice for Number of Strands in a Litz-Wire Transformer Winding”, *IEEE Trans. Power Electron.*, v. 14, n. 2, pp. 283–291, 1999.

# Apêndice A

## Resultados obtidos com o FDIPA-H

Tabela A.1: Resultados obtidos para os problemas de Hock e Schittkowski.

Problema			Iterações			Objetivo		
prob	nvar	nres	FDIPA	FAIPA	FDIPA-H	FDIPA	FAIPA	FDIPA-H
1	2	0	41	41	27	1.7608e-23	1.7608e-23	0.0000e+00
2	2	0	14*	14*	11*	4.9412e+00	4.9412e+00	4.9412e+00
3	2	0	16	16	6	4.6564e-12	4.6564e-12	3.0437e-09
4	2	0	6	6	6	2.6667e+00	2.6667e+00	2.6667e+00
5	2	0	8	8	6	-1.9132e+00	-1.9132e+00	-1.9132e+00
6	2	1	11	12	5	1.1511e-25	2.2795e-21	0.0000e+00
7	2	1	12	12	10	-1.7321e+00	-1.7321e+00	-1.7321e+00
8	2	2	7	9	5	-1.0000e+00	-1.0000e+00	-1.0000e+00
9	2	1	7	7	6	-5.0000e-01	-5.0000e-01	-5.0000e-01
10	2	1	29	20	9	-1.0000e+00	-1.0000e+00	-1.0000e+00
11	2	1	11	12	9	-8.4985e+00	-8.4985e+00	-8.4985e+00
12	2	1	20	9	11	-3.0000e+01	-3.0000e+01	-3.0000e+01
13	2	1	34	27	65	1.0030e+00	1.0035e+00	1.0000e+00
14	2	2	11	16	9	1.3935e+00	1.3935e+00	1.3935e+00
15	2	2	12	48	14	3.6038e+02	3.6038e+02	3.6038e+02
16	2	2	7*	7*	6*	2.3145e+01	2.3145e+01	2.3145e+01
17	2	2	31	13	28	1.0000e+00	1.0000e+00	1.0000e+00
18	2	2	14	16	10	5.0000e+00	5.0000e+00	5.0000e+00
19	2	2	33	100	29	-6.9618e+03	-6.9618e+03	-6.9618e+03
20	2	3	6*	6*	6*	4.0199e+01	4.0199e+01	4.0199e+01
21	2	1	12	12	9	-9.9960e+01	-9.9960e+01	-9.9960e+01
22	2	2	12	14	11	1.0000e+00	1.0000e+00	1.0000e+00
23	2	5	13	15	12	2.0000e+00	2.0000e+00	2.0000e+00
24	2	3	15	13	12	-1.0000e+00	-1.0000e+00	-1.0000e+00
25	3	0	2*	2*	2*	3.2835e+01	3.2835e+01	3.2835e+01
26	3	1	100	43	25	1.5100e-05	4.9718e-11	1.4843e-17
27	3	1	100	21	9	4.0000e-02	4.0000e-02	4.0000e-02
28	3	1	4	4	3	0.0000e+00	0.0000e+00	0.0000e+00
29	3	1	22	15	8	-2.2627e+01	-2.2627e+01	-2.2627e+01

Tabela A.1: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	FDIPA	FAIPA	FDIPA-H	FDIPA	FAIPA	FDIPA-H
30	3	1	12	11	8	1.0000e+00	1.0000e+00	1.0000e+00
31	3	1	10	14	6	6.0000e+00	6.0000e+00	6.0000e+00
32	3	2	17	17	16	1.0000e+00	1.0000e+00	1.0000e+00
33	3	2	15	16	7*	-4.5858e+00	-4.5858e+00	-4.0000e+00
34	3	2	19	17	10	-8.3403e-01	-8.3403e-01	-8.3403e-01
35	3	1	8	8	7	1.1111e-01	1.1111e-01	1.1111e-01
36	3	1	14	13	13	-3.3000e+03	-3.3000e+03	-3.3000e+03
37	3	2	18	17	10	-3.4560e+03	-3.4560e+03	-3.4560e+03
38	4	0	21	21	9	3.1462e-18	3.1462e-18	8.5109e-29
39	4	2	17	17	10	-1.0000e+00	-1.0000e+00	-1.0000e+00
40	4	3	8	7	5	-2.5000e-01	-2.5000e-01	-2.5000e-01
41	4	1	16	16	12	1.9259e+00	1.9259e+00	1.9259e+00
42	4	2	10	10	64	1.3858e+01	1.3858e+01	1.3858e+01
43	4	3	18	13	8	-4.4000e+01	-4.4000e+01	-4.4000e+01
44	4	6	22	22	16	-1.5000e+01	-1.5000e+01	-1.5000e+01
45	5	0	24	24	12	1.0000e+00	1.0000e+00	1.0000e+00
46	5	2	100	40	25	2.9479e-05	1.8371e-10	1.1731e-16
47	5	3	19	14	11	1.5579e-17	4.2717e-25	0.0000e+00
48	5	2	4	4	3	0.0000e+00	3.6978e-31	0.0000e+00
49	5	2	29	29	12	7.0642e-13	7.0642e-13	7.4825e-12
50	5	3	17	17	11	1.3902e-19	1.3902e-19	1.7256e-31
51	5	3	3	3	3	2.4652e-32	1.2326e-32	0.0000e+00
52	5	3	7	7	3	5.3266e+00	5.3266e+00	5.3266e+00
53	5	3	9	9	5	4.0930e+00	4.0930e+00	4.0930e+00
54	6	1	58*	58*	17	-8.6741e-01	-8.6741e-01	-9.0807e-01
55	6	6	-	7	9	ERRO	6.6667e+00	6.6667e+00
56	7	4	22	16	6	-3.4560e+00	-3.4560e+00	-3.4560e+00
57	2	1	37	41	16	2.8499e-02	2.8501e-02	2.8460e-02
59	2	3	33*	19*	11*	-6.7495e+00	-6.7495e+00	-6.7495e+00
60	3	1	9	10	8	3.2568e-02	3.2568e-02	3.2568e-02
61	3	2	-	-	7	ERRO	ERRO	-1.4365e+02
62	3	1	7	7	5	-2.6273e+04	-2.6273e+04	-2.6273e+04
63	3	2	77	11	7	9.6172e+02	9.6172e+02	9.6172e+02
64	3	1	54	31	29	6.2998e+03	6.2998e+03	6.2998e+03
65	3	1	14	16	12	9.5353e-01	9.5353e-01	9.5353e-01
66	3	2	12	13	11	5.1816e-01	5.1816e-01	5.1816e-01
68	4	2	34	29	33	-9.2043e-01	-9.2043e-01	-9.2043e-01
69	4	2	28	13	17	-9.5671e+02	-9.5671e+02	-9.5671e+02
70	4	1	-	-	-	ERRO	ERRO	ITER
71	4	2	24	16	22	1.7014e+01	1.7014e+01	1.7014e+01
72	4	2	173	86	56	7.2768e+02	7.2768e+02	7.2768e+02
73	4	3	17	17	18	2.9894e+01	2.9894e+01	2.9894e+01
74	4	5	27	16	41	5.1265e+03	5.1265e+03	5.1265e+03
75	4	5	37	23	-	5.1744e+03	5.1744e+03	ITER

Tabela A.1: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	FDIPA	FAIPA	FDIPA-H	FDIPA	FAIPA	FDIPA-H
76	4	3	11	11	9	-4.6818e+00	-4.6818e+00	-4.6818e+00
77	5	2	20	21	9	2.4151e-01	2.4151e-01	2.4151e-01
78	5	3	10	10	7	-2.9197e+00	-2.9197e+00	-2.9197e+00
79	5	3	12	11	7	7.8777e-02	7.8777e-02	7.8777e-02
80	5	3	11	9	7	5.3950e-02	5.3950e-02	5.3950e-02
81	5	3	21	11	6	5.3950e-02	5.3950e-02	5.3950e-02
83	5	6	19	20	20	-3.0666e+04	-3.0666e+04	-3.0666e+04
84	5	6	22	15	18	-5.2803e+06	-5.1127e+06	-5.2803e+06
86	5	10	17	21	12	-3.2349e+01	-3.2349e+01	-3.2349e+01
88	2	1	28	20	24	1.3627e+00	1.3627e+00	1.3627e+00
89	3	1	28	22	33	1.3627e+00	1.3627e+00	1.3627e+00
90	4	1	22	20	26	1.3627e+00	1.3627e+00	1.3627e+00
91	5	1	28	30	31	1.3627e+00	1.3627e+00	1.3627e+00
92	6	1	24	31	37	1.3627e+00	1.3627e+00	1.3627e+00
93	6	2	18	15	12	1.3508e+02	1.3508e+02	1.3508e+02
95	6	4	9	9	8	1.5620e-02	1.5620e-02	1.5620e-02
96	6	4	8	8	8	1.5620e-02	1.5620e-02	1.5620e-02
97	6	4	51*	37*	26*	4.0712e+00	4.0712e+00	4.0712e+00
98	6	4	36*	67	-	4.0712e+00	3.1358e+00	ITER
99	7	2	41	-	5	-8.3108e+08	ERRO	-8.3108e+08
100	7	4	35	16	50	6.8063e+02	6.8063e+02	6.8063e+02
101	7	6	58	-	40	1.8098e+03	ERRO	1.8098e+03
102	7	6	-	-	44	INICIAL	ERRO	9.1188e+02
103	7	6	79	-	52	5.4367e+02	ERRO	5.4367e+02
104	8	6	45	29	18	3.9512e+00	3.9512e+00	3.9512e+00
105	8	1	100	100	97	1.1384e+03	1.1384e+03	1.1384e+03
106	8	6	-	-	-	ITER	ITER	ITER
107	9	6	13	16	15	5.0541e+03	5.0541e+03	5.0541e+03
108	9	13	23*	19*	-	-5.0000e-01	-5.0000e-01	INICIAL
109	9	10	-	-	-	ITER	ITER	ITER
110	10	0	8	8	6	-4.5778e+01	-4.5778e+01	-4.5778e+01
111	10	3	100	100	13	-4.7760e+01	-4.7761e+01	-4.7761e+01
112	10	3	29	29	12	-4.7761e+01	-4.7761e+01	-4.7761e+01
113	10	8	24	26	19	2.4267e+01	2.4267e+01	2.4267e+01
114	10	11	-	100	-	ITER	-1.7687e+03	ITER
116	13	15	-	-	-	INICIAL	INICIAL	INICIAL
117	15	5	36	47	21	3.2349e+01	3.2349e+01	3.2349e+01
118	15	29	46	47	22	6.6482e+02	6.6482e+02	6.6482e+02
119	16	8	48	39	31	2.4490e+02	2.4490e+02	2.4490e+02

(\*): O algoritmo converge para um ponto diferente do reportado por Hock e Schittkowski [49].

ERRO: Falha por causa do sistema linear singular ou na busca linear.

ITER: Não converge em 100 iterações.

INICIAL: Não encontra ponto inicial viável.

Tabela A.2: Resultados obtidos para os problemas do CUTE, primeira parte.

Problema			Iterações			Objetivo		
prob	nvar	nres	KN	LQ	FD-H	KNITRO	LOQO	FDIPA-H
artif	5000	5000	112	41	15	1.9499e-13	1.1355e-19	0.0000e+00
arwhead	5000	0	5	23	5	1.3769e-09	1.3096e-09	0.0000e+00
aug2d	20400	9996	-	17	3	ERRO	1.6874e+06	1.6874e+06
aug2dc	20400	10194	-	19	24	ERRO	1.8184e+06	1.8184e+06
aug2dcqp	20400	30594	-	28	-	ERRO	6.4982e+06	TEMPO
aug2dqp	20400	30594	-	27	-	ERRO	6.2370e+06	TEMPO
bdexp	5000	0	13	33	15	1.7195e-03	4.5495e-10	7.4146e-06
bdqrtic	1000	0	12	13	11	3.9838e+03	3.9838e+03	3.9838e+03
bdvalue	5002	5002	1	8	3	9.9178e-12	9.2947e-12	0.0000e+00
bigbank	2230	1420	38	36	66	-4.2057e+06	-4.2057e+06	-4.2057e+06
biggsb1	1000	0	23	30	32	1.5125e-02	1.5000e-02	1.5001e-02
blockqp1	2005	1001	19	18	20	-9.9649e+02	-9.9650e+02	-9.9650e+02
blockqp2	2005	1001	18	12	18	-9.9609e+02	-9.9510e+02	-9.9610e+02
blockqp3	2005	1001	123	33	20	-4.9749e+02	-4.9750e+02	-4.9750e+02
blockqp4	2005	1001	17	18	19	-4.9809e+02	-4.9810e+02	-4.9810e+02
blockqp5	2005	1001	104	33	58	-4.9749e+02	-4.9750e+02	-4.9749e+02
bloweya	2002	1002	12	76	10	-1.5634e-02	-4.5526e-02	-4.5531e-02
bloweyb	2002	1002	14	138	7	2.7125e-02	-3.0453e-02	-3.0453e-02
bloweyc	2002	1002	7	31	10	-1.3252e-02	-3.0298e-02	-3.0364e-02
bratu1d	1003	2	9	14	20	-8.5189e+00	-8.5189e+00	-8.5189e+00
bratu2d	5184	5184	4	11	4	1.9359e-10	4.6797e-16	0.0000e+00
bratu2dt	5184	5184	4	14	-	8.0199e-05	1.2300e-08	ERRO
bratu3d	4913	4913	5	11	4	9.1027e-13	1.9789e-18	0.0000e+00
broydn3d	10000	10000	7	13	5	2.6310e-13	4.7086e-23	0.0000e+00
broydn7d	1000	0	-	-	71	ERRO	ERRO	4.6918e+02
broydnbd	5000	5000	10	15	6	1.4422e-16	6.1768e-24	0.0000e+00
brybnd	5000	0	10	15	12	1.4422e-16	6.1768e-24	2.2105e-12
catenary	501	170	-	41	1968	ERRO	-3.4840e+05	-3.4840e+05
cbratu2d	1058	1058	1	8	3	4.5931e-12	3.7410e-16	0.0000e+00
chainwoo	1000	0	-	60	23	ERRO	6.3625e+01	4.5728e+00
chenhark	1000	0	646	18	15	-1.9948e+00	-2.0000e+00	-2.0000e+00
clnlbeam	1503	1004	22	119	-	3.4488e+02	3.4488e+02	ITER
clplatea	5041	71	14	10	7	-1.2592e-02	-1.2592e-02	-1.2592e-02
clplateb	5041	71	59	12	6	-6.9882e+00	-6.9882e+00	-6.9882e+00
clplatec	5041	71	3	9	3	-5.0207e-03	-5.0207e-03	-5.0207e-03
corkscrw	9006	7009	-	33	-	ERRO	9.0688e+01	ERRO
cosine	10000	0	33	13	7	-9.9990e+03	-9.9990e+03	-9.9990e+03
cragglvy	5000	0	15	17	19	1.6882e+03	1.6882e+03	1.6882e+03
curly10	10000	0	23	18	13	-1.0032e+06	-1.0032e+06	-1.0025e+06
curly20	10000	0	-	18	27	ERRO	-1.0032e+06	-1.0026e+06
curly30	10000	0	?	18	19	-1.0032e+06	-1.0032e+06	-1.0021e+06
cvxbqp1	10000	0	-	18	14	ERRO	2.2502e+06	2.2502e+06
cvxqp1	1000	500	12	29	-	1.0875e+06	1.0875e+06	ITER
cvxqp2	10000	2500	-	25	20	ERRO	8.1842e+07	8.1842e+07

Tabela A.2: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	KN	LQ	FD-H	KNITRO	LOQO	FDIPA-H
cvxqp3	10000	7500	-	-	-	ERRO	TEMPO	TEMPO
dixmaana	3000	0	8	12	4	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanb	3000	0	7	13	6	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanc	3000	0	9	14	6	1.0000e+00	1.0000e+00	1.0000e+00
dixmaand	3000	0	10	15	6	1.0000e+00	1.0000e+00	1.0000e+00
dixmaane	3000	0	36	18	5	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanf	3000	0	?	18	17	1.0000e+00	1.0000e+00	1.0000e+00
dixmaang	3000	0	33	20	34	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanb	3000	0	35	22	51	1.0000e+00	1.0000e+00	1.0000e+00
dixmaani	3000	0	24	18	13	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanb	3000	0	62	22	29	1.0000e+00	1.0000e+00	1.0000e+00
dixmaank	3000	0	55	24	16	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanl	3000	0	61	24	23	1.0000e+00	1.0000e+00	1.0000e+00
dqdrtic	5000	0	8	12	3	5.7374e-18	1.3534e-17	6.0779e-62
dqrtic	5000	0	38	64	35	1.1827e-05	1.6546e-17	2.9732e-08
drcav1lq	10816	10816	-	-	-	TEMPO	TEMPO	TEMPO
drcav2lq	10816	10816	-	-	-	TEMPO	TEMPO	TEMPO
drcav3lq	10816	10816	-	-	-	TEMPO	TEMPO	TEMPO
drcavty1	10816	204	-	-	-	TEMPO	TEMPO	ERRO
drcavty2	10816	204	-	-	-	TEMPO	TEMPO	TEMPO
drcavty3	10816	204	-	-	-	TEMPO	TEMPO	TEMPO
dtoc1l	14995	10000	-	17	6	ERRO	1.2534e+02	1.2534e+02
dtoc1na	1495	1000	7	14	6	1.2702e+01	1.2702e+01	1.2702e+01
dtoc1nb	1495	1000	6	12	6	1.5938e+01	1.5938e+01	1.5938e+01
dtoc1nc	1495	1000	10	10	50	2.4970e+01	2.4970e+01	2.4970e+01
dtoc1nd	745	500	23	31	2138	1.2644e+01	1.2644e+01	1.3169e+01
dtoc2	5998	4000	5	20	6	5.0868e-01	5.0868e-01	5.3151e-01
dtoc3	14999	10000	-	54	3	ERRO	2.3526e+02	2.3526e+02
dtoc4	14999	10000	-	20	-	ERRO	2.8685e+00	TEMPO
dtoc5	9999	5000	5	19	179	1.5337e+00	1.5351e+00	1.5351e+00
dtoc6	9999	5000	13	27	12	1.3485e+05	1.3485e+05	1.3482e+05
edensch	2000	0	9	11	9	1.2003e+04	1.2003e+04	1.2003e+04
eigena2	110	55	70	24	2	1.1538e-05	2.2962e-09	1.2622e-29
engval1	5000	0	10	16	10	5.5487e+03	5.5487e+03	5.5487e+03
fletcbv3	10000	0	-	-	-	ERRO	ERRO	ITER
fminsrf2	15625	0	-	394	19	ERRO	1.0000e+00	1.0000e+00
fminsurf	1024	0	293	55	17	1.0000e+00	1.0000e+00	1.0000e+00
freuroth	5000	0	12	21	7	6.0816e+05	6.0816e+05	6.0816e+05
gausselm	1496	4219	136	-	36	-1.7491e+01	ERRO	-1.0000e+00
gilbert	1000	1	33	37	21	4.8203e+02	4.8203e+02	4.8203e+02
gouldqp2	699	349	22	33	25	1.9400e-04	1.8800e-04	1.8801e-04
gouldqp3	699	349	22	23	-	2.0654e+00	2.0652e+00	ITER
hager1	10001	5001	4	10	3	8.8080e-01	8.8080e-01	8.8080e-01
hager2	10001	5001	4	10	3	4.3208e-01	4.3208e-01	4.3208e-01

Tabela A.2: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	KN	LQ	FD-H	KNITRO	LOQO	FDIPA-H
hager3	10001	5001	4	21	3	1.4096e-01	1.4096e-01	1.4096e-01
hager4	10001	5001	-	17	21	ERRO	2.7940e+00	2.7932e+00
hues-mod	10000	2	-	283	-	ERRO	3.4824e+07	TEMPO
huestis	10000	2	-	63	-	ERRO	3.4824e+11	TEMPO
indef	1000	0	-	-	-	ERRO	ERRO	ITER
kissing	127	903	87	33	85	8.4464e-01	1.0000e+00	1.0000e+00
ksip	20	1001	33	47	19	5.7580e-01	5.7580e-01	5.7580e-01
liarwhd	10000	0	13	22	13	1.1049e-09	7.9714e-23	0.0000e+00
lminsurf	15625	496	-	148	24	ERRO	9.0000e+00	9.0000e+00
manne	1095	731	-	-	-	TEMPO	ERRO	ITER
mccormck	50000	0	-	10	10	ERRO	-4.5662e+04	-4.5662e+04
morebv	5002	2	1	8	2	9.9198e-12	9.2971e-12	5.8083e-15
mosarqp1	2500	700	28	18	18	-9.5287e+02	-9.5288e+02	-9.5288e+02
mosarqp2	900	600	24	22	17	1.5975e+03	-1.5975e+03	-1.5954e+03
msqrta	1024	1024	48	33	6	7.0868e-12	8.1984e-20	0.0000e+00
msqrta	1024	1024	47	26	6	4.0102e-11	2.6933e-19	0.0000e+00
nondia	10000	0	5	13	17	5.7555e-10	2.4635e-22	0.0000e+00
nondquar	10000	0	163	24	19	3.1020e-06	7.1767e-11	7.1432e-11
nonscomp	10000	0	-	33	33	ERRO	2.4094e-09	1.9841e-11
optcdeg2	1202	803	29	74	-	2.2959e+02	2.2957e+02	ITER
optcdeg3	1202	803	36	55	-	4.6148e+01	4.6146e+01	ITER
orthrdm2	4003	2000	6	333	-	1.5553e+02	1.0539e+04	TEMPO
orthregc	10005	5000	-	23	-	ERRO	1.8959e+02	TEMPO
orthregd	10005	5000	-	442	-	ERRO	4.2458e+04	TEMPO
orthrgdm	10003	5000	-	12	-	ERRO	1.5138e+03	TEMPO
orthrgds	10003	5000	-	295	-	ERRO	2.6035e+04	TEMPO
penalty1	1000	0	46	56	30	9.6862e-03	9.6862e-03	9.6862e-03
pentdi	1000	0	19	20	14	7.4500e-01	-7.5000e-01	-7.5000e-01
porous1	5184	5184	183	40	9	1.6060e-15	7.8150e-13	0.0000e+00
porous2	5184	5184	356	38	10	2.3851e-17	2.1622e-20	0.0000e+00
powell20	1000	1000	55	27	-	5.2146e+07	5.2146e+07	INICIAL
power	1000	0	8	13	3	4.0154e-15	7.2502e-20	1.5109e-56
quartc	10000	0	40	69	38	3.6890e-05	2.4061e-18	3.3382e-08
reading1	10002	5001	-	225	-	ERRO	-1.6048e-01	TEMPO
reading2	15003	10002	-	24	2	ERRO	-1.2582e-02	-2.6444e-03
scon1dls	1002	2	-	292	-	TEMPO	4.6155e-20	ITER
scosine	10000	0	-	80	43	ERRO	-9.9990e+03	-9.9312e+03
scurlly10	10000	0	-	103	31	TEMPO	-1.0032e+06	-3.7678e+05
scurlly20	10000	0	-	94	93	TEMPO	-1.0032e+06	-9.7475e+05
scurlly30	10000	0	-	88	89	TEMPO	-1.0032e+06	-9.6392e+05
semicon1	1002	1002	-	292	-	TEMPO	4.6155e-20	ERRO
semicon2	1002	1002	-	70	34	TEMPO	1.4906e-16	0.0000e+00
sinquad	10000	0	217	70	259	2.2195e-06	6.1655e-12	5.1591e-14
sinrosnb	1000	2000	-	6	2	ERRO	-9.9901e+04	-9.9901e+04

Tabela A.2: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	KN	LQ	FD-H	KNITRO	LOQO	FDIPA-H
sipow1	2	10000	-	17	16	ERRO	-1.0000e+00	-1.0000e+00
sipow1m	2	10000	-	16	14	ERRO	-1.0000e+00	-1.0000e+00
sipow2	2	5001	-	16	15	-1.0000e+00	-1.0000e+00	-1.0000e+00
sipow2m	2	5001	22	16	15	-1.0000e+00	-1.0000e+00	-1.0000e+00
sipow3	4	10000	-	19	12	ERRO	5.3565e-01	5.3569e-01
sipow4	4	10000	-	19	14	ERRO	2.7283e-01	2.7283e-01
sosqp2	20000	10001	-	18	-	ERRO	-4.9987e+03	TEMPO
sreadin3	10002	5002	-	41	-	ERRO	-7.3083e-05	TEMPO
srosenbr	10000	0	26	28	21	8.5935e-09	1.3791e-21	1.1155e-26
svanberg	5000	5000	-	20	45	ERRO	8.3614e+03	8.3614e+03
tfi2	3	10001	-	25	44	ERRO	6.4904e-01	6.4904e-01
trainf	20008	10010	-	81	-	ERRO	3.1034e+00	TEMPO
trainh	20008	10010	-	92	-	ERRO	1.2370e+01	TEMPO
tridia	10000	0	8	12	3	5.2351e-13	9.9931e-19	6.1973e-57
ubh1	18009	12012	-	35	30	ERRO	1.1160e+00	1.1160e+00
ubh5	20010	14013	-	285	24	ERRO	1.1160e+00	3.7200e-01
woods	10000	0	63	48	12	1.3249e-14	4.4737e-22	2.2393e-25
yao	2002	2003	20	202	-	1.8695e+02	1.9770e+02	ITER

ERRO: Falha por causa do sistema linear singular ou na busca linear.

ITER: Não converge em 3000 iterações.

TEMPO: Não converge por superar o limite de tempo de 20 minutos.

INICIAL: Não encontra ponto inicial viável.

Tabela A.3: Resultados obtidos para os problemas do CUTE, segunda parte.

Problema			Iterações			Objetivo		
prob	nvar	nres	SN	FD	FA	SNOPT	FDIPA	FAIPA
artif	5000	5000	-	16	-	TEMPO	0.0000e+00	TEMPO
arwhead	5000	0	-	8	8	TEMPO	0.0000e+00	0.0000e+00
aug2d	20400	9996	-	-	-	TEMPO	MEMO	MEMO
aug2dc	20400	10194	-	-	-	TEMPO	MEMO	MEMO
aug2dcqp	20400	30594	-	-	-	TEMPO	MEMO	MEMO
aug2dqp	20400	30594	-	-	-	TEMPO	MEMO	MEMO
bdexp	5000	0	-	20	20	TEMPO	2.5293e-05	2.5293e-05
bdqrtic	1000	0	1194	84	84	3.9838e+03	3.9838e+03	3.9838e+03
bdvalue	5002	5002	0	6	6	1.0394e-11	0.0000e+00	0.0000e+00
bigbank	2230	1420	-	-	-	TEMPO	ERRO	ERRO
biggsb1	1000	0	-	-	-	ERRO	TEMPO	TEMPO
blockqp1	2005	1001	2008	-	-	-9.9650e+02	TEMPO	TEMPO
blockqp2	2005	1001	2362	-	12	-9.9610e+02	ERRO	-9.9610e+02
blockqp3	2005	1001	2015	-	-	-4.9750e+02	TEMPO	TEMPO
blockqp4	2005	1001	2356	-	-	-4.9810e+02	ERRO	ERRO
blockqp5	2005	1001	2021	-	-	-4.9750e+02	TEMPO	TEMPO
bloweya	2002	1002	603	27	27	-8.0400e-06	-4.0188e-02	-4.0188e-02
bloweyb	2002	1002	401	13	13	-3.2458e-20	-2.7205e-02	-2.7205e-02
bloweyc	2002	1002	402	26	26	-3.2080e-05	-2.6972e-02	-2.6972e-02
bratu1d	1003	2	-	-	-	ERRO	TEMPO	TEMPO
bratu2d	5184	5184	-	11	10	TEMPO	0.0000e+00	0.0000e+00
bratu2dt	5184	5184	-	-	-	TEMPO	ERRO	TEMPO
bratu3d	4913	4913	-	12	12	TEMPO	0.0000e+00	0.0000e+00
broydn3d	10000	10000	-	-	-	TEMPO	MEMO	MEMO
broydn7d	1000	0	2867	449	449	3.8234e+00	3.8964e+02	3.8964e+02
broydnbd	5000	5000	-	12	12	TEMPO	0.0000e+00	0.0000e+00
brybnd	5000	0	-	165	165	TEMPO	3.0545e+00	3.0545e+00
catenary	501	170	-	-	-	TEMPO	ERRO	ERRO
cbratu2d	1058	1058	2417	9	9	2.1754e-08	0.0000e+00	0.0000e+00
chainwoo	1000	0	-	443	443	ERRO	1.0000e+00	1.0000e+00
chenhark	1000	0	-	-	-	ERRO	TEMPO	TEMPO
clnlbeam	1503	1004	1466	-	-	3.5000e+02	ERRO	ERRO
clplatea	5041	71	-	1629	1629	TEMPO	-1.2592e-02	-1.2592e-02
clplateb	5041	71	-	665	672	TEMPO	-6.9882e+00	-7.1311e+00
clplatec	5041	71	-	-	-	TEMPO	ITER	ITER
corkscrw	9006	7009	46387	-	-	9.0688e+01	MEMO	MEMO
cosine	10000	0	-	20	20	TEMPO	-9.9990e+03	-9.9990e+03
cragglvy	5000	0	-	-	-	TEMPO	ERRO	ERRO
curly10	10000	0	-	-	-	TEMPO	ITER	ITER
curly20	10000	0	-	-	-	TEMPO	ITER	ITER
curly30	10000	0	-	-	-	TEMPO	ITER	ITER
cvxbqp1	10000	0	10000	-	-	2.2502e+06	MEMO	MEMO
cvxqp1	1000	500	1593	-	-	1.0875e+06	TEMPO	TEMPO
cvxqp2	10000	2500	-	-	-	TEMPO	MEMO	MEMO

Tabela A.3: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	SN	FD	FA	SNOPT	FDIPA	FAIPA
cvxqp3	10000	7500	10217	-	-	1.1571e+08	MEMO	MEMO
dixmaana	3000	0	3007	8	8	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanb	3000	0	3022	15	15	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanc	3000	0	3028	31	31	1.0000e+00	1.0000e+00	1.0000e+00
dixmaand	3000	0	3031	39	39	1.0000e+00	1.0000e+00	1.0000e+00
dixmaane	3000	0	3170	374	374	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanf	3000	0	3169	333	333	1.0000e+00	1.0000e+00	1.0000e+00
dixmaang	3000	0	3177	203	203	1.0000e+00	1.0000e+00	1.0000e+00
dixmaanb	3000	0	3130	381	381	1.0000e+00	1.0000e+00	1.0000e+00
dixmaani	3000	0	-	-	-	TEMPO	ITER	ITER
dixmaanb	3000	0	3149	142	142	1.0902e+00	1.1770e+00	1.1770e+00
dixmaank	3000	0	-	130	130	TEMPO	1.5337e+00	1.5337e+00
dixmaanl	3000	0	3293	-	-	1.0000e+00	ITER	ITER
dqdrtic	5000	0	-	6	6	TEMPO	7.6787e-43	7.6787e-43
dqrtic	5000	0	-	177	177	TEMPO	2.5260e-07	2.5260e-07
drcav1lq	10816	10816	-	-	-	TEMPO	MEMO	MEMO
drcav2lq	10816	10816	-	-	-	TEMPO	MEMO	MEMO
drcav3lq	10816	10816	-	-	-	TEMPO	MEMO	MEMO
drcavty1	10816	204	-	-	-	TEMPO	TEMPO	TEMPO
drcavty2	10816	204	-	-	-	TEMPO	TEMPO	TEMPO
drcavty3	10816	204	-	-	-	TEMPO	TEMPO	TEMPO
dtoc1l	14995	10000	-	-	-	TEMPO	MEMO	MEMO
dtoc1na	1495	1000	2347	15	15	1.2702e+01	1.2702e+01	1.2702e+01
dtoc1nb	1495	1000	2750	16	16	1.5938e+01	1.5938e+01	1.5938e+01
dtoc1nc	1495	1000	2702	30	28	2.4970e+01	2.4970e+01	2.4970e+01
dtoc1nd	745	500	1070	86	62	1.2502e+01	1.2646e+01	1.2733e+01
dtoc2	5998	4000	4594	-	-	5.0868e-01	TEMPO	TEMPO
dtoc3	14999	10000	11271	-	-	2.3541e+02	MEMO	MEMO
dtoc4	14999	10000	-	-	-	TEMPO	MEMO	MEMO
dtoc5	9999	5000	-	-	-	TEMPO	MEMO	TEMPO
dtoc6	9999	5000	-	-	-	TEMPO	MEMO	TEMPO
edensch	2000	0	2035	41	41	1.2003e+04	1.2003e+04	1.2003e+04
eigena2	110	55	197	-	-	1.7785e-12	TEMPO	TEMPO
engval1	5000	0	-	34	34	TEMPO	5.5487e+03	5.5487e+03
fletcbv3	10000	0	-	-	-	TEMPO	ERRO	ERRO
fminsrf2	15625	0	-	496	496	TEMPO	1.0000e+00	1.0000e+00
fminsurf	1024	0	1337	214	214	1.0000e+00	1.0000e+00	1.0000e+00
freuroth	5000	0	-	-	-	TEMPO	ERRO	ERRO
gausselm	1496	4219	2321	-	-	-1.0000e+00	TEMPO	TEMPO
gilbert	1000	1	1046	51	51	4.8203e+02	4.8203e+02	4.8203e+02
gouldqp2	699	349	456	-	-	1.8875e-04	TEMPO	TEMPO
gouldqp3	699	349	469	-	-	2.0652e+00	TEMPO	TEMPO
hager1	10001	5001	-	-	-	TEMPO	TEMPO	TEMPO
hager2	10001	5001	-	16	16	TEMPO	4.3208e-01	4.3208e-01

Tabela A.3: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	SN	FD	FA	SNOPT	FDIPA	FAIPA
hager3	10001	5001	-	16	16	TEMPO	1.4096e-01	1.4096e-01
hager4	10001	5001	-	-	-	TEMPO	MEMO	MEMO
hues-mod	10000	2	-	-	-	TEMPO	MEMO	MEMO
huestis	10000	2	-	-	-	TEMPO	MEMO	MEMO
indef	1000	0	-	-	-	ERRO	ERRO	ERRO
kissing	127	903	-	12	16	ERRO	1.0000e+00	1.0000e+00
ksip	20	1001	2438	42	42	5.7580e-01	5.7580e-01	5.7580e-01
liarwhd	10000	0	-	19	19	TEMPO	3.7114e-20	3.7114e-20
lminsurf	15625	496	-	-	-	TEMPO	TEMPO	TEMPO
manne	1095	731	729	-	-	-9.7457e-01	ERRO	ERRO
mccormck	50000	0	-	-	-	TEMPO	MEMO	MEMO
morebv	5002	2	0	2	2	1.0395e-11	1.0392e-11	1.0392e-11
mosarqp1	2500	700	4578	24	24	-9.5288e+02	-9.5288e+02	-9.5288e+02
mosarqp2	900	600	2574	183	161	-1.5975e+03	-1.5954e+03	-1.5954e+03
msqrta	1024	1024	-	-	-	ERRO	ERRO	ERRO
msqrta	1024	1024	-	-	-	ERRO	ERRO	ERRO
nondia	10000	0	-	21	21	TEMPO	1.1552e-22	1.1552e-22
nondquar	10000	0	-	2047	2047	TEMPO	2.6077e-06	2.6077e-06
nonscomp	10000	0	-	-	-	TEMPO	MEMO	MEMO
optcdeg2	1202	803	2902	-	-	2.2957e+02	TEMPO	TEMPO
optcdeg3	1202	803	2778	-	-	4.6146e+01	TEMPO	TEMPO
orthrdm2	4003	2000	4023	-	-	1.5553e+02	ERRO	ERRO
orthregc	10005	5000	-	-	-	TEMPO	TEMPO	TEMPO
orthregd	10005	5000	-	-	-	TEMPO	ERRO	ERRO
orthrgdm	10003	5000	-	-	-	TEMPO	TEMPO	ERRO
orthrgds	10003	5000	-	-	-	TEMPO	ERRO	ERRO
penalty1	1000	0	1170	69	-	9.6863e-03	9.6862e-03	ITER
pentdi	1000	0	4	-	-	-7.5000e-01	TEMPO	TEMPO
porous1	5184	5184	-	-	-	TEMPO	ERRO	TEMPO
porous2	5184	5184	-	15	-	TEMPO	0.0000e+00	TEMPO
powell20	1000	1000	501	11	11	5.2146e+07	6.0275e-07	1.5159e-06
power	1000	0	2626	1527	1527	2.5732e-16	2.1904e-16	2.1904e-16
quartc	10000	0	-	192	192	TEMPO	1.3784e-07	1.3784e-07
reading1	10002	5001	-	-	-	TEMPO	MEMO	MEMO
reading2	15003	10002	17172	-	-	-8.1409e-03	MEMO	MEMO
scon1dls	1002	2	-	-	-	ERRO	TEMPO	TEMPO
scosine	10000	0	-	-	-	TEMPO	ERRO	ERRO
scurly10	10000	0	-	-	-	TEMPO	ERRO	ERRO
scurly20	10000	0	-	-	-	TEMPO	ERRO	ERRO
scurly30	10000	0	-	-	-	TEMPO	ERRO	ERRO
semicon1	1002	1002	-	-	-	ERRO	ERRO	ERRO
semicon2	1002	1002	-	-	-	ERRO	ERRO	ERRO
sinquad	10000	0	-	297	297	TEMPO	8.8351e-13	8.8351e-13
sinrosnb	1000	2000	0	2	2	-9.9901e+04	-9.9901e+04	-9.9901e+04

Tabela A.3: Continuação da página anterior.

Problema			Iterações			Objetivo		
prob	nvar	nres	SN	FD	FA	SNOPT	FDIPA	FAIPA
sipow1	2	10000	850	-	-	-1.0000e+00	MEMO	MEMO
sipow1m	2	10000	851	-	-	-1.0000e+00	MEMO	MEMO
sipow2	2	5001	739	11	24	-1.0000e+00	-1.0000e+00	-1.0000e+00
sipow2m	2	5001	739	11	18	-1.0000e+00	-1.0000e+00	-9.9999e-01
sipow3	4	10000	75	-	-	5.3565e-01	MEMO	MEMO
sipow4	4	10000	20	-	-	2.7283e-01	MEMO	MEMO
sosqp2	20000	10001	-	-	-	TEMPO	MEMO	MEMO
sreadin3	10002	5002	-	-	-	TEMPO	MEMO	MEMO
srosenbr	10000	0	-	36	36	TEMPO	9.3799e-22	9.3799e-22
svanberg	5000	5000	-	-	-	TEMPO	MEMO	MEMO
tfi2	3	10001	62	-	-	6.4903e-01	MEMO	MEMO
trainf	20008	10010	-	-	-	TEMPO	MEMO	MEMO
trainh	20008	10010	-	-	-	TEMPO	MEMO	MEMO
tridia	10000	0	-	1098	1098	TEMPO	1.3632e-14	1.3632e-14
ubh1	18009	12012	-	-	-	TEMPO	MEMO	MEMO
ubh5	20010	14013	-	-	-	TEMPO	MEMO	MEMO
woods	10000	0	-	29	29	TEMPO	5.6261e-14	5.6261e-14
yao	2002	2003	2	-	-	2.7313e+02	TEMPO	TEMPO

ERRO: Falha por causa do sistema linear singular ou na busca linear.

ITER: Não converge em 3000 iterações.

TEMPO: Não converge por superar o limite de tempo de 20 minutos.

INICIAL: Não encontra ponto inicial viável.

MEMO: Supera a capacidade de memória do computador.