

UNIVERSIDADE FEDERAL DE GOIÁS
IME - INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
PROGRAMA DE PÓS GRADUAÇÃO EM MATEMÁTICA

**SECANT-TYPE METHODS WITH FEASIBLE
INEXACT PROJECTION FOR SOLVING
CONSTRAINED MIXED GENERALIZED
EQUATIONS**

DOCTORAL THESIS BY
PAULO CÉSAR DA SILVA JÚNIOR

SUPERVISED BY
Prof. Dr. **Orizon Pereira Ferreira**

GOIÂNIA, GOIÁS, BRAZIL
AGOSTO 2025



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese Outro*: _____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Paulo César da Silva Júnior

3. Título do trabalho

Secant-type method with feasible inexact projection for solving constrained mixed generalized equations

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Orizon Pereira Ferreira, Professora do Magistério Superior**, em 25/08/2025, às 07:39, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paulo Cesar Da Silva Junior, Discente**, em 27/08/2025, às 10:21, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5598071** e o código CRC **2B960B05**.

PAULO CÉSAR DA SILVA JÚNIOR

SECANT-TYPE METHODS WITH FEASIBLE INEXACT PROJECTION
FOR SOLVING CONSTRAINED MIXED GENERALIZED EQUATIONS

Tese apresentada ao Programa de Pós-Graduação em
Matemática do Instituto de Matemática e Estatística da
Universidade Federal de Goiás, como requisito parcial
para obtenção do título de Doutor em Matemática.

Área de concentração: Otimização

Orientador: Prof. Dr. Orizon Pereira Ferreira

Coorientador: Prof. Dr. Gilson do Nascimento Silva

Goiânia
2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Silva Júnior, Paulo César da

Secant-type method with feasible inexact projection for solving constrained mixed generalized equations [manuscrito] / Paulo César da Silva Júnior. - 2025.

LXXXI, 81 f.

Orientador: Prof. Dr. Orizon Pereira Ferreira; co-orientador Dr. Gilson do Nascimento Silva.

Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Matemática e Estatística (IME), Programa de Pós-Graduação em Matemática, Goiânia, 2025.

Bibliografia.

1. Equação generalizada mista. 2. Métodos do tipo Newton. 3. Projeções inexatas. 4. Regularidade métrica. 5. Princípio da Contração. I. Ferreira, Orizon Pereira, orient. II. Título.



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

ATA DE DEFESA DE TESE

Ata Nº 26 da sessão de Defesa de Tese de **Paulo César da Silva Júnior** que confere o título de Doutor em **Matemática**, na área de concentração em **Otimização**.

Ao **vigésimo primeiro dia do mês de agosto do ano de dois mil e vinte e cinco**, a partir das **14h00**, via web vídeo conferência, realizou-se a sessão pública de Defesa de Tese intitulada “**Secant-type method with feasible inexact projection for solving constrained mixed generalized equations**”. Os trabalhos foram instalados pelo Orientador, Professor Doutor **Orizon Pereira Ferreira - IME/UFG** com a participação dos demais membros da Banca Examinadora: Professor Doutor **Gilson do Nascimento Silva- DM/UFPI (coorientador)**, membro titular externo; Professor Doutor **Jefferson Divino Gonçalves de Melo - IME/UFG**, membro titular interno; Professor Doutor **Leandro da Fonseca Prudente - IME/UFG**, membro titular interno; Professor Doutor **Roberto Andreani - IMECC/UNICAMP**, membro titular externo e Professor Doutor **Paulo Sérgio Marques dos Santos - DEMAT/UFDP**, membro titular externo. Durante a arguição os membros da banca **não fizeram** sugestão de alteração do título do **trabalho**. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor **Orizon Pereira Ferreira**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, ao **vigésimo primeiro dia do mês de agosto do ano de dois mil e vinte e cinco**.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Jefferson Divino Goncalves De Melo, Professor do Magistério Superior**, em 21/08/2025, às 16:05, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leandro Da Fonseca Prudente, Professor do Magistério Superior**, em 21/08/2025, às 16:05, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Orizon Pereira Ferreira, Professora do Magistério Superior**, em 21/08/2025, às 16:17, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Roberto Andreani, Usuário Externo**, em 22/08/2025, às 07:42, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paulo Sérgio Marques dos Santos, Usuário Externo**, em 22/08/2025, às 10:37, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gilson do Nascimento Silva, Usuário Externo**, em 22/08/2025, às 13:04, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5478876** e o código CRC **E4F3DCD8**.

Referência: Processo nº 23070.034423/2025-61

SEI nº 5478876

Dedicado a:

TODAS AS PESSOAS QUE CONTRIBUÍRAM PARA A
REALIZAÇÃO DESTE TRABALHO, EM ESPECIAL,
MINHAS MÃES, MARIA DE SOUZA (IN MEMORIAM) E MARIA
MARQUES DA SILVA
MINHA IRMÃ, FLÁVIA CRISTINA DA SILVA ALVES
MEU ORIENTADOR, ORIZON PEREIRA FERREIRA
MEU COORIENTADOR, GILSON DO NASCIMENTO SILVA

Agradecimentos

Primeiramente, agradeço a Deus por me permitir realizar mais este sonho, por Sua presença constante em minha vida, por me conceder sabedoria para enfrentar as dificuldades e forças para jamais desistir.

À minha família e aos amigos, em especial à minha saudosa mãezinha Maria de Souza, à minha segunda mãe Maria Marques da Silva e à minha irmã Flávia Cristina da Silva Alves, meu porto seguro, que estiveram ao meu lado em todos os momentos. Agradeço por todo amor, apoio, palavras de incentivo e, sobretudo, por sempre acreditarem em mim.

Ao meu orientador, professor Orizon Pereira Ferreira, expresso minha sincera gratidão por todo o conhecimento compartilhado, pelo direcionamento firme e encorajador, pela confiança, paciência incansável e disponibilidade para esclarecer dúvidas. Agradeço, ainda, por sua dedicação ao longo do desenvolvimento desta tese e por ter me acolhido como orientando.

Ao meu coorientador, professor Gilson do Nascimento Silva, agradeço pela valiosa contribuição ao trabalho.

Aos funcionários do IME/UFG, meu agradecimento pela cordialidade no dia a dia. Aos professores da instituição que contribuíram para minha formação acadêmica, meu sincero reconhecimento.

Estendo também meus agradecimentos aos docentes do grupo de Otimização do IME/UFG, pelas contribuições enriquecedoras durante os seminários.

Aos professores Leandro da Fonseca Prudente, Roberto Andreani, Paulo Sérgio Marques dos Santos e Jefferson D. Gonçalves de Melo, agradeço por aceitarem o convite para compor a banca examinadora e pelas observações e sugestões valiosas. Da mesma forma, agradeço aos professores suplentes João Carlos de Oliveira Souza e Maurício Silva Louzeiro.

Aos amigos que compartilharam comigo esta caminhada, pelos momentos de estudo, apoio e convivência durante esses anos, deixo meu muito obrigado.

E a todos que, de alguma forma, direta ou indiretamente, contribuíram para a realização desta tese, mesmo que não mencionados nominalmente, receba meu mais sincero agradecimento.

Abstract

This thesis addresses the solution of a mixed generalized equation of the form Find $x \in C$ such that $f(x) + g(x) + F(x) \ni 0$, where $f : \Omega \rightarrow \mathbb{R}^n$ is continuously differentiable, $g : \Omega \rightarrow \mathbb{R}^n$ is continuous (but not necessarily differentiable), $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping with closed, nonempty graph, $\Omega \subseteq \mathbb{R}^n$ is open, and $C \subset \Omega$ is a closed convex set.

To handle the non-smoothness of g and the complexity of the set-valued term F , this work introduces and analyzes two iterative methods based on partial linearizations and feasible inexact projections. The first is a secant-type method, which utilizes first- and second-order divided differences of g to approximate its local behavior. The second is a quasi-Newton method employing a Broyden update to approximate the Jacobian of f , thus avoiding its exact computation.

The convergence analysis of both methods is conducted under assumptions of metric regularity for the associated linearized mappings. Theoretical results establish local convergence of the sequences generated by the algorithms. For the secant method, it is shown that if the projection errors vanish, the convergence becomes superlinear, and in the exact projection case, the rate becomes quadratic. For the quasi-Newton method, assuming bounded deterioration of the Jacobian approximations, the sequence converges q -linearly.

The analysis is supported by variational tools such as metric regularity, strong metric regularity, linearization error bounds, the Contraction Principle Theorem and a key perturbed metric regularity theorem. These tools enable a rigorous treatment of the impact of approximation errors in both the projections and Jacobian estimates, ensuring robustness and stability of the proposed methods.

Overall, this thesis contributes new algorithms and a complete convergence theory for solving nonsmooth and set-valued generalized equations with convex constraints. The methods presented extend classical Newton-type strategies to more general, practical contexts involving nonsmooth components, variational terms, and computational inexactness.

Keywords: Mixed generalized equation, Secant-type method, Feasible inexact projection, Metric regularity, Contraction Principle, Partial linearization, Broyden rule, quasi-Newton method.

Resumo

Esta tese trata da solução de uma equação generalizada mista da forma Encontrar $x \in C$ tal que $f(x)+g(x)+F(x) \ni 0$, onde $f : \Omega \rightarrow \mathbb{R}^n$ é uma função continuamente diferenciável, $g : \Omega \rightarrow \mathbb{R}^n$ é uma função contínua (mas não necessariamente diferenciável), $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ é uma aplicação ponto-conjunto com gráfico fechado e não vazio, $\Omega \subseteq \mathbb{R}^n$ é um conjunto aberto, e $C \subset \Omega$ é um conjunto convexo e fechado.

Para lidar com a não suavidade de g e com a complexidade da aplicação ponto-conjunto F , este trabalho introduz e analisa dois métodos iterativos baseados em linearizações parciais e projeções inexatas viáveis. O primeiro é um método do tipo secante, que utiliza diferenças divididas de primeira e segunda ordem de g para aproximar seu comportamento local. O segundo é um método quasi-Newton que emprega a atualização de Broyden para aproximar o jacobiano de f , evitando, assim, sua computação exata.

A análise de convergência de ambos os métodos é realizada sob hipóteses de regularidade métrica para as aplicações linearizadas associadas. Resultados teóricos estabelecem a convergência local das sequências geradas pelos algoritmos. Para o método secante, mostra-se que, se os erros de projeção forem nulos, a convergência é superlinear, e no caso de projeções exatas, a taxa de convergência torna-se quadrática. Para o método quasi-Newton, assumindo deterioração limitada das aproximações do jacobiano, a sequência converge q -linearmente.

A análise é fundamentada em ferramentas variacionais como regularidade métrica, regularidade métrica forte, estimativas do erro de linearização, Teorema do Princípio da Contração e um teorema fundamental de regularidade métrica perturbada. Essas ferramentas permitem um tratamento rigoroso do impacto dos erros de aproximação tanto nas projeções quanto nas estimativas do jacobiano, garantindo robustez e estabilidade dos métodos propostos.

De forma geral, esta tese contribui com novos algoritmos e uma teoria completa de convergência para a resolução de equações generalizadas com componentes não suaves e multivalorados sujeitas a restrições convexas. Os métodos apresentados estendem estratégias clássicas do tipo Newton para contextos mais gerais e práticos, que envolvem não suavidade, termos variacionais e inexatidões computacionais.

Palavras-chave: Equação generalizada mista, Métodos do tipo Newton, Projeções inexatas, Regularidade métrica, Princípio da Contração, Linearização Parcial, Métodos quasi-Newton, Broyden.

Contents

1	Introduction	3
2	Basic results	11
2.1	First notations	11
2.2	First and second order divided differences	11
2.3	Distance and Excess	15
2.4	Metric Regularity	18
2.4.1	Linearization Error and Partial Linearization	23
2.4.2	Graph, Domain, Range and inverse of a set-valued function	23
2.4.3	Monotone maximal mapping	24
2.4.4	Aubin Property	28
2.4.5	The Contraction Principle Theorem	32
2.4.6	Perturbed Metric Regularity Theorem	33
2.5	The Feasible Inexact Projection Mapping	36
2.6	Broyden Update Rule	40
2.6.1	The Bounded Deterioration Condition	41
2.7	Convergence q-linear and Superlinear	43
3	Secant-type method with feasible inexact projection for solving constrained mixed generalized equations	45
3.1	The Secant-InexP Algorithm	45
3.2	The Convergence Theorem	50
3.2.1	The Uniqueness	52
3.2.2	Application	55

4	Broyden quasi-Newton secant-type method with feasible inexact projection for solving constrained mixed generalized equations	57
4.1	The Inexact Broyden quasi-Newton secant-type Algorithm	58
4.2	The Convergence Theorem	60
5	Conclusion	67
5.1	Summary of Contributions	67
5.2	Implications and Significance	68
5.3	Future Research Directions	68
5.4	Final Remarks	69

Chapter 1

Introduction

Structure of the Thesis

The primary objective of this thesis is to investigate novel versions of secant-type methods for solving constrained generalized equations. Specifically, the proposed method synergistically combines the traditional secant method, applied to generalized equations, with the conditional gradient method (also known as the Frank-Wolfe method). This fusion aims to leverage the strengths of both approaches to enhance convergence properties and computational efficiency in solving complex, constrained problems. To establish the convergence of the proposed algorithm, we employ the contraction mapping principle, a fundamental tool in fixed-point theory. Additionally, by assuming a Lipschitz condition on the gradient of the involved functions and utilizing the metric regularity property of the underlying mappings, we demonstrate that the sequence generated by our algorithm is well defined and exhibits local convergence. In particular, the convergence rate achieved is linear or superlinear, depending on the specific conditions satisfied by the problem.

The thesis is organized into several chapters, each addressing distinct aspects of the research problem:

- **Chapter 2:** This chapter revisits essential notation, definitions, and preliminary results that form the basis for subsequent analysis. This ensures that the reader is equipped with the necessary theoretical background to comprehend the advanced methods discussed in the later chapters.
- **Chapter 3:** Dedicated to the development of a *secant-type method* to solve *mixed generalized equations subject to a set of constraints*, this chapter introduces the problem formulation and the proposed iterative scheme. The primary focus is on addressing the following problem:

$$\text{Find } x \in C \text{ such that } f(x) + g(x) + F(x) \ni 0, \quad (1.1)$$

where:

- $f : \Omega \rightarrow \mathbb{R}^n$ is a continuously differentiable function,
- $g : \Omega \rightarrow \mathbb{R}^n$ is a continuous function (not necessarily differentiable),
- $\Omega \subseteq \mathbb{R}^n$ is an open set,
- $C \subset \Omega$ is a closed and convex set,
- $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping with a closed, nonempty graph.

Although the study of generalized equations began in the early 1970s with seminal contributions from S. M. Robinson [52, 53, 55, 56] and N. H. Josephy [41, 42], it remains a vibrant area of research today. Contemporary works, such as those by Adly, Van Ngai, and Nguyen [3], Gilson, Leo, and Andreani [4], Dontchev [22], and Dontchev and Rockafellar [24], continue to advance the theoretical and computational methodologies for solving generalized equations. It is investigated the following iterative scheme:

$$f(x_k) + g(x_k) + (f'(x_k) + [x_{k-1}, x_k; g])(y_k - x_k) + F(y_k) \ni 0, \quad k = 0, 1, \dots, \quad (1.2)$$

where $[x_{k-1}, x_k; g]$ represents the first order divided difference of the function g at the points x_{k-1}, x_k .

It is demonstrated that with some metrically regular hypothesis, some bounded properties and f' is Lipschitz continuous, the sequence $\{x_k\}$ generated by the presented algorithm converges linearly to x_* and sometimes the sequence $\{x_k\}$ generated by the presented algorithm converges superlinearly to x_* .

- **Chapter 4:** This chapter introduces a novel *secant-type method* augmented with a Broyden-type quasi-Newton update, inspired by the work of Roberto Andreani, Rui Marques Carvalho, Leonardo Delarmelina Secchin, and Gilson do Nascimento Silva [5]. The method addresses the same following constrained mixed generalized equation (1.1).

This chapter builds on Chapter 3 by integrating a secant-type iterative scheme with a quasi-Newton approach, thereby enhancing convergence properties while maintaining computational efficiency. The iterative process involves computing intermediate points that satisfy a linearized inclusion, followed by projecting these points onto the feasible set C using an inexact projection procedure inspired by the conditional gradient method. It is investigated the following iterative scheme:

$$f(x_k) + g(x_k) + (B_k + [x_{k-1}, x_k; g])(y_k - x_k) + F(y_k) \ni 0, \quad k = 0, 1, \dots, \quad (1.3)$$

where $\{B_k\}$ denotes a sequence of bounded linear mappings between \mathbb{R}^n , which adhere to the classical Broyden update rule. It is demonstrated that if the multifunction $f + F$ is metrically regular at the point x for 0, and if the derivative mapping f' is Lipschitz continuous, then the sequence $\{x_k\}$ generated by (1.3) converges linearly to x_* , as shown in [6, Theorem 4.3]. More broadly, Adly and Huynh [2] proposed quasi-Newton schemes analogous to (1.3) for solving (1.1), permitting the function f to be potentially non-differentiable. In this context, it is imposed a regularity metric condition concerning a specific form of semismooth regularization of $f + F$. It is established that if $\{B_k\}$ satisfies an appropriate modified Broyden update, then the sequence $\{x_k\}$ generated by (1.3) converges linearly to a solution x_* of (1.1) [2, Theorem 4.3]. A similar methodology was employed in [9, 43].

Problem Formulation and Motivation

The generalized equation framework, as encapsulated in equation (1.1), provides a unified mathematical model that captures a broad spectrum of problems, including nonlinear equation systems, equilibrium models, complementarity problems and variational inequality problems. Specifically, the generalized equation can be viewed as an abstraction that extends beyond traditional equation solving paradigms by incorporating set-valued mappings and additional nonlinear components.

A particular case of (1.1) arises when the constraints are relaxed or removed and $g(x) = 0$. Setting $C = \mathbb{R}^n$ simplifies (1.1) to an *unconstrained generalized equation*, formulated as:

$$\text{Find } x \in \mathbb{R}^n \text{ such that } f(x) + F(x) \ni 0. \quad (1.4)$$

This problem was first systematically studied in the works of Josephy [41, 42] and Robinson [54], laying the groundwork for subsequent advancements.

The constrained generalized equation (1.1) further broadens the applicability by introducing feasible regions defined by the convex set C . This inclusion is critical for modeling scenarios where solutions must adhere to specific constraints, such as non-negativity, boundedness, or more complex feasibility conditions. For example, if $g(x) = 0$ and F is the normal cone mapping N_D of a convex set $D \subset \mathbb{R}^n$, then (1.1) specializes to a *constrained variational inequality*:

$$\text{Find } x \in C \cap D \text{ such that } \langle f(x), y - x \rangle \geq 0, \quad \forall y \in D, \quad (1.5)$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product in \mathbb{R}^n .

A particularly notable instance of (1.5) is the *constrained nonlinear complementarity problem*, defined as:

$$\text{Find } x \in C \text{ such that } x \in \mathbb{R}_+^n, \quad f(x) \in \mathbb{R}_+^n, \quad \text{and } \langle f(x), x \rangle = 0,$$

where \mathbb{R}_+^n denotes the non-negative orthant in \mathbb{R}^n . The problem (1.4) and its variants have been extensively studied over the past decades, with significant contributions in the literature [14], highlighting their relevance in both theoretical research and practical applications.

When the constraint set C coincides with the entire space \mathbb{R}^n , the problem reduces to a *constrained system of nonlinear equations*. This reduction underscores the flexibility of generalized equations in encapsulating various types of problem. This system is pivotal in many disciplines, prompting the development of numerous numerical methods for its resolution [11, 33, 34, 44, 48].

Newton's Method for Generalized Equations

Newton's method, a cornerstone in numerical analysis, has been extended to tackle generalized equations due to its superior convergence properties. For the unconstrained generalized equation (1.4), Newton's iterative scheme is defined as:

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad \forall k \geq 0. \quad (1.6)$$

At each iteration, the method solves a *partially linearized inclusion* based on the current iterate x_k . This approach generalizes the traditional Newton's method for solving nonlinear equations by accommodating set-valued mappings and additional nonlinearities.

The versatility of (1.6) is evident in its ability to model various iterative procedures in numerical nonlinear programming:

- **Unconstrained Nonlinear Equations:** When $F \equiv \{0\}$, (1.6) simplifies to the classic Newton's method for solving $f(x) = 0$.
- **Nonlinear Equations with Inequality Constraints:** If F represents the product of the negative orthant \mathbb{R}_-^s in \mathbb{R}^s with the origin in \mathbb{R}^{m-s} , i.e., $F = \mathbb{R}_-^s \times \{0\}^{m-s}$, then (1.6) becomes Newton's method for solving systems of nonlinear equalities and inequalities [17].
- **Sequential Quadratic Programming (SQP):** When (1.4) represents the Karush-Kuhn-Tucker (KKT) optimality conditions for a nonlinear programming problem, (1.6) corresponds to the well-known SQP method [24, p. 384]; see also [21, 39].

The attractiveness of Newton-type methods lies in their superior local convergence properties, including superlinear rates under appropriate regularity conditions. Consequently, extensive research has been dedicated to refining these methods for generalized equations [6, 7, 13, 20, 23, 28–30, 51].

Secant-Type and Quasi-Newton Methods

Building upon the foundation of Newton's method, this thesis explores *secant-type methods* for solving constrained generalized equations. These methods approximate the Jacobian matrix $f'(x)$ iteratively, eliminating the need for explicit derivative computations and thereby enhancing computational efficiency, especially in large-scale problems.

In Chapter 3, we develop a secant-type method for the constrained mixed generalized equation (1.1), incorporating elements from both secant and conditional gradient methods. This hybrid approach is designed to maintain feasibility with respect to the constraint set C while effectively approximating the Jacobian. The iterative scheme for solving the constrained generalized equation (1.1) is given by:

$$f(x_k) + g(x_k) + (f'(x_k) + [x_{k-1}, x_k; g])(y_k - x_k) + F(y_k) \ni 0, \quad \forall k \geq 0, \quad (1.7)$$

where:

- $[x_{k-1}, x_k; g]$ denotes the first-order divided difference of the function g with respect to the points x_{k-1} and x_k ,
- y_k is the intermediate point computed at each iteration.

Chapter 4 introduces an enhanced *secant-type method* integrated with a *Broyden-type quasi-Newton update*, inspired by Andreani, Carvalho, Secchin, and Silva [5]. The iterative scheme for solving the constrained generalized equation (1.1) is given by:

$$f(x_k) + g(x_k) + (B_k + [x_{k-1}, x_k; g])(y_k - x_k) + F(y_k) \ni 0, \quad \forall k \geq 0, \quad (1.8)$$

where:

- B_k is an approximation of the Jacobian $f'(x_k)$, updated iteratively using a modified Broyden update rule,
- $[x_{k-1}, x_k; g]$ denotes the first-order divided difference of the function g with respect to the points x_{k-1} and x_k ,
- y_k is the intermediate point computed at each iteration.

Due to the possibility of y_k being infeasible with respect to the constraint set C , we employ the conditional gradient method to perform an inexact projection, thereby obtaining the next iterate x_{k+1} that remains within C . This projection step is crucial for maintaining feasibility throughout the iterative process.

The local convergence of quasi-Newton approaches like (1.8) is established under specific assumptions, including bounded deterioration of the update matrix and Lipschitz continuity of the gradient. These conditions ensure that the algorithm remains stable and converges towards a solution at a linear or superlinear rate.

Contributions of the Thesis

The thesis makes several significant contributions to the field of numerical analysis and optimization:

- (i) **Development of Hybrid Secant-Type Methods:** We propose new versions of secant-type methods tailored for constrained generalized equations, integrating the conditional gradient method to handle feasibility constraints effectively.
- (ii) **Theoretical Convergence Analysis:** Utilizing the contraction mapping principle, we rigorously establish the convergence of the proposed algorithms. Under Lipschitz continuity of the gradient and metric regularity of the mapping, we demonstrate that the generated sequence is well-defined and converges locally with desirable rates.
- (iii) **Algorithmic Enhancements with Quasi-Newton Updates:** Inspired by Broyden's method, we incorporate quasi-Newton updates into the secant-type framework, improving the approximation of the Jacobian and enhancing convergence properties.

Literature Review

The study of generalized equations has its roots in the early 1970s, with pioneering works by S. M. Robinson [52, 53, 55, 56] and N. H. Josephy [41, 42]. These foundational studies laid the groundwork for understanding the theoretical underpinnings of generalized equations and their solution methodologies.

Over the decades, researchers have explored various iterative methods for solving generalized equations, focusing on improving convergence rates, handling constraints, and enhancing computational efficiency. Notable contributions include:

- **Newton-Type Methods:** Dontchev and Rockafellar [24] extensively studied Newton's method for generalized equations, providing comprehensive convergence analyses under metric regularity conditions.

- **Quasi-Newton Approaches:** Works by Adly and Huynh [2] and Andreani et al. [5] introduced quasi-Newton schemes that approximate the Jacobian matrix iteratively, reducing computational overhead while maintaining favorable convergence properties.
- **Constrained Optimization:** Studies by Dontchev [21, 23] and others have focused on the interplay between generalized equations and constrained optimization problems, emphasizing the role of metric regularity in ensuring solution stability and convergence.
- **Conditional Gradient Methods:** The integration of conditional gradient (Frank-Wolfe) methods into the framework of generalized equations, as seen in Chapter 3, has opened new avenues for handling feasibility constraints efficiently.

These contributions collectively underscore the importance of generalized equations in modeling and solving a wide array of problems in optimization, equilibrium analysis, and beyond. Our work builds upon these foundations, introducing novel hybrid methods that enhance both theoretical robustness and practical applicability.

Motivation for the Proposed Method

The generalized equation framework is inherently versatile, capturing a multitude of problem types within a unified mathematical structure. However, the presence of constraints and the set-valued nature of the mappings involved introduce significant challenges in devising efficient and robust solution methods.

Traditional Newton-type methods, while powerful, require the explicit computation of derivatives and may face difficulties in handling constraints and ensuring feasibility. Quasi-Newton methods alleviate some of these challenges by iteratively approximating the Jacobian matrix, reducing computational burden. However, they still necessitate strategies to maintain feasibility with respect to constraints.

The conditional gradient method offers an efficient means to perform projections onto feasible sets, especially when exact projections are computationally intensive or impractical. By integrating this method into a secant-type iterative scheme, we aim to create an algorithm that not only approximates the Jacobian effectively but also ensures that all iterates remain within the feasible region.

Moreover, the use of the contraction mapping principle provides a robust theoretical foundation for establishing convergence, ensuring that the sequence of iterates converges reliably to a solution under mild regularity conditions.

Conclusion

The convergence and efficiency of iterative methods for solving constrained generalized equations are pivotal for addressing a broad spectrum of applications in optimization and equilibrium modeling. By developing a hybrid secant-type method augmented with quasi-Newton updates and incorporating the conditional gradient method for feasible projections, this thesis contributes to the advancement of numerical techniques in this domain. The theoretical convergence analyses, grounded in metric regularity and Lipschitz continuity assumptions, provide confidence in the reliability of the proposed algorithms. Coupled with numerical validations, the work presented herein offers both theoretical insights and practical tools for tackling complex constrained generalized equations.

Chapter 2

Basic results

In this chapter, we review some notations, definitions and preliminary results used throughout this text.

2.1 First notations

Let $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^n . The sets

$$B_\delta(x) := \{y \in \mathbb{R}^n : \|x - y\| < \delta\}$$

and

$$B_\delta[x] := \{y \in \mathbb{R}^n : \|x - y\| \leq \delta\}$$

represent the *open* and *closed balls* centered at the point x with radius $\delta > 0$, respectively. We denote by $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ the *space of all continuous linear mappings* $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$. The *norm of the mapping* A is defined as

$$\|A\| := \sup\{\|Ax\| : \|x\| \leq 1\}.$$

Let $\Omega \subseteq \mathbb{R}^n$ be an open set. The *derivative of a differentiable function* $h : \Omega \rightarrow \mathbb{R}^n$ at a point x is characterized as the continuous linear mapping $h'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

2.2 First and second order divided differences

In the subsequent discussion, we will revisit the concept of *divided differences*, which serves as a fundamental tool for analyzing the incremental variations of a function across points within its domain. Consider a function $g : \Omega \rightarrow \mathbb{R}^n$ defined on a subset $\Omega \subseteq \mathbb{R}^n$. Divided

differences provide a systematic method for approximating derivatives and yield insights into the local behavior of the function. The *first-order divided difference* of g at points x and y in Ω is defined as an operator that linearly maps the vector difference $y - x$ to the corresponding difference $g(y) - g(x)$. When g is differentiable at a point x , this first-order divided difference coincides with the derivative $g'(x)$. Extending this concept, the *second-order divided difference*, which involves three points x, y , and z in Ω , quantifies the variation in the first-order divided differences and is analogous to the second derivative when g is twice differentiable. The formal definition is as follows:

Definition 2.2.1 Let $g : \Omega \rightarrow \mathbb{R}^n$ be a function and $x, y \in \Omega \subseteq \mathbb{R}^n$. An operator $[x, y; g]$ belonging to $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is a **first order divided difference of the function g at the points $x, y \in \Omega$** , if

$$[x, y; g](y - x) = g(y) - g(x), \quad \forall x, y \in \Omega, \quad x \neq y,$$

and, when g differentiable at x ,

$$[x, x; g] = g'(x).$$

An operator $[x, y, z; g]$ belonging to $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))$ is a **second order divided difference of g at the points $x, y, z \in \Omega$** , if

$$[x, y, z; g](z - x) = [y, z; g] - [x, y; g], \quad \forall x, y, z \in \Omega, \quad x \neq y, \quad x \neq z, \quad y \neq z,$$

and, if g is twice differentiable at x ,

$$[x, x, x; g] = \frac{1}{2}g''(x).$$

Below are three examples that illustrate the application of first and second-order divided differences for different types of functions.

Example 2.2.2 (Linear Function) Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear function defined by $g(x) = Ax$, where A is a constant $n \times n$ matrix.

The first-order divided difference, for any distinct points $x, y \in \mathbb{R}^n$ is given by

$$[x, y; g] = A.$$

This is because

$$[x, y; g](y - x) = A(y - x) = g(y) - g(x).$$

In conclusion, since g is linear, the divided difference operator is constant and equal to the matrix A . The second-order divided difference, for any distinct points $x, y, z \in \mathbb{R}^n$, is given by

$$[x, y, z; g] = 0.$$

This is because the first-order divided differences are all equal to A , so their difference is zero.

Example 2.2.3 (Quadratic Function in \mathbb{R}^n) Let $g(x) = \frac{1}{2}x^\top Qx + c^\top x + d$, where Q is a symmetric $n \times n$ matrix, $c \in \mathbb{R}^n$, and $d \in \mathbb{R}$.

We again use the definition of the first-order divided difference:

$$[x, y; g](y - x) = g(y) - g(x), \quad \text{for } x \neq y.$$

Step 1: Compute $g(y) - g(x)$.

Compute $g(y)$ and $g(x)$:

$$g(y) = \frac{1}{2}y^\top Qy + c^\top y + d,$$

$$g(x) = \frac{1}{2}x^\top Qx + c^\top x + d.$$

Subtract $g(x)$ from $g(y)$:

$$g(y) - g(x) = \left(\frac{1}{2}y^\top Qy - \frac{1}{2}x^\top Qx \right) + (c^\top y - c^\top x).$$

Step 2: Simplify $g(y) - g(x)$.

Using the identity for symmetric Q :

$$y^\top Qy - x^\top Qx = (y - x)^\top Q(y + x),$$

we rewrite the difference as:

$$g(y) - g(x) = \frac{1}{2}(y - x)^\top Q(y + x) + c^\top (y - x).$$

Step 3: Relate to the Divided Difference.

The first-order divided difference satisfies:

$$[x, y; g](y - x) = g(y) - g(x).$$

Substitute $g(y) - g(x)$ from Step 2:

$$[x, y; g](y - x) = \frac{1}{2}(y - x)^\top Q(y + x) + c^\top (y - x).$$

Step 4: Solve for $[x, y; g]$.

To isolate $[x, y; g]$, factor out $(y - x)$ from the right-hand side:

$$[x, y; g](y - x) = \left(\frac{1}{2}Q(y + x) + c \right)^\top (y - x).$$

Since Q is symmetric, $Q^\top = Q$, and we can write:

$$[x, y; g](y - x) = \left(\frac{1}{2}Q(y + x) + c \right)^\top (y - x).$$

Thus, the first-order divided difference is:

$$[x, y; g] = \frac{1}{2}Q(y + x) + c.$$

The second-order divided difference, for distinct points $x, y, z \in \mathbb{R}^n$, is given by

$$[x, y, z; g] = \frac{Q}{2}.$$

This is because $[x, y, z; g](z - x) = [y, z; g] - [x, y; g]$.

$$\text{So, } [x, y, z; g](z - x) = [y, z; g] - [x, y; g] = (Q \frac{z+y}{2} + c) - (Q \frac{y+x}{2} + c) = \frac{Q}{2}(z - x).$$

Example 2.2.4 Let $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a function given by $g(x_1, x_2) = (g_1(x_1, x_2), g_2(x_1, x_2))$, where

$$g_1(x_1, x_2) = x_1^2 - x_2 + 1 + \frac{1}{9}|x_1 - 1|,$$

$$g_2(x_1, x_2) = x_2^2 + x_1 - 7 + \frac{1}{9}|x_2|.$$

For $x, y \in \mathbb{R}^2$, $x = (x_1, x_2)$, $y = (y_1, y_2)$, $x_i \neq y_i$, $i = 1, 2$. We have:

$[x, y; g] \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ as:

$$[x, y; g]_{i1} = \frac{g_i(x_1, y_2) - g_i(y_1, y_2)}{x_1 - y_1}, i = 1, 2,$$

$$[x, y; g]_{i2} = \frac{g_i(x_1, x_2) - g_i(x_1, y_2)}{x_2 - y_2}, i = 1, 2.$$

$$[x, y; g]_{11} = \frac{g_1(x_1, y_2) - g_1(y_1, y_2)}{x_1 - y_1} = \frac{x_1^2 - y_1^2 + \frac{1}{9}(|x_1 - 1| - |y_1 - 1|)}{x_1 - y_1}.$$

$$[x, y; g]_{21} = \frac{g_2(x_1, y_2) - g_2(y_1, y_2)}{x_1 - y_1} = \frac{x_1 - y_1}{x_1 - y_1} = 1.$$

$$[x, y; g]_{12} = \frac{g_1(x_1, x_2) - g_1(x_1, y_2)}{x_2 - y_2} = -\frac{x_2 - y_2}{x_2 - y_2} = -1.$$

$$[x, y; g]_{22} = \frac{g_2(x_1, x_2) - g_2(x_1, y_2)}{x_2 - y_2} = \frac{x_2^2 - y_2^2 + \frac{1}{9}(|x_2| - |y_2|)}{x_2 - y_2}.$$

So,

$$[x, y; g] = \begin{bmatrix} \frac{x_1^2 - y_1^2 + \frac{1}{9}(|x_1 - 1| - |y_1 - 1|)}{x_1 - y_1} & -1 \\ 1 & \frac{x_2^2 - y_2^2 + \frac{1}{9}(|x_2| - |y_2|)}{x_2 - y_2} \end{bmatrix}$$

Note that

$$[x, y; g](y-x) = \begin{bmatrix} \frac{x_1^2 - y_1^2 + \frac{1}{9}(|x_1 - 1| - |y_1 - 1|)}{x_1 - y_1} & -1 \\ 1 & \frac{x_2^2 - y_2^2 + \frac{1}{9}(|x_2| - |y_2|)}{x_2 - y_2} \end{bmatrix} \begin{bmatrix} y_1 - x_1 \\ y_2 - x_2 \end{bmatrix} =$$

$$\begin{bmatrix} -x_1^2 + y_1^2 - \frac{1}{9}(|x_1 - 1| - |y_1 - 1|) - y_2 + x_2 \\ y_1 - x_1 - x_2^2 + y_2^2 - \frac{1}{9}(|x_2| - |y_2|) \end{bmatrix} = [g(y) - g(x)]^\top.$$

2.3 Distance and Excess

The following definitions related to sets and distances are fundamental tools that will be used throughout this report to analyze the convergence properties of the algorithm and its results. They provide a precise language for quantifying the relationship between points and sets such as relationship between two sets, which is essential for understanding the conditions under which our iterative method approaches a solution.

Definition 2.3.1 For sets C and D in \mathbb{R}^n the **distance from x to D** and the **excess of C beyond D** are defined, respectively, by

$$d(x, D) := \inf_{\hat{x} \in D} \|x - \hat{x}\| \quad \text{and} \quad e(C, D) := \sup_{x \in C} d(x, D), \quad (2.1)$$

where $d(x, \emptyset) = +\infty$, $e(\emptyset, D) = 0$ when $D \neq \emptyset$ and $e(\emptyset, \emptyset) = +\infty$.

Below are some examples that illustrate the definition of excess of a set beyond another one.

Example 2.3.2 Points in \mathbb{R}^2

Let $C = \{(2, 1)\}$ and $D = \{(0, 0)\}$.

$$d((2, 1), D) = \sqrt{5}, \quad e(C, D) = \sqrt{5}.$$

The distance is calculated by applying the Euclidean distance formula from $(2, 1)$ to $(0, 0)$, and since C is a single point, the excess is the same as the distance.

Example 2.3.3 Line Segment in \mathbb{R}^2

Let $C = [1, 2] \times \{0\}$ and $D = \{(0, 0)\}$.

$$d((x, 0), D) = x, \quad e(C, D) = 2.$$

The distance from $(x, 0)$ to $(0, 0)$ is x , and the supremum over the interval $[1, 2]$ is 2.

Example 2.3.4 Sphere and Point in \mathbb{R}^3

Let $C = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$ and $D = \{(0, 0, 0)\}$.

$$d(x, D) = 1, \quad e(C, D) = 1.$$

Every point on the unit sphere is at distance 1 from the origin.

Example 2.3.5 Cube and a Plane in \mathbb{R}^3 Let $C = [1, 2] \times [1, 2] \times [1, 2]$ and $D = \{(x, y, 0)\}$.

$$d((x, y, z), D) = |z|, \quad e(C, D) = 2.$$

The distance from (x, y, z) to the xy -plane is $|z|$, and the supremum of z over $[1, 2]$ is 2.

Example 2.3.6 A Point and a Line in \mathbb{R}^2

Let $C = \{(1, 2)\}$ and $D = \{(x, 0) \mid x \in \mathbb{R}\}$.

$$\begin{aligned} d((1, 2), D) &= \inf_{\hat{x} \in D} \|(1, 2) - \hat{x}\| = \|(1, 2) - (1, 0)\| = |2 - 0| = 2, \\ e(C, D) &= \sup_{x \in C} d(x, D) = d((1, 2), D) = 2. \end{aligned}$$

The point $(1, 2)$ is located directly 2 units above the line $D = \{(x, 0) \mid x \in \mathbb{R}\}$. Since the shortest distance from a point to a line is the perpendicular distance, the nearest point on D is $(1, 0)$. Hence, $d((1, 2), D) = 2$. Since C contains only this one point, the excess is the same as the distance, i.e., $e(C, D) = 2$.

Example 2.3.7 A Ball and a Point in \mathbb{R}^3

Let $C = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1\}$ and $D = \{(2, 0, 0)\}$.

$$\begin{aligned} d((-1, 0, 0), D) &= \|(-1, 0, 0) - (2, 0, 0)\| = \|(-3, 0, 0)\| = 3, \\ e(C, D) &= \sup_{x \in C} d(x, D) = 3. \end{aligned}$$

The set C is the unit ball centered at the origin, and D is the point $(2, 0, 0)$. The farthest point on the surface of the ball from D is $(-1, 0, 0)$, which is directly opposite D . The distance from $(-1, 0, 0)$ to $(2, 0, 0)$ is 3, hence $d((-1, 0, 0), D) = 3$. The excess is the maximum distance from any point in C to D , which is also 3.

Example 2.3.8 A Plane and a Line in \mathbb{R}^3

Let $C = \{(x, y, 0) \mid x^2 + y^2 \leq 1\}$ and $D = \{(0, 0, z) \mid z \in \mathbb{R}\}$.

$$\begin{aligned}d((x, y, 0), D) &= \sqrt{x^2 + y^2}, \\e(C, D) &= 1.\end{aligned}$$

The set C is the unit disk in the xy -plane, and D is the z -axis. The distance from any point $(x, y, 0) \in C$ to the z -axis is simply the distance in the xy -plane, which is $\sqrt{x^2 + y^2}$. The farthest point in C from D occurs at the boundary of the disk, where $x^2 + y^2 = 1$. Hence, $e(C, D) = 1$.

Example 2.3.9 A Cube and a Point in \mathbb{R}^4

Let $C = [1, 2] \times [1, 2] \times [1, 2] \times [1, 2]$ and $D = \{(0, 0, 0, 0)\}$.

$$\begin{aligned}d((x_1, x_2, x_3, x_4), D) &= \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}, \\e(C, D) &= 4.\end{aligned}$$

The set C is a 4-dimensional cube, and D is the origin. The distance from any point $(x_1, x_2, x_3, x_4) \in C$ to the origin is the Euclidean distance $\sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$. The maximum distance occurs when $x_1 = x_2 = x_3 = x_4 = 2$, giving $d((2, 2, 2, 2), D) = \sqrt{16} = 4$. Thus, the excess is 4.

Example 2.3.10 In this example we show, through the figure 2.1 the excess of C beyond D and the excess of D beyond C .

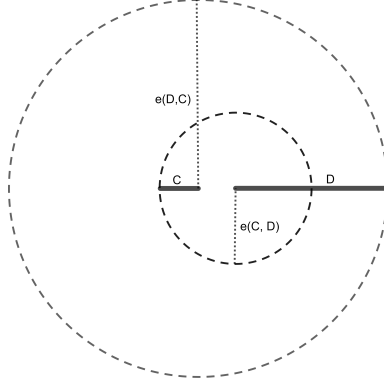


Figure 2.1: *Excess of C beyond D and excess of D beyond C*

2.4 Metric Regularity

In the following, we present the concept of metric regularity which plays an important role for understanding the behavior of set-valued mappings near a given point. More specifically, it describes how the inverse image of a point under a set-valued mapping behaves locally in terms of distance. By ensuring that the distance from points to their preimages (under the mapping) behaves in a controlled manner, metric regularity provides useful insights into the existence of solutions in optimization problems. It will be provided an in-depth exploration of metrically regular set-valued mappings, beginning with its formal definition.

Definition 2.4.1 *Let $\Omega \subset \mathbb{R}^n$ be open and nonempty. A set-valued mapping $G : \Omega \rightrightarrows \mathbb{R}^m$ is said to be **metrically regular at $\bar{x} \in \Omega$ for $\bar{u} \in \mathbb{R}^m$** if the following conditions are satisfied:*

- (i) $\bar{u} \in G(\bar{x})$.
- (ii) *The graph of G is locally closed at (\bar{x}, \bar{u}) .*
- (iii) *There exist constants $\kappa \geq 0$, $a > 0$, and $b > 0$ such that $B_a(\bar{x}) \subset \Omega$ and*

$$d(x, G^{-1}(u)) \leq \kappa d(u, G(x)), \quad \forall (x, u) \in B_a(\bar{x}) \times B_b(\bar{u}),$$

where $d(\cdot, \cdot)$ denotes the Euclidean distance.

*In addition, if the mapping $B_b(\bar{u}) \ni u \mapsto G^{-1}(u) \cap B_a(\bar{x})$ is single-valued, then G is called **strongly metrically regular at $\bar{x} \in \Omega$ for $\bar{u} \in \mathbb{R}^m$** with associated constants $\kappa \geq 0$, $a > 0$, and $b > 0$.*

When the mapping $B_b[\bar{u}] \ni u \mapsto G^{-1}(u) \cap B_a[\bar{x}]$ in Definition 2.4.1 is single-valued, then for the sake of simplicity we hereafter adopt the notation $w = G^{-1}(u) \cap B_a[\bar{x}]$ instead of $\{w\} = G^{-1}(u) \cap B_a[\bar{x}]$.

To elucidate the concept of metric regularity, we present three distinct examples. Each example highlights different scenarios where metric regularity holds, providing a comprehensive understanding of its applications and implications.

Example 2.4.2 (Affine Mappings) Consider the affine mapping $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by

$$G(x) = Ax + b,$$

where $A \in \mathbb{R}^{m \times n}$ is a constant matrix and $b \in \mathbb{R}^m$.

(i) **Metric Regularity:** The mapping G is metrically regular at any point $\bar{x} \in \mathbb{R}^n$ for $\bar{u} \in \mathbb{R}^m$ if and only if the matrix A has full row rank, i.e., $\text{rank}(A) = m$. This condition ensures that the inverse mapping $G^{-1}(u) = \{x \in \mathbb{R}^n \mid Ax = u - b\}$ is well-defined and behaves continuously with respect to u .

(ii) **Constants Identification:** If A has full row rank, there exists a constant $\kappa > 0$ such that

$$d(x, G^{-1}(u)) \leq \kappa \|u - G(x)\|, \quad \forall x \in B_a(\bar{x}), u \in B_b(\bar{u}),$$

where a and b can be chosen based on the norms of A and its pseudoinverse.

(iii) **Strong Metric Regularity:** Since G is single-valued and linear with full row rank, it is also strongly metrically regular. The inverse mapping $G^{-1}(u)$ is unique within the neighborhood $B_a(\bar{x})$.

Proof. (a) **Metric Regularity iff A has full row rank**

Necessity: Suppose G is metrically regular at \bar{x} for \bar{u} . Assume, for contradiction, that A does not have full row rank, i.e., $\text{rank}(A) < m$. Then, there exists a non-zero vector $v \in \mathbb{R}^m$ such that $A^T v = 0$. Consider a sequence $u_k = \bar{u} + \frac{1}{k}v$ and $x_k \in G^{-1}(u_k)$. Since $Ax_k = u_k - b = \bar{u} - b + \frac{1}{k}v = A\bar{x} + \frac{1}{k}v$, it follows that $A(x_k - \bar{x}) = \frac{1}{k}v$. However, $A^T v = 0$ implies v is orthogonal to the range of A , which contradicts $A(x_k - \bar{x}) = \frac{1}{k}v \neq 0$ for all k . Thus, A must have full row rank.

Sufficiency: Assume A has full row rank. Then, for each $u \in \mathbb{R}^m$, the equation $Ax = u - b$ has either a unique solution (if $n = m$) or infinitely many solutions forming an affine subspace (if $n > m$). Given the full row rank, the mapping G is open onto its image, and the inverse mapping is Lipschitz continuous with constant $\kappa = \|A^\dagger\|$, where A^\dagger is the Moore-Penrose pseudoinverse of A . Therefore, metric regularity holds.

(b) Existence of Constants κ , a , and b

Since A has full row rank, the pseudoinverse A^\dagger exists and satisfies $AA^\dagger = I_m$. The Lipschitz constant κ can be taken as $\|A^\dagger\|$, which bounds the distance from any x to the inverse image $G^{-1}(u)$. Specifically, for any $x \in B_a(\bar{x})$ and $u \in B_b(\bar{u})$, we have:

$$d(x, G^{-1}(u)) = \|A^\dagger(u - b - Ax)\| \leq \|A^\dagger\| \cdot \|u - G(x)\| = \kappa\|u - G(x)\|.$$

Appropriate choices of a and b ensure that x remains within the domain where A maintains full row rank and the Lipschitz condition holds uniformly.

(c) Strong Metric Regularity

Given that G is single-valued (since it's linear and A has full row rank), and the inverse mapping G^{-1} is unique within $B_a(\bar{x})$, the mapping G is strongly metrically regular. The uniqueness of $G^{-1}(u)$ within the neighborhood satisfies the single-valuedness condition required for strong metric regularity. ■

Example 2.4.3 (Nonlinear Mappings with Invertible Jacobian) Consider a nonlinear, single-valued mapping $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by

$$G(x) = F(x),$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuously differentiable.

- (i) **Local Invertibility:** Suppose that F is locally invertible around a point $\bar{x} \in \mathbb{R}^n$, and the Jacobian matrix $F'(\bar{x})$ is invertible, i.e., $\det(F'(\bar{x})) \neq 0$. By the Implicit Function Theorem, G is locally invertible near $(\bar{x}, F(\bar{x}))$.
- (ii) **Metric Regularity:** Under these conditions, G is metrically regular at \bar{x} for $\bar{u} = F(\bar{x})$. The inverse mapping G^{-1} exists locally and is Lipschitz continuous.
- (iii) **Strong Metric Regularity:** Since G is single-valued and its Jacobian is invertible at \bar{x} , G is also strongly metrically regular at \bar{x} for \bar{u} .

Proof. **(a) Local Invertibility via Implicit Function Theorem**

Since F is continuously differentiable and its Jacobian $F'(\bar{x})$ is invertible, the Implicit Function Theorem guarantees that there exists a neighborhood U of \bar{x} and a neighborhood V of $\bar{u} = F(\bar{x})$ such that for every $u \in V$, there exists a unique $x \in U$ satisfying $G(x) = u$. This establishes the local invertibility of G around (\bar{x}, \bar{u}) .

(b) Metric Regularity

Given the local invertibility and the smoothness of F , the inverse mapping $G^{-1} : V \rightarrow U$ exists and is continuously differentiable. Therefore, G^{-1} is Lipschitz continuous on V . Specifically, there exists a constant $\kappa > 0$ such that for all $u_1, u_2 \in V$,

$$\|G^{-1}(u_1) - G^{-1}(u_2)\| \leq \kappa \|u_1 - u_2\|.$$

This Lipschitz continuity directly implies metric regularity of G at \bar{x} for \bar{u} , with constants κ , a , and b determined by the neighborhoods U and V .

(c) Strong Metric Regularity

Since G is single-valued and its inverse G^{-1} is unique within the neighborhood U , the mapping G satisfies the single-valuedness condition required for strong metric regularity. Therefore, G is strongly metrically regular at \bar{x} for \bar{u} . ■

Example 2.4.4 (Normal Cone Mapping) Let $C \subset \mathbb{R}^n$ be a nonempty, closed, and convex set. Define the normal cone mapping $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by

$$G(x) = N_C(x),$$

where $N_C(x)$ denotes the normal cone to C at x .

(i) **Metric Regularity:** The normal cone mapping G is metrically regular at a point $\bar{x} \in C$ for $\bar{u} \in N_C(\bar{x})$ if and only if C satisfies the *Clarke regularity* (or *linear regularity*) around \bar{x} . This typically holds if C is smooth (i.e., has a continuously differentiable boundary) around \bar{x} .

(ii) **Examples of Regular Sets:**

- If C is the non-negative orthant \mathbb{R}_+^n , then $G(x)$ is metrically regular at any boundary point where one or more components of x are zero.
- If C is a strictly convex set with a smooth boundary, metric regularity holds at every boundary point.

(iii) **Strong Metric Regularity:** When C is smooth and strictly convex around \bar{x} , the normal cone mapping G is strongly metrically regular at \bar{x} for $\bar{u} \in N_C(\bar{x})$. This implies the uniqueness of the projection onto C in a neighborhood of \bar{x} .

Proof. **(a) Metric Regularity and Clarke Regularity**

Necessity: Suppose G is metrically regular at \bar{x} for \bar{u} . If C were not Clarke regular around \bar{x} , there would exist directions in which the normal cone does not behave continuously, violating the distance estimate required for metric regularity. Hence, C must satisfy Clarke regularity for G to be metrically regular.

Sufficiency: Assume C is Clarke regular around \bar{x} . Clarke regularity ensures that the normal cone mapping behaves continuously and satisfies the distance estimate:

$$d(x, G^{-1}(u)) \leq \kappa d(u, G(x)),$$

for some constant $\kappa > 0$ and for all (x, u) in neighborhoods around (\bar{x}, \bar{u}) . This directly implies that G is metrically regular at \bar{x} for \bar{u} .

(b) Examples of Regular Sets

- **Non-negative Orthant \mathbb{R}_+^n :** At any boundary point \bar{x} where one or more components are zero, the normal cone $N_{\mathbb{R}_+^n}(\bar{x})$ consists of vectors with non-negative entries corresponding to the zero components of \bar{x} . The orthant is polyhedral and satisfies Clarke regularity at all points, including boundary points. Therefore, G is metrically regular at these points.
- **Strictly Convex Sets with Smooth Boundary:** For strictly convex sets with a continuously differentiable boundary, the normal cone at any boundary point is uniquely determined by the outward normal vector. The smoothness ensures that small perturbations in the point result in small perturbations in the normal vector, satisfying the distance estimate required for metric regularity.

(c) Strong Metric Regularity

When C is smooth and strictly convex around \bar{x} , the projection onto C is uniquely defined and continuously differentiable in a neighborhood of \bar{x} . This uniqueness ensures that the inverse mapping $G^{-1}(u)$ is single-valued for u near \bar{u} . Consequently, G satisfies the conditions for strong metric regularity at \bar{x} for \bar{u} . ■

Remark 2.4.5 These examples illustrate the versatility and applicability of metric regularity in various contexts:

- **Affine - Linear Mappings:** Demonstrates that metric regularity is closely tied to the properties of the underlying linear operator, specifically its rank. Full row rank ensures that the mapping behaves well in terms of invertibility and stability.
- **Nonlinear Mappings with Invertible Jacobian:** Extends the concept to nonlinear settings, where local properties of the Jacobian play a crucial role in determining metric regularity. This is foundational in the analysis of Newton-type methods for solving nonlinear equations.

- **Normal Cone Mapping:** Highlights the role of geometry and regularity of convex sets in establishing metric regularity. This is particularly relevant in optimization and variational inequality problems, where normal cones characterize optimality conditions.

Understanding these examples provides deeper insight into how metric regularity facilitates the convergence analysis of iterative methods for solving generalized equations, ensuring robustness and reliability in diverse problem settings.

2.4.1 Linearization Error and Partial Linearization

The next definitions involves linearization error and partial linearization. Linearization error of a function $f + g$ at a point y in the domain Ω provides a way of measuring how accurately the linear approximation of $f + g$ at y represents the actual function in the neighborhood of x and z . Specifically, the linearization error captures the difference between the function $f + g$ and its first-order approximation using the derivative $f'(x)$ and an approximation of the derivative of the non-differentiable function g . The second concept, partial linearization, is a generalized form of linearization that is particularly useful when dealing with sets of functions or when we need to approximate a function with additional terms. Specifically, it applies to the sum of a differentiable function f , a non-differentiable function g , and a set-valued mapping F . In this case, the partial linearization of $f + g + F$ at y involves approximating the function $f + g + F$ by a linear map plus an additional term involving F .

Definition 2.4.6 Let $f : \Omega \rightarrow \mathbb{R}^n$ be a differentiable function and $g : \Omega \rightarrow \mathbb{R}^n$ a not necessarily differentiable function. The **linearization error of $f + g$** at $x, z \in \Omega$ is the mapping $E_{f+g}(x, \cdot, z) : \Omega \rightarrow \mathbb{R}^n$ defined by

$$E_{f+g}(x, y, z) := f(y) + g(y) - [f(x) + g(x) + (f'(x) + [z, x; g]) (y - x)]. \quad (2.2)$$

Definition 2.4.7 The **partial linearization of $f + g + F$** at $x, z \in \Omega$ is the set-valued mapping $L_{f+g+F}(x, \cdot, z) : \Omega \rightrightarrows \mathbb{R}^n$ defined by

$$L_{f+g+F}(x, y, z) := f(x) + g(x) + (f'(x) + [z, x; g]) (y - x) + F(y). \quad (2.3)$$

2.4.2 Graph, Domain, Range and inverse of a set-valued function

In the study of optimization, variational analysis, and nonsmooth analysis, set-valued mappings serve as fundamental objects that generalize classical functions. These mappings, which associate each point in the domain with a set of possible values in the range, allow us to model situations where solutions are not unique or where multi-valued relationships

between variables exist. To better understand and work with such mappings, we need to define some basic concepts, such as the **graph**, **domain**, **range**, and **inverse** of a set-valued mapping.

Definition 2.4.8 The **graph** of the set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is defined as

$$\text{gph } F := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : y \in F(x)\}.$$

The **domain** and the **range** of F are, respectively, the sets

$$\text{dom } F = \{x \in \mathbb{R}^n : F(x) \neq \emptyset\} \text{ and } \text{rge } F = \{y \in \mathbb{R}^n : y \in F(x) \text{ for some } x \in \mathbb{R}^n\}.$$

The **inverse** of F is the set-valued mapping $F^{-1} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined as

$$F^{-1}(y) = \{x \in \mathbb{R}^n : y \in F(x)\}.$$

2.4.3 Monotone maximal mapping

Monotonicity is a fundamental property that arises naturally in many practical contexts, particularly in problems where the relationship between variables should preserve a consistent directional effect. For example, in optimization or equilibrium models, we often require that increasing or decreasing one variable should lead to a predictable and consistent change in the corresponding solution set, thus maintaining stability and coherence within the system. In this report, we will explore how monotonicity and related properties, such as maximal monotonicity, influence the behavior of partial linearization and contribute to achieving strong metrical regularity. These concepts will be formally examined in the next proposition.

Definition 2.4.9 A mapping $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be **monotone** if

$$\langle y' - y, x' - x \rangle \geq 0, \quad \forall (x, y), (x', y') \in \text{gph } G.$$

If G is monotone and its graph is maximal with respect to this property, i.e., it is not properly contained in the graph of any other monotone operator, then we say that G is **maximal monotone**.

Example 2.4.10 Identity Operator in \mathbb{R}^n

Let $G(x) = x$. For any $(x, y), (x', y') \in \text{gph } G$, we have $y = x$ and $y' = x'$. The monotonicity condition becomes:

$$\langle x' - x, x' - x \rangle = \|x' - x\|^2 \geq 0.$$

Thus, $G(x) = x$ is monotone. Any attempt to extend the graph by adding a new point would break monotonicity, hence the graph is maximal.

Example 2.4.11 Zero Operator in \mathbb{R}^n

Let $G(x) = 0$. For any $(x, y), (x', y') \in \text{gph } G$, we have $y = 0$ and $y' = 0$. The monotonicity condition becomes:

$$\langle 0 - 0, x' - x \rangle = 0 \geq 0.$$

Thus, $G(x) = 0$ is monotone. Extending the graph by adding a point with $\tilde{y} \neq 0$ would violate monotonicity, so the graph is maximal.

Example 2.4.12 Subdifferential of a Convex Function in \mathbb{R}^n

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function and let $G(x) = \partial f(x)$, where $\partial f(x)$ is the subdifferential. The monotonicity condition holds as follows: For $(x, y) \in \text{gph } G$, we have $y \in \partial f(x)$, which means

$$f(x') \geq f(x) + \langle y, x' - x \rangle.$$

Similarly, for (x', y') , we have $y' \in \partial f(x')$, which means

$$f(x) \geq f(x') + \langle y', x - x' \rangle.$$

Adding these two inequalities gives:

$$0 \leq \langle y' - y, x' - x \rangle.$$

Hence, $G(x) = \partial f(x)$ is monotone. Extending the graph would violate this property, making the subdifferential maximal monotone.

When the mapping $B_b[\bar{u}] \ni u \mapsto G^{-1}(u) \cap B_a[\bar{x}]$ in Definition 2.4.1 is single-valued, then for the sake of simplicity we hereafter adopt the notation $w = G^{-1}(u) \cap B_a[\bar{x}]$ instead of $\{w\} = G^{-1}(u) \cap B_a[\bar{x}]$. Next, we present a class of mappings f, g and F for which the partial linearization of $f + g + F$ (see (2.3)) is strongly metrically regular.

Proposition 2.4.13 *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuously differentiable function, $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a continuous function and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximal monotone mapping. Assume that $z, x_* \in \mathbb{R}^n$, $\alpha > 0$ and $\beta > 0$ satisfy the following conditions:*

$$\langle f'(x_*)p, p \rangle \geq \beta \|p\|^2, \quad \langle [z, x_*; g]q, q \rangle \geq \alpha \|q\|^2, \quad \forall p, q \in \mathbb{R}^n. \quad (2.4)$$

Then, $\text{rge } L_{f+g+F}(x_, \cdot, z) = \mathbb{R}^n$, and for each $\bar{x} \in \mathbb{R}^n$ and $\bar{u} \in L_{f+g+F}(x_*, \bar{x}, z)$, the set-valued mapping $L_{f+g+F}(x_*, \cdot, z) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is strongly metrically regular at $\bar{x} \in \mathbb{R}^n$ for $\bar{u} \in \mathbb{R}^n$, with constants $\kappa = 1/(\alpha + \beta)$, $a = +\infty$, and $b = +\infty$.*

Proof. First, we will prove that $\text{rge } L_{f+g+F}(x_*, \cdot, z) = \mathbb{R}^n$. For that we take $\mu \in \mathbb{R}$ satisfying

$$0 < \mu < \frac{2\alpha + 2\beta}{(\|f'(x_*)\| + \|[z, x_*; g]\|)^2}. \quad (2.5)$$

Moreover, take $\hat{x} \in \mathbb{R}^n$ and define the mapping

$$\mathbb{R}^n \ni y \mapsto \Phi(y) := (I + \mu F)^{-1} (\mu \hat{x} + y - \mu [f(x_*) + g(x_*) + (f'(x_*) + [z, x_*; g]) (y - x_*)]). \quad (2.6)$$

Since F is a maximal monotone mapping, according to [24, Theorem 6C.4, p. 387] the mapping $(I + \mu F)^{-1}$ is single-valued and Lipschitz continuous on \mathbb{R}^n with constant 1. Thus, by using (2.6), for any $y, w \in \mathbb{R}^n$ we have

$$\|\Phi(y) - \Phi(w)\|^2 \leq \|y - w - \mu (f'(x_*) + [z, x_*; g]) (y - w)\|^2. \quad (2.7)$$

On the other hand, by using triangular inequality, some algebraic manipulations show that

$$\begin{aligned} \|y - w - \mu (f'(x_*) + [z, x_*; g]) (y - w)\|^2 &\leq \|y - w\|^2 - 2\mu \langle f'(x_*) (y - w), y - w \rangle - \\ &2\mu \langle [z, x_*; g] (y - w), y - w \rangle + \mu^2 (\|f'(x_*)\| + \|[z, x_*; g]\|)^2 \|y - w\|^2. \end{aligned}$$

Thus, using both inequalities (2.4) into the last inequality and combining the result with (2.7) we conclude that

$$\|\Phi(y) - \Phi(w)\| \leq (1 - 2\beta\mu - 2\alpha\mu + \mu^2 (\|f'(x_*)\| + \|[z, x_*; g]\|)^2)^{1/2} \|y - w\|.$$

Since μ satisfies (2.5) we conclude that $(1 - 2\beta\mu - 2\alpha\mu + \mu^2 (\|f'(x_*)\| + \|[z, x_*; g]\|)^2) < 1$. Therefore, by the Banach contraction principle (see [24, Theorem 1A.3, p. 17]), there exists a unique $x \in \mathbb{R}^n$ such that $x = \Phi(x)$, which together (2.6) implies that

$$(I + \mu F)x = \mu \hat{x} + x - \mu [f(x_*) + g(x_*) + (f'(x_*) + [z, x_*; g]) (x - x_*)].$$

Hence, expression (2.3) yields $\hat{x} = L_{f+g+F}(x_*, x, z)$. Therefore, $\text{rge } L_{f+g+F}(x_*, \cdot, z) = \mathbb{R}^n$.

We proceed to prove that the graph of $L_{f+g+F}(x_*, \cdot, z)$ is locally closed at (\bar{x}, \bar{u}) , i.e., there exists a neighborhood \mathcal{U} of (\bar{x}, \bar{u}) such that the intersection $\text{gph } L_{f+g+F}(x_*, \cdot, z) \cap \mathcal{U}$ is closed. Indeed, let $\{(\bar{x}_k, \bar{u}_k)\} \subset \text{gph } L_{f+g+F}(x_*, \cdot, z) \cap \mathcal{U}$ be a sequence such that $\lim_{k \rightarrow +\infty} \bar{x}_k = \bar{x}$ and $\lim_{k \rightarrow +\infty} \bar{u}_k = \bar{u}$. By the definition of graph of a set-valued mapping, we have $\bar{u}_k \in L_{f+g+F}(x_*, \bar{x}_k, z)$ for $k = 0, 1, \dots$. Hence, from (2.3) we obtain

$$\bar{u}_k \in f(x_*) + g(x_*) + (f'(x_*) + [z, x_*, z]) (\bar{x}_k - x_*) + F(\bar{x}_k), \quad k = 0, 1, \dots$$

Because F is maximal monotone, according to [8, Proposition 6.1.3, p. 185] it has closed graph, and thus we can take the limit in the last inclusion to conclude that

$$\bar{u} \in f(x_*) + g(x_*) + (f'(x_*) + [z, x_*, z]) (\bar{x} - x_*) + F(\bar{x}).$$

This implies that $\bar{u} \in L_{f+g+F}(x_*, \bar{x}, z)$, i.e., $(\bar{x}, \bar{u}) \in \text{gph } L_{f+g+F}(x_*, \cdot, z)$. Therefore, the graph of $L_{f+g+F}(x_*, \cdot, z)$ is locally closed at (\bar{x}, \bar{u}) .

Now, we will prove that the mapping $\mathbb{R}^n \ni x \mapsto L_{f+g+F}(x_*, x, z)$ is metrically regular at $\bar{x} \in \mathbb{R}^n$ for $\bar{u} \in \mathbb{R}^n$ with constants $\kappa = 1/(\alpha + \beta)$, $a = +\infty$, and $b = +\infty$. For this, take arbitrary $x, u \in \mathbb{R}^n$. As $\text{rge } L_{f+g+F}(x_*, \cdot, z) = \mathbb{R}^n$, there exists $y \in \mathbb{R}^n$ such that $u \in L_{f+g+F}(x_*, y, z)$. Thus, it follows from (2.3) that there exists $w_y \in F(y)$ such that

$$u = f(x_*) + g(x_*) + (f'(x_*) + [z, x_*; g]) (y - x_*) + w_y. \quad (2.8)$$

Moreover, for every arbitrary $v \in L_{f+g+F}(x_*, x, z)$, we can find $w_x \in F(x)$ such that

$$v = f(x_*) + g(x_*) + (f'(x_*) + [z, x_*; g]) (x - x_*) + w_x. \quad (2.9)$$

Thus, the monotonicity of F together with (2.8) and (2.9) yield

$$\begin{aligned} \langle (f'(x_*) + [z, x_*; g]) (x - y), x - y \rangle &\leq \langle (f'(x_*) + [z, x_*; g]) (x - y), x - y \rangle + \langle w_x - w_y, x - y \rangle \\ &= \langle v - u, x - y \rangle \\ &\leq \|v - u\| \|x - y\|. \end{aligned}$$

On the other hand, (2.4) implies in $\beta\|x - y\|^2 + \alpha\|x - y\|^2 \leq \langle (f'(x_*) + [z, x_*; g]) (x - y), x - y \rangle$, which combined with the last inequality gives

$$\beta\|x - y\|^2 + \alpha\|x - y\|^2 \leq \|v - u\| \|x - y\|. \quad (2.10)$$

We have two possibilities: $x = y$ and $x \neq y$. First, assume that $x = y$. In this case, as $u \in L_{f+g+F}(x_*, y, z)$ we have $x \in L_{f+g+F}(x_*, u, z)^{-1}$. Thus, we conclude that

$$d(x, L_{f+g+F}(x_*, u, z)^{-1}) = 0 \leq \frac{1}{\alpha + \beta} d(u, L_{f+g+F}(x_*, x, z)). \quad (2.11)$$

Now assume that $x \neq y$. In this case, it follows from (2.10) that

$$\|x - y\| \leq \frac{1}{\alpha + \beta} \|v - u\|, \quad \forall v \in L_{f+g+F}(x_*, x, z).$$

Because $u \in L_{f+g+F}(x_*, y, z)$, (2.1) and the previous inequality imply that

$$d(x, L_{f+g+F}(x_*, u, z)^{-1}) \leq \frac{1}{\alpha + \beta} d(u, L_{f+g+F}(x_*, x, z)), \quad \forall x, u \in \mathbb{R}^n. \quad (2.12)$$

Since we already know that $L_{f+g+F}(x_*, \cdot, z)$ is locally closed at (\bar{x}, \bar{u}) , by using (2.11) and (2.12), we conclude that $L_{f+g+F}(x_*, \cdot, z)$ is metrically regular at $\bar{x} \in \mathbb{R}^n$ for $\bar{u} \in \mathbb{R}^n$, with constants $\kappa = 1/(\alpha + \beta)$, $a = +\infty$, and $b = +\infty$.

To conclude the proof, it remains to prove that the mapping $u \mapsto L_{f+g+F}(x_*, u, z)^{-1}$ is single-valued from \mathbb{R}^n to \mathbb{R}^n . Take $u \in \mathbb{R}^n$, $x_1 \in L_{f+g+F}(x_*, u, z)^{-1}$ and $x_2 \in L_{f+g+F}(x_*, u, z)^{-1}$. For $i = 1, 2$, take $v_i \in F(x_i)$ such that $u = f(x_*) + g(x_*) +$

$(f'(x_*) + [z, x_*; g])(x_i - x_*) + v_i$. Thus, using (2.4) and taking into account that F is monotone we conclude that

$$(\alpha + \beta) \|x_1 - x_2\|^2 \leq \langle (f'(x_*) + [z, x_*; g])(x_1 - x_2), x_1 - x_2 \rangle + \langle v_1 - v_2, x_1 - x_2 \rangle = 0,$$

yielding that $x_1 = x_2$. Therefore, $L_{f+g+F}(x_*, \cdot, z)^{-1}$ is single-valued, concluding the proof. ■

2.4.4 Aubin Property

Definition 2.4.14 *A mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is said to have the Aubin property at $\bar{y} \in \mathbb{R}^m$ for $\bar{x} \in \mathbb{R}^n$ if $\bar{x} \in S(\bar{y})$, the graph of S is locally closed at (\bar{y}, \bar{x}) , and there is a constant $\kappa \geq 0$ together with neighborhoods U of \bar{x} and V of \bar{y} such that*

$$e(S(y') \cap U, S(y)) \leq \kappa \|y' - y\|, \quad \forall y', y \in V. \quad (2.13)$$

Proposition 2.4.15 *(alternative description of Aubin property). A mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ has the Aubin property at \bar{y} for \bar{x} with constant $\kappa > 0$ if and only if $\bar{x} \in S(\bar{y})$, $\text{gph} S$ is locally closed at (\bar{y}, \bar{x}) , and there exist neighborhoods U of \bar{x} and V of \bar{y} such that*

$$e(S(y') \cap U; S(y)) \leq \kappa \|y' - y\|, \quad \forall y' \in \mathbb{R}^m, \text{ and } y \in V. \quad (2.14)$$

Proof. Clearly, (2.14) implies (2.13). Assume (2.13) with corresponding U and V and choose positive a and b such that $B_a[\bar{x}] \subset U$ and $B_b[\bar{y}] \subset V$. Let $0 < a' < a$ and $0 < b' < b$ be such that

$$2\kappa b' + a' \leq \kappa b. \quad (2.15)$$

For any $y \in B_{b'}[\bar{y}]$ we have from (2.13) that

$$d(\bar{x}, S(y)) \leq \kappa \|y - \bar{y}\| \leq \kappa b',$$

hence

$$e(B_{a'}[\bar{x}], S(y)) \leq \kappa b' + a'. \quad (2.16)$$

The last inequality comes from $d(x, S(y)) \leq d(x, \bar{x}) + d(\bar{x}, S(y))$.

Take any $y' \in \mathbb{R}^m$. If $y' \in B_b[\bar{y}]$ the inequality in (2.14) comes from (2.13) and there is nothing more to prove. Assume $\|y' - y\| > b$. Then $\|y - y'\| = \|\bar{y} - y' + y - \bar{y}\| \geq \|\bar{y} - y'\| - \|y - \bar{y}\| > b - b'$. From (2.15),

$$2\kappa b' + a' \leq \kappa b \Rightarrow \kappa b' + a' \leq \kappa(b - b') \leq \kappa \|y' - y\|.$$

Using (2.16), we obtain

$$e(B_{a'}[\bar{x}], S(y)) \leq \kappa \|y' - y\|,$$

and since $S(y') \cap B_{a'}[\bar{x}] \subset B_{a'}[\bar{x}]$, we come again to (2.14). ■

Theorem 2.4.16 (equivalence of metric regularity and the inverse Aubin property). *A set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is metrically regular at \bar{x} for \bar{y} if and only if its inverse $F^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ has the Aubin property at \bar{y} for \bar{x} .*

Proof. The local closedness of the graph of F at $(\bar{x}; \bar{y})$ is equivalent to the local closedness of the graph of F^{-1} at $(\bar{y}; \bar{x})$. By assumption, F is metrically regular at \bar{x} for \bar{y} . Hence, $\bar{y} \in F(\bar{x})$, $\exists \kappa \geq 0, a > 0, b > 0$ ($U = B_a[\bar{x}], V = B_b[\bar{y}]$) such that

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)), \quad \forall (x, y) \in U \times V. \quad (2.17)$$

Without loss of generality, assume $b < \frac{a}{\kappa}$. We will prove next that

$$e(F^{-1}(y) \cap U, F^{-1}(y')) \leq \kappa \|y - y'\| \quad \forall y, y' \in V, \text{ with } U = B_a[\bar{x}], V = B_b[\bar{y}]. \quad (2.18)$$

Choose $y, y' \in B_b[\bar{y}]$. If $F^{-1}(y) \cap B_a[\bar{x}] = \emptyset$, then $d(\bar{x}; F^{-1}(y)) \geq a$. From the inequality (2.17) with $x = \bar{x}$ we have

$$a \leq d(\bar{x}, F^{-1}(y)) \leq \kappa d(y, F(\bar{x})) \leq \kappa \|y - \bar{y}\| \leq \kappa b < a,$$

a contradiction. Hence, there exists $x \in F^{-1}(y) \cap B_a[\bar{x}]$, and for any such x we have from (2.17) that

$$d(x, F^{-1}(y')) \leq \kappa d(y', F(x)) \leq \kappa \|y' - y\|.$$

Taking the supremum with respect to $x \in F^{-1}(y) \cap B_a[\bar{x}]$ we obtain (2.18).

Conversely, suppose there are neighborhoods U of \bar{x} and V of \bar{y} along with a constant $\kappa > 0$ such that (2.18) is satisfied. Take U and V smaller if necessary so that, according to Proposition 2.4.15, we have

$$e(F^{-1}(y) \cap U; F^{-1}(y')) \leq \kappa \|y - y'\|, \quad \forall y \in \mathbb{R}^m, \text{ and } y' \in V. \quad (2.19)$$

Let $x \in U$ and $y \in V$. If $F(x) \neq \emptyset$, then for any $y \in F(x)$ we have $x \in F^{-1}(y) \cap U$. From (2.19), we obtain

$$d(x, F^{-1}(y')) \leq e(F^{-1}(x) \cap U, F^{-1}(y')) \leq \kappa \|y - y'\|.$$

This holds for every $y \in F(x)$, hence, by taking the infimum with respect to $y \in F(x)$ in the last expression we get

$$d(x, F^{-1}(y')) \leq \kappa d(y', F(x)).$$

(If $F(x) = \emptyset$, then because of the convention $d(y; \emptyset) = \infty$, this inequality holds automatically.)

Hence, F is metrically regular at \bar{x} for \bar{y} with a constant κ . ■

Here, we provide several examples of applications that exhibit metric regularity.

Example 2.4.17 $S : \mathbb{R}^m \rightarrow \mathbb{R}^n$ linear continuous.

Take $\bar{y} \in \mathbb{R}^m$ and $\bar{x} \in \mathbb{R}^n$, with $\bar{x} = S(\bar{y})$. Let $V = \mathbb{R}^m$ be a neighborhood of \bar{y} and $U = \mathbb{R}^n$ a neighborhood of \bar{x} . Take $y', y \in V$. Then,

$$e(S(y') \cap U, S(y)) = e(S(y'), S(y)) = d(S(y'), S(y)) = \|S(y') - S(y)\| = \|S(y' - y)\| \leq \|S\| \|y' - y\|.$$

Hence, S is Aubin at \bar{y} for \bar{x} . So, S^{-1} is metrically regular at \bar{x} for \bar{y} .

Example 2.4.18 $S : \mathbb{R}^m \rightarrow \mathbb{R}^n$ locally Lipschitz.

Take $\bar{y} \in \mathbb{R}^m$ and $\bar{x} \in \mathbb{R}^n$, with $\bar{x} = S(\bar{y})$. Let V be a neighborhood of \bar{y} and U a neighborhood of \bar{x} , such that S is Lipschitz in V with constant L . Take $y', y \in V$. Then,

$$e(S(y') \cap U, S(y)) = e(S(y'), S(y)) = d(S(y'), S(y)) = \|S(y') - S(y)\| \leq L \|y' - y\|.$$

Hence, S is Aubin at \bar{y} for \bar{x} . So, S^{-1} is metrically regular at \bar{x} for \bar{y} .

Example 2.4.19 $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear invertible.

Take $\bar{x}, \bar{y} \in \mathbb{R}^n$, with $\bar{y} = S(\bar{x})$. Let V be a neighborhood of \bar{y} and U a neighborhood of \bar{x} . Take $(x, y) \in U \times V$.

$$d(x, S(y)) = \|x - S(y)\| = \|SS^{-1}(x) - S(y)\| = \|S(S^{-1}(x) - y)\| \leq \|S\| \|y - S^{-1}(x)\|.$$

So, S is metrically regular at \bar{x} for \bar{y} .

Example 2.4.20 $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ invertible, locally Lipschitz.

Take $\bar{x}, \bar{y} \in \mathbb{R}^n$, with $\bar{y} = S(\bar{x})$. Let V be a neighborhood of \bar{y} and U a neighborhood of \bar{x} . Take $(x, y) \in U \times V$.

$$d(x, S(y)) = \|x - S(y)\| = \|SS^{-1}(x) - S(y)\| \leq \kappa \|y - S^{-1}(x)\|.$$

So, S is metrically regular at \bar{x} for \bar{y} .

Example 2.4.21 $F^{-1} : \mathbb{R} \rightrightarrows \mathbb{R}$ defined by

$$F^{-1}(y) = \begin{cases} \{0, 1 + \sqrt{y}\}, & \text{se } y \geq 0 \\ 0, & \text{se } y < 0. \end{cases}$$

This mapping has the Aubin property at 0 for 0.

In fact, let U be a neighbourhood of 0, for instance $U = (-2, 2)$ and V be a neighbourhood of 0, for instance $V = (-\frac{1}{4}, \frac{1}{4})$.

Take $y, y' \in V$.

If $y, y' \geq 0$, we have:

$F^{-1}(y) = \{0, 1 + \sqrt{y}\}$ and $F^{-1}(y') = \{0, 1 + \sqrt{y'}\}$. Since $y' \in [0, \frac{1}{4})$, we have $F^{-1}(y') < \frac{3}{2} \Rightarrow F^{-1}(y') \cap U = F^{-1}(y')$.

So, $e(F^{-1}(y') \cap U, F^{-1}(y)) = \sup_{s \in S(y') \cap U} d(s, F^{-1}(y)) = \frac{3}{2}$.

Note that $\|y - y'\| < \frac{1}{4}$.

Hence, taking $\kappa = 6$, we have $e(F^{-1}(y') \cap U, F^{-1}(y)) = \frac{3}{2} \leq \kappa \cdot \|y - y'\|$.

If $y \in (-\frac{1}{4}, 0)$ and $y' \geq 0$, we have:

$F^{-1}(y) = \{0, 1 + \sqrt{y'}\}$ and $F^{-1}(y') = 0 \Rightarrow F^{-1}(y') \cap U = \{0\}$.

So, $e(F^{-1}(y') \cap U, F^{-1}(y)) = d(0, F^{-1}(y)) = \frac{3}{2}$.

Note that $\|y - y'\| < \frac{1}{2}$.

Hence, taking $\kappa = 3$, we have $e(F^{-1}(y') \cap U, F^{-1}(y)) = \frac{3}{2} \leq \kappa \cdot \|y - y'\|$.

If $y, y' \in (-\frac{1}{4}, 0)$, we have:

$F^{-1}(y) = 0$ and $F^{-1}(y') = 0 \Rightarrow F^{-1}(y') \cap U = \{0\}$.

So, $e(F^{-1}(y') \cap U, F^{-1}(y)) = 0$.

Note that $\|y - y'\| < \frac{1}{4}$.

Hence, taking any $\kappa > 0$, we have $e(F^{-1}(y') \cap U, F^{-1}(y)) = 0 \leq \kappa \cdot \|y - y'\|$.

Conclusion: We can take $\kappa = 6$ for all cases, and this proof that F^{-1} has the Aubin property at 0 for 0.

Then, $F : \mathbb{R} \rightrightarrows \mathbb{R}$ is metrically regular at 0 for 0, where,

$$F(x) = \begin{cases} \emptyset, & \text{se } x \leq 0 \\ \mathbb{R}, & \text{se } x = 0 \\ \emptyset, & \text{se } 0 < x < 1 \\ (x - 1)^2, & \text{se } x \geq 1. \end{cases}$$

2.4.5 The Contraction Principle Theorem

In the study of fixed-point theory, one of the most fundamental and widely applicable results is the *Contraction Principle* (also known as the Banach Fixed-Point Theorem), which guarantees the existence of a unique fixed point for certain types of mappings. This principle has found applications in various fields such as differential equations, optimization, and mathematical economics.

The theorem, in its classical form, addresses single-valued contractions in a complete metric space, ensuring the existence and uniqueness of fixed points. However, the set-valued version of the contraction principle is an important generalization that extends this result to set-valued mappings (multifunctions), where the mapping may associate each point with a set of values rather than a single value.

The *Contraction Principle* in this setting introduces conditions under which a set-valued mapping Φ on \mathbb{R}^n admits a fixed point, meaning there exists a point \hat{x} such that $\hat{x} \in \Phi(\hat{x})$. The theorem provides specific conditions involving the distance from \bar{x} to the set $\Phi(\bar{x})$ and a contraction property within a neighborhood around \bar{x} .

The theorem can be particularly useful in dealing with non-linear systems or in contexts where solutions are not guaranteed to be unique or where the solution space is more complex than a single point. This generalization of the fixed-point theorem offers powerful tools for handling these broader cases and has implications in fields like control theory and game theory, where set-valued mappings naturally arise.

The theorem is stated as follows:

Theorem 2.4.22 (Contraction Principle) *Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping and $\bar{x} \in \mathbb{R}^n$. Suppose that there exist scalars $r > 0$ and $\lambda \in (0, 1)$ such that the set $\text{gph } \Phi \cap (B_r[\bar{x}] \times B_r[\bar{x}])$ is closed and*

$$(a) \quad d(\bar{x}, \Phi(\bar{x})) \leq r(1 - \lambda);$$

$$(b) \quad e(\Phi(x) \cap B_r[\bar{x}], \Phi(x')) \leq \lambda \|x - x'\|, \text{ for all } x, x' \in B_r[\bar{x}].$$

Then Φ has a fixed point in $B_r[\bar{x}]$; that is, there exists $\hat{x} \in B_r[\bar{x}]$ such that $\hat{x} \in \Phi(\hat{x})$. If Φ is single-valued, then \hat{x} is the unique fixed point of Φ in $B_r[\bar{x}]$.

Proof. By assumption (a), there exists $x_1 \in \Phi(\bar{x})$ such that $\|\bar{x} - x_1\| \leq r(1 - \lambda)$. Proceeding by induction, suppose that there exists $x_{k+1} \in \Phi(x_k) \cap B_r[\bar{x}]$, for $k = 1, 2, \dots, n - 1$ with

$$\|x_{k+1} - x_k\| \leq r(1 - \lambda)\lambda^k.$$

By assumption (b),

$$d(x_n, \Phi(x_n)) \leq e(\Phi(x_{n-1}) \cap B_r[\bar{x}], \Phi(x_n)) \leq \lambda \|x_{n-1} - x_n\| < r(1 - \lambda)\lambda^n.$$

This implies that there exists $x_{n+1} \in \Phi(x_n)$ such that

$$\|x_{n+1} - x_n\| \leq r(1 - \lambda)\lambda^n.$$

By the triangle inequality,

$$\|x_{n+1} - \bar{x}\| \leq \left(\sum_{k=1}^n \|x_{k+1} - x_k\| \right) + \|x_1 - \bar{x}\| < r(1 - \lambda) \sum_{k=0}^n \lambda^k \leq r(1 - \lambda) \sum_{k=0}^{\infty} \lambda^k = r.$$

Hence, $x_{n+1} \in B_r[\bar{x}]$ and the induction step is complete.

For $n > m$, we have

$$\|x_n - x_m\| \leq \sum_{k=m}^{n-1} \|x_{k+1} - x_k\| \leq r(1 - \lambda) \sum_{k=m}^{n-1} \lambda^k \leq r(1 - \lambda) \sum_{k=m}^{\infty} \lambda^k = r\lambda^m.$$

Thus, the sequence $\{x_k\}$ is a Cauchy sequence converging to some $\hat{x} \in B_r[\bar{x}]$.

By assumption (b),

$$d(x_n, \Phi(\hat{x})) \leq e(\Phi(x_{n-1}) \cap B_r[\bar{x}], \Phi(\hat{x})) \leq \lambda \|\hat{x} - x_{n-1}\|.$$

The triangle inequality implies that

$$d(\hat{x}, \Phi(\hat{x})) \leq \|\hat{x} - x_n\| + d(x_n, \Phi(\hat{x})) \leq \|x_n - \hat{x}\| + \lambda \|x_{n-1} - \hat{x}\|,$$

which approaches zero as n increases. Since $\Phi(\hat{x})$ is closed, we conclude that $\hat{x} \in \Phi(\hat{x})$.

If Φ is single-valued, assumption (b) implies that \hat{x} is the unique fixed point of Φ in $B_r[\bar{x}]$.

This completes the proof. ■

2.4.6 Perturbed Metric Regularity Theorem

Theorem 2.4.23 (Perturbed metric regularity) *Consider a mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with closed graph and a point $(\bar{x}, \bar{y}) \in \text{gph}H$ at which H is metrically regular, that is, there exist positive constants a, b and κ such that*

$$d(x, H^{-1}(y)) \leq \kappa d(y, H(x)) \quad \text{for all } (x, y) \in B_a[\bar{x}] \times B_b[\bar{y}]. \quad (2.20)$$

Let $\mu > 0$ be such that $\kappa\mu < 1$ and $\kappa' > \kappa$. Then for every positive α and β such that

$$\alpha \leq a/2, \quad \mu\alpha + 2\beta \leq b \quad \text{and} \quad 2\kappa'\beta \leq \alpha(1 - \kappa\mu), \quad (2.21)$$

and for every function $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ satisfying

$$\|h(\bar{x})\| \leq \beta, \quad (2.22)$$

and

$$\|h(x) - h(x')\| \leq \mu \|x - x'\|, \quad \forall x, x' \in B_\alpha[\bar{x}], \quad (2.23)$$

the mapping $h + H$ has the following property: for every $y, y' \in B_\beta[\bar{y}]$ and every $x \in (h + H)^{-1}(y) \cap B_\alpha[\bar{x}]$ there exists $x' \in (h + H)^{-1}(y')$ such that

$$\|x - x'\| \leq \frac{\kappa'}{1 - \kappa\mu} \|y - y'\|. \quad (2.24)$$

In addition, if the mapping H is strongly metrically regular at \bar{x} for \bar{y} ; that is, the mapping $y \mapsto H^{-1}(y) \cap B_\alpha[\bar{x}]$ is single-valued and Lipschitz continuous on $B_b[\bar{y}]$ with a Lipschitz constant κ , then for μ, κ', α and β as above and any function h satisfying (2.22) and (2.23), the mapping $y \mapsto (H + h)^{-1}(y) \cap B_\alpha[\bar{x}]$ is a Lipschitz continuous function on $B_\beta[\bar{y}]$ with a Lipschitz constant $\frac{\kappa'}{1 - \kappa\mu}$.

Proof. Choose μ and κ' as required and then α and β to satisfy (2.21). For any $x \in B_\alpha[\bar{x}]$ and $y \in B_\beta[\bar{y}]$, using (2.22), (2.23) and the triangle inequality, we obtain

$$\| -h(x) + y - \bar{y} \| \leq \|h(\bar{x})\| + \|h(\bar{x}) - h(x)\| + \|y - \bar{y}\| \leq \beta + \mu \|x - \bar{x}\| + \beta \leq 2\beta + \mu\alpha \leq b, \quad (2.25)$$

where the last inequality follows from the second inequality in (2.21). Fix $y' \in B_\beta[\bar{y}]$ and consider the mapping

$$\Phi_{y'} : x \mapsto H^{-1}(-h(x) + y') \text{ for } x \in B_\alpha[\bar{x}].$$

Clearly, $\text{gph}\Phi_{y'}$ is closed. Let $y \in B_\beta[\bar{y}], y \neq y'$ and let $x \in (h + H)^{-1}(y) \cap B_\alpha[\bar{x}]$. We will apply Theorem 2.4.22 with the closed ball $B_\alpha[\bar{x}]$ to show that there is a fixed point $x' \in \Phi_{y'}(x')$ in the closed ball centered at x with radius

$$\epsilon := \frac{\kappa' \|y - y'\|}{1 - \kappa\mu}. \quad (2.26)$$

Then, from the third inequality in (2.21), we obtain

$$\epsilon \leq \frac{\kappa'(2\beta)}{1 - \kappa\mu} \leq \alpha.$$

Hence, from the first inequality in (2.21), we get $B_\epsilon[x] \subset B_\alpha[\bar{x}]$. Since $y \in h(x) + H(x)$ and (x, y) satisfies (2.25), from the assumed metric regularity of H we get

$$\begin{aligned} d(x, \Phi_{y'}(x)) &= d(x, H^{-1}(-h(x) + y')) \leq \kappa d(-h(x) + y', H(x)) \\ &= \kappa d(y', h(x) + H(x)) \leq \kappa \|y - y'\| \\ &< \kappa' \|y - y'\| = \epsilon(1 - \kappa\mu). \end{aligned}$$

For any $u, v \in B_\epsilon[x]$, using (2.23), we have

$$\begin{aligned} e(\Phi_{y'}(u) \cap B_\epsilon[x], \Phi_{y'}(v)) &\leq e(H^{-1}(-h(u) + y') \cap B_\alpha[\bar{x}], H^{-1}(-h(v) + y')) \\ &\leq \kappa \|h(u) - h(v)\| \leq \kappa \mu \|u - v\|. \end{aligned}$$

Applying Theorem 2.4.22 to the mapping $\Phi_{y'}$, with \bar{x} identified with x and constants $r = \epsilon$ and $\lambda = \kappa \mu$, we obtain the existence of a fixed point $x' \in \Phi_{y'}(x') = H^{-1}(-h(x') + y')$, which is equivalent to $x' \in (h + H)^{-1}(y')$, within distance ϵ given by (2.26) from x . So,

$$\|x - x'\| \leq \epsilon = \frac{\kappa'}{1 - \kappa \mu} \|y - y'\|.$$

For the second part of the theorem, suppose that $y \mapsto s(y) := H^{-1}(y) \cap B_\alpha[\bar{x}]$ is a Lipschitz continuous function on $B_b[\bar{y}]$ with a Lipschitz constant κ . Choose μ, κ', α and β as in the statement and let h satisfy (2.22) and (2.23).

For any $y \in B_\beta[\bar{y}]$, since $\bar{x} \in (h + H)^{-1}(\bar{y} + h(\bar{x})) \cap B_\alpha[\bar{x}]$, from (2.24) we obtain that there exists $x \in (h + H)^{-1}(y)$ such that

$$\|x - \bar{x}\| \leq \frac{\kappa'}{1 - \kappa \mu} \|y - (\bar{y} + h(\bar{x}))\|.$$

Since $\|y - (\bar{y} + h(\bar{x}))\| \leq \|y - \bar{y}\| + \|h(\bar{x})\| \leq 2\beta$, by (2.22) we get $\|x - \bar{x}\| \leq \alpha$, that is, $(h + H)^{-1}(y) \cap B_\alpha[\bar{x}] \neq \emptyset$. Hence, the domain of the mapping $(h + H)^{-1} \cap B_\alpha[\bar{x}]$ contains $B_\beta[\bar{y}]$.

If $x \in (h + H)^{-1}(y) \cap B_\alpha[\bar{x}]$, then $x \in H^{-1}(y - h(x)) \cap B_\alpha[\bar{x}] \subset H^{-1}(y - h(x)) \cap B_\alpha[\bar{x}] = s(y - h(x))$, since $y - h(x) \in B_b[\bar{y}]$ according to (2.25). Hence,

$$H^{-1}(y - h(x)) \cap B_\alpha[\bar{x}] = s(y - h(x)) = x. \quad (2.27)$$

Assume that there exist $y \in B_\beta$ and $x, x' \in (h + H)^{-1}(y) \cap B_\alpha(\bar{x})$ such that $x \neq x'$. From (2.25), we have that both $y - h(x)$ and $y - h(x')$ are in $B_b[\bar{y}]$. Then from (2.27) we get

$$\begin{aligned} \|x - x'\| &= \|s(-h(x') + y) - s(-h(x) + y)\| \\ &\leq \kappa \| -h(x) + y + h(x') - y \| = \kappa \|h(x) - h(x')\| \\ &\leq \kappa \mu \|x - x'\| < \|x - x'\|, \end{aligned}$$

which is a contradiction. Hence, the mapping $y \mapsto g(y) := (h + H)^{-1}(y) \cap B_\alpha[\bar{x}]$ is single-valued, that is, a function, defined on $B_\beta[\bar{y}]$. Let $y, y' \in B_\beta[\bar{y}]$. Utilizing the equality $g(y) = s(-h(g(y)) + y)$, see (2.27), we have

$$\begin{aligned} \|g(y) - g(y')\| &= \|s(-h(g(y)) + y) - s(-h(g(y')) + y')\| \\ &\leq \kappa \| -h(g(y)) + y + h(g(y')) - y' \| \leq \kappa \|h(g(y)) - h(g(y'))\| + \kappa \|y - y'\| \\ &\leq \kappa \mu \|g(y) - g(y')\| + \kappa \|y - y'\|. \end{aligned}$$

Thus

$$(1 - \kappa\mu)\|g(y) - g(y')\| \leq \kappa\|y - y'\| \Rightarrow \|g(y) - g(y')\| \leq \frac{\kappa'}{1 - \kappa\mu}\|y - y'\|;$$

that is, g is Lipschitz continuous with Lipschitz constant $\frac{\kappa'}{1 - \kappa\mu}$. The proof is complete. ■

Remark 2.4.24 Assuming $h(\bar{x}) = 0$ in the last theorem, the concept of perturbed metric regularity indicates that if H is (strongly) metrically regular at \bar{x} for \bar{y} , and if h has a sufficiently small Lipschitz constant, then the perturbed function $h + H$ is also (strongly) metrically regular at \bar{x} for \bar{y} . Specifically, the claim related to inequality (2.24) implies that $(h + H)^{-1}$ possesses the Aubin property at \bar{y} for \bar{x} which is equivalent to the metric regularity of $h + H$ at \bar{x} for \bar{y} . This leads us to the (extended) Lyusternik-Graves theorem as stated in [24, Theorem 5E.1]. For the strong regularity aspect, we derive a version of Robinson's Theorem, as referenced in [24, Theorem 5F.1].

2.5 The Feasible Inexact Projection Mapping

The proposed methods studies in the following chapters require an inner subroutine named *CondG procedure* to compute a feasible inexact projection onto the set C , which was also used in [45] (see also, [12, p. 215], e [18, 33]). Hence, we also need first describe this subroutine. For that, we first recall the concept of feasible inexact projection introduced in [18].

Definition 2.5.1 Let $C \subset \mathbb{R}^n$ be a closed and convex set, $x \in C$ and $\theta \geq 0$. The **feasible inexact projection mapping** relative to x and forcing term θ , denoted by $P_C(\cdot, x, \theta) : \mathbb{R}^n \rightarrow C$ is the set-valued mapping defined as

$$P_C(y, x, \theta) := \{w \in C : \langle y - w, z - w \rangle \leq \theta\|y - x\|^2, \quad \forall z \in C\}.$$

Each point $w \in P_C(y, x, \theta)$ is called a feasible inexact projection of y onto C with respect to x and forcing term θ .

Remark 2.5.2 Since $C \subset \mathbb{R}^n$ is a closed and convex set, Proposition 2.1.3 of [12, p. 201] implies that $P_C(y) \in P_C(y, x, \theta)$ for each $y \in \mathbb{R}^n$, where $P_C(\cdot)$ denotes the exact projection mapping. Therefore, $P_C(y, x, \theta) \neq \emptyset$ for all $y \in \mathbb{R}^n$ and $x \in C$. If in Definition 2.5.1 we have $\theta = 0$, then $P_C(y, x, 0) = \{P_C(y)\}$ for all $y \in \mathbb{R}^n$ and $x \in C$. In this case, we may write $P_C(y, x, 0) = P_C(y)$ instead of $P_C(y, x, 0) = \{P_C(y)\}$.

The next result plays an important role in our theory. It is a basic property of the feasible inexact projection whose proof can be found in [33, Lemma 4]; see also [18].

Lemma 2.5.3 *Let $y, \tilde{y} \in \mathbb{R}^n$, $x, \tilde{x} \in C$ and $\theta \geq 0$. Then, for any $u \in P_C(y, x, \theta)$ we have*

$$\|u - P_C(\tilde{y}, \tilde{x}, 0)\| \leq \|y - \tilde{y}\| + \sqrt{2\theta}\|y - x\|.$$

The CondG procedure, which is based on conditional gradient method (for a general overview of this method; see [12]), is the procedure for obtain a feasible inexact projections onto special compact sets C . The CondG procedure with input data v , u and θ it is formally described as follows:

Algorithm 2.5.4 CondG

Input: Take $\theta \in \mathbb{R}_+$, $v \in \mathbb{R}^n$ and $v \neq u, w \in C$. Set $w_0 = w$ and $\ell = 0$.

Step 1. Use a Linear Optimization oracle to compute an optimal solution z_ℓ and the optimal value s_ℓ^* as

$$z_\ell := \arg \min_{z \in C} \langle w_\ell - v, z - w_\ell \rangle, \quad s_\ell^* := \langle w_\ell - v, z_\ell - w_\ell \rangle. \quad (2.28)$$

Step 2. If $-s_\ell^* \leq \theta\|v - u\|^2$, then **stop** and returns $w^+ := w_\ell$.

Step 3. Compute $\alpha_\ell \in (0, 1]$ and $w_{\ell+1}$ as

$$w_{\ell+1} := w_\ell + \alpha_\ell(z_\ell - w_\ell), \quad \alpha_\ell := \min \left\{ 1, \frac{-s_\ell^*}{\|z_\ell - w_\ell\|^2} \right\}. \quad (2.29)$$

Step 4. Set $\ell \leftarrow \ell + 1$, and go to step 1.

Output: $w^+ := w_\ell$.

We now highlight the key aspects of algorithm above. The algorithm aims to approximate the projection of v onto the set C by iteratively improving the estimate w_ℓ . At each iteration, the algorithm solves the linearized subproblem (2.28), which involves minimizing the linear approximation of the distance function over C . The optimal value s_ℓ^* obtained in (2.28) measures the potential reduction in the objective function. If this reduction is smaller than the specified tolerance, i.e., $-s_\ell^* \leq \theta\|v - u\|^2$, the algorithm terminates, returning $w^+ = w_\ell$ as the approximate projection. If the condition is not met, the algorithm updates the current estimate by moving towards z_ℓ using a step size α_ℓ . The choice of α_ℓ ensures sufficient decrease in the objective function while keeping $w_{\ell+1}$ within C due to the convexity of C . Under mild assumptions, it can be shown that the sequence $(s_\ell^*)_{\ell \in \mathbb{N}}$ converges to zero (see, e.g., [10, Proposition A.2]), guaranteeing that the algorithm will terminate after a finite number of iterations. In essence, the algorithm above provides an efficient method for computing an approximate projection onto C that satisfies the inexact projection condition required by our main algorithm.

Next, we simulate the Cond G Algorithm. We aim to minimize a linear function over a convex feasible set, specifically the unit square in \mathbb{R}^2 . The goal is to demonstrate how the Conditional Gradient Algorithm can be applied in this context, following the steps of the algorithm to compute an optimal solution.

Example 2.5.5 We want to minimize a linear function $\langle w - v, z - w \rangle$ over a convex feasible set $C \subseteq \mathbb{R}^2$.

Problem Setup

- Let C be the unit square in \mathbb{R}^2 :

$$C = \{z = (z_1, z_2) \in \mathbb{R}^2 \mid 0 \leq z_1 \leq 1, 0 \leq z_2 \leq 1\}.$$

- Let the starting point $w = (0.5, 0.5)$.

- Let $v = (0, 0)$.

- Set $u = (1, 1)$.

- Choose a threshold $\theta = 0.01$.

Applying the Conditional Gradient Algorithm

Step 1: Linear Optimization Oracle (Initial Iteration, $\ell = 0$)

The algorithm solves the linear optimization problem:

$$z_0 = \arg \min_{z \in C} \langle w_0 - v, z - w_0 \rangle.$$

This means minimizing the function $\langle (0.5, 0.5) - (0, 0), (z_1, z_2) - (0.5, 0.5) \rangle$ over $z \in C$. Simplifying, this becomes:

$$\langle (0.5, 0.5), (z_1 - 0.5, z_2 - 0.5) \rangle = 0.5(z_1 - 0.5) + 0.5(z_2 - 0.5).$$

To minimize this function, we choose $z_0 = (0, 0)$, since it gives the smallest values for both $z_1 - 0.5$ and $z_2 - 0.5$ over the unit square.

So, the solution is:

$$z_0 = (0, 0).$$

The optimal value of the objective function is:

$$s_0^* = \langle (0.5, 0.5), (0, 0) - (0.5, 0.5) \rangle = -0.5(-0.5) - 0.5(-0.5) = -0.5.$$

Step 2: Stopping Condition

The algorithm checks if $-s_0^* \leq \theta \|v - u\|^2$.

- We know $s_0^* = -0.5$.

- The norm $\|v - u\|^2 = \|(1, 1) - (0, 0)\|^2 = 1^2 + 1^2 = 2$.

- We check if $0.5 \leq 0.01 \times 2$, i.e., if $0.5 \leq 0.02$, which is **false**.

Thus, the algorithm continues to the next iteration.

Step 3: Compute w_1 and α_0

Compute α_0 as:

$$\alpha_0 = \min \left\{ 1, \frac{-s_0^*}{\|z_0 - w_0\|^2} \right\}.$$

- $s_0^* = -0.5$.

- $\|z_0 - w_0\|^2 = \|(0, 0) - (0.5, 0.5)\|^2 = 0.5^2 + 0.5^2 = 0.5$.

- So, $\alpha_0 = \min \left\{ 1, \frac{0.5}{0.5} \right\} = 1$.

Update w_1 as:

$$w_1 = w_0 + \alpha_0(z_0 - w_0) = (0.5, 0.5) + 1 \times ((0, 0) - (0.5, 0.5)) = (0, 0).$$

Step 4: Iteration Update

Set $\ell = 1$ and return to Step 1 with $w_1 = (0, 0)$.

Step 1: Linear Optimization Oracle (Iteration $\ell = 1$)

Now, solve the linear optimization problem:

$$z_1 = \arg \min_{z \in C} \langle w_1 - v, z - w_1 \rangle = \arg \min_{z \in C} \langle (0, 0) - (0, 0), (z_1, z_2) - (0, 0) \rangle.$$

Simplifying:

$$\langle (0, 0), (z_1, z_2) \rangle = 0.$$

The optimal solution is any (z_1, z_2) over the unit square C .

The optimal value is:

$$s_1^* = \langle (0, 0), (z_1, z_2) \rangle = 0.$$

Step 2: Stopping Condition

The algorithm checks if $-s_1^* \leq \theta \|v - u\|^2$.

- We know $s_1^* = 0$, and as a result, $-s_1^* = 0 \leq \theta \|v - u\|^2 = 0.02$, which is **true**.

Thus, the algorithm stops and returns $w^+ = w_1 = (0, 0)$.

Conclusion

The algorithm terminates with the solution $w^+ = (0, 0)$ after two iterations.

2.6 Broyden Update Rule

The subsequent definition pertains to the Broyden update rule, a significant numerical optimization tool that serves as an iterative approximation for the Jacobian matrix f' in quasi-Newton methods. This rule is employed to update the approximate Jacobian in each iteration.

Definition 2.6.1 The Broyden update rule is defined as

$$B_{k+1} := B_k + \frac{(z_k - B_k s_k) s_k^T}{\|s_k\|^2}, \quad (2.30)$$

where $z_k := f(y_k) - f(x_k)$ and $s_k := y_k - x_k$. A practical usual choice for B_0 is $B_0 = f'(x_0)$.

Example 2.6.2 Consider the function $f(x) = \begin{bmatrix} 2x_1 - x_2 \\ -x_1 + 3x_2 \end{bmatrix}$ and two points $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $y_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$.

We aim to apply the Broyden update rule to approximate the Jacobian matrix.

Step 1: Compute z_0 and s_0

$$z_0 = f(y_0) - f(x_0) = \begin{bmatrix} 2(0.5) - 0.5 \\ -0.5 + 3(0.5) \end{bmatrix} - \begin{bmatrix} 2(1) - 1 \\ -1 + 3(1) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -0.5 \\ -1 \end{bmatrix}.$$

$$s_0 = y_0 - x_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.5 \\ -0.5 \end{bmatrix}$$

Step 2: Initialize B_0

A common choice for B_0 is the Jacobian $f'(x_0)$, which for $f(x)$ is:

$$B_0 = f'(x_0) = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}.$$

Step 3: Apply the Broyden update rule

Using the formula $B_{k+1} = B_k + \frac{(z_k - B_k s_k) s_k^T}{\|s_k\|^2}$, we first compute $B_0 s_0$:

$$B_0 s_0 = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} -0.5 \\ -0.5 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

Now compute the difference $z_0 - B_0 s_0$:

$$z_0 - B_0 s_0 = \begin{bmatrix} -0.5 \\ -1 \end{bmatrix} - \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix}.$$

The norm $\|s_0\|^2$ is:

$$\|s_0\|^2 = (-0.5)^2 + (-0.5)^2 = 0.5.$$

Now update B_1 :

$$B_1 = B_0 + \frac{\begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \begin{bmatrix} -0.5 & -0.5 \end{bmatrix}}{0.5} = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} + \begin{bmatrix} -0.5 & -0.5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1.5 & -1.5 \\ -1 & 3 \end{bmatrix}.$$

Final Matrix:

The updated matrix B_1 is:

$$B_1 = \begin{bmatrix} 1.5 & -1.5 \\ -1 & 3 \end{bmatrix}.$$

In the context of quasi-Newton methods, such as the Broyden algorithm, the iterative update of the matrix B_k is a central component. These methods rely on the idea that the sequence of matrices $\{B_k\}$ generated at each step serves as an effective approximation of the true Jacobian of the function f by the solution point x_* . The quality of this approximation is paramount for ensuring the convergence and efficiency of the algorithm.

2.6.1 The Bounded Deterioration Condition

The **bounded deterioration condition**, in the next definition, is a crucial property that addresses a key challenge in these methods: controlling the error of the Jacobian approximation. It ensures that the increase in the approximation error is bounded by a term related to the proximity of the current iterates x_k and y_k to the solution x_* , thus preserving the desirable convergence properties of the method.

Definition 2.6.3 *If there exists a constant $c > 0$ such that, for each $k \geq 0$,*

$$\|B_{k+1} - f'(x_*)\| \leq \|B_k - f'(x_*)\| + c(\|x_k - x_*\| + \|y_k - x_*\|). \quad (2.31)$$

*then we say that B_{k+1} satisfies **the bounded deterioration condition**.*

Lemma 2.6.4 Let $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. If $x \in \mathbb{R}^n \setminus \{0\}$, then

$$\left\| A - \frac{\langle x, \cdot \rangle Ax}{\|x\|^2} \right\| = \begin{cases} 0 & \text{if } n = 1 \\ \|A\| & \text{if } n > 1 \end{cases}. \quad (2.32)$$

Proof. Let $z \in \text{span}(x) = \{\lambda x \mid \lambda \in \mathbb{R}\}$. Then there is some $\lambda_0 \in \mathbb{R}$ such that $z = \lambda_0 x$. Hence,

$$\left\| Az - \frac{\langle x, z \rangle Ax}{\|x\|^2} \right\| = \left\| \lambda_0 Ax - \frac{\lambda_0 \langle x, x \rangle Ax}{\|x\|^2} \right\| = 0.$$

If $n = 1$, then $\mathbb{R}^n = \text{span}(x)$, and from the above equality we obtain (2.32). Otherwise, assume that $n > 1$. For any $z \in \mathbb{R}^n$, $\|z\| \leq 1$, one has

$$\left\| z - \frac{\langle x, z \rangle x}{\|x\|^2} \right\|^2 = \|z\|^2 - \frac{\langle x, z \rangle^2}{\|x\|^2} \leq 1.$$

Hence,

$$\left\| A - \frac{\langle x, \cdot \rangle Ax}{\|x\|^2} \right\| = \sup_{\|w\| \leq 1} \left\| A \left(w - \frac{\langle x, w \rangle x}{\|x\|^2} \right) \right\| \leq \|A\| \sup_{\|w\| \leq 1} \left\| w - \frac{\langle x, w \rangle x}{\|x\|^2} \right\| \leq \|A\|.$$

For any $z \in \{x\}^\perp = \{w \in \mathbb{R}^n \mid \langle w, x \rangle = 0\} \neq \emptyset$, one has

$$\left\| Az - \frac{\langle x, z \rangle Ax}{\|x\|^2} \right\| = \|Az\|,$$

and therefore, (2.32) follows. ■

Proposition 2.6.5 Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuously differentiable function, such that f' is Lipschitz continuous with constant L , in a convex neighborhood U of a point x_* . Given $B_k \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ and $x_k, y_k \in U$, $y_k \neq x_k$, the operator B_{k+1} , defined as in (2.30) satisfies

$$\|B_{k+1} - f'(x_*)\| \leq \|B_k - f'(x_*)\| + \frac{L}{2}(\|y_k - x_*\| + \|x_* - x_k\|), \quad (2.33)$$

i.e., B_{k+1} satisfies the bounded deterioration condition.

Proof. By assumption,

$$\|f'(u) - f'(v)\| \leq L\|u - v\|, \quad \forall u, v \in U.$$

Let $y_k, x_k \in U, y_k \neq x_k$ and let B_{k+1} be defined as in (2.6.1). Then

$$\begin{aligned} B_{k+1} - f'(x_*) &= B_k - f'(x_*) + \frac{(z_k - B_k s_k) s_k^T}{\|s_k\|^2} \\ &= B_k - f'(x_*) - \frac{(B_k - f'(x_*)) s_k s_k^T}{\|s_k\|^2} + \frac{(z_k - f'(x_*) s_k) s_k^T}{\|s_k\|^2}. \end{aligned}$$

Thus,

$$\|B_{k+1} - f'(x_*)\| \leq \left\| (B_k - f'(x_*)) - \frac{(B_k - f'(x_*)) s_k s_k^T}{\|s_k\|^2} \right\| + \frac{\|z_k - f'(x_*) s_k\|}{\|s_k\|}.$$

By Lemma 2.6.4,

$$\left\| (B_k - f'(x_*)) - \frac{(B_k - f'(x_*)) s_k s_k^T}{\|s_k\|^2} \right\| \leq \|B_k - f'(x_*)\|.$$

Using the mean value theorem, we obtain

$$\begin{aligned} \|z_k - f'(x_*) s_k\| &= \|f(y_k) - f(x_k) - f'(x_*) s_k\| \\ &= \left\| \int_0^1 [f'(x_k + t(y_k - x_k))(y_k - x_k) - f'(x_*) s_k] dt \right\| \\ &\leq \|s_k\| \int_0^1 \|f'(x_k + t(y_k - x_k)) - f'(x_*)\| dt \\ &\leq L \|s_k\| \int_0^1 ((1-t)\|x_k - x_*\| + t\|y_k - x_*\|) dt \\ &= \frac{L}{2} \|s_k\| (\|y_k - x_*\| + \|x_k - x_*\|). \end{aligned}$$

This yields (2.33). ■

2.7 Convergence q-linear and Superlinear

Definition 2.7.1 A sequence $\{x_k\}$ is **convergent q-linearly** to \bar{x} when there exist a natural K and a real $\alpha \in [0, 1)$ such that

$$\|x_{k+1} - \bar{x}\| \leq \alpha \|x_k - \bar{x}\| \text{ for } k = K, K+1, \dots$$

Example 2.7.2 $x_k = \frac{1}{k+5}$ is not convergent q-linearly to 0.

$$\text{In fact, } \frac{\|x_{k+1}\|}{\|x_k\|} = \frac{k+5}{k+6} \rightarrow 1.$$

Example 2.7.3 $x_k = \frac{1}{3^k}$ is convergent q-linearly to 0.

$$\text{In fact, } \frac{\|x_{k+1}\|}{\|x_k\|} = \frac{1}{3}.$$

Definition 2.7.4 A sequence $\{x_k\}$ is convergent **super linearly** to \bar{x} when

$$\frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} \rightarrow 0.$$

Example 2.7.5 $x_k = \frac{1}{2^{k^2}}$ is convergent super linearly to 0.

$$\text{In fact, } \frac{\|x_{k+1}\|}{\|x_k\|} = \frac{x^{k^2}}{x^{(k+1)^2}} = \frac{1}{2^{2k+1}} \rightarrow 0.$$

Chapter 3

Secant-type method with feasible inexact projection for solving constrained mixed generalized equations

In this chapter we present an *secant method* to solve the constrained problem (1.1). Under metric regularity of the partial linearization of $f+g+F$, Lipschitz continuity of the derivative f' and boundedness of the second order divided difference of g , we will analyze the sequence generated by the inexact secant method.

3.1 The Secant-InexP Algorithm

The conceptual secant method with a feasible inexact projection, named the *Secant-InexP* method for solving the mixed generalized equation (1.1) and with $x_{-1}, x_0 \in C$ and $(\theta_k)_{k \in \mathbb{N}} \subset [0, +\infty)$ as the input data, is formally described as follows:

Algorithm 3.1.1 Secant-InexP

Step 0. Let $x_{-1}, x_0 \in C$ and $(\theta_k)_{k \in \mathbb{N}} \subset [0, +\infty)$ be given, and set $k = 0$.

Step 1. If $f(x_k) + g(x_k) + F(x_k) \ni 0$, then **stop**; otherwise, compute $y_k \in \mathbb{R}^n$ such that

$$f(x_k) + g(x_k) + (f'(x_k) + [x_{k-1}, x_k; g])(y_k - x_k) + F(y_k) \ni 0. \quad (3.1)$$

Step 2. If $y_k \in C$, set $x_{k+1} = y_k$; otherwise take any $x_{k+1} \in C$ satisfying

$$x_{k+1} \in P_C(y_k, x_k, \theta_k). \quad (3.2)$$

Step 3. Set $k \leftarrow k + 1$, and go to **Step 1**.

In **Step 1** of Algorithm 3.1.1, we first check if the current iterate x_k is a solution of (1.1). If not, then we compute y_k satisfying the inclusion (3.1). Since the point y_k in **Step 1** may not belong to the constraint set C , the Secant-InexP method uses Algorithm 2.5.4 to compute x_{k+1} in C . It is worth noting that x_{k+1} obtained in (3.2) is an approximate solution to the projection subproblem $\min_{p \in C} \|p - y_k\|^2/2$, satisfying the stopping criterion $\langle y_k - x_{k+1}, p - x_{k+1} \rangle \leq \theta_k \|y_k - x_k\|^2$, for any $p \in C$.

Remark 3.1.2 Whenever $g = 0$ and $F \equiv \{0\}$ the above algorithm merges into the one proposed in [33]. On the other hand, if $C = \mathbb{R}^n$ then Algorithm 3.1.1 merges into the finite-dimensional version of the secant method proposed in [32].

In order to analyze the convergence properties of Algorithm 3.1.1, we will need some preliminary results. For that, we assume throughout our presentation that the following assumptions hold:

(A1) there exist $L > 0$ and $M > 0$ such that

$$\|f'(x) - f'(y)\| \leq L\|x - y\|, \quad \|[x, y, z; g]\| \leq M, \quad \forall x, y, z \in \Omega, \quad (3.3)$$

where Ω is the domain of f and g .

(A2) for each $z \in \Omega$, the set-valued mapping $\Omega \ni y \mapsto L_{f+g+F}(x_*, y, z)$ is metrically regular at x_* for 0, with constants $\kappa > 0$, $a > 0$, and $b > 0$.

For simplifying notation, we fix constants $\bar{\delta} > 0$, $\bar{\Theta} > 0$ and $\bar{\theta}$ satisfying the following conditions:

$$\bar{\delta} < \bar{\Theta} := \min \left\{ \frac{2(1 - \sqrt{2\bar{\theta}})}{\kappa \left[(3 - \sqrt{2\bar{\theta}})L + 4M \right]}, \sqrt{\frac{2b}{3L + 4M}} \right\}, \quad \bar{\theta} := \sup_k \theta_k < \frac{1}{2}, \quad (3.4)$$

where $(\theta_k)_{k \in \mathbb{N}}$ is a sequence of nonnegative numbers. We begin by establishing two technical results that will be useful for our analysis.

Lemma 3.1.3 *Let $\bar{\delta}$ and $\bar{\Theta}$ be as in (3.4). For all $x \in B_{\bar{\delta}}(x_*)$, the following inequality holds:*

$$\|f(x_*) - f(x) - f'(x)(x_* - x)\| \leq \frac{L}{2} \|x_* - x\|^2.$$

Moreover, if x, z and p are distinct points in $B_{\bar{\delta}}[x_*]$ then

$$\|E_{f+g}(x, p, z) - E_{f+g}(x_*, p, z)\| \leq b.$$

Proof. Since $x_* + (1 - \tau)(x - x_*) \in B_{\bar{\delta}}(x_*)$ for all $\tau \in [0, 1]$ and f is continuously differentiable in Ω , the first inequality of the lemma can be obtained by combining the fundamental theorem of calculus and the first inequality in **(A1)** in a straightforward way. We proceed to prove the second inequality. For this purpose, let x, z and p be distinct points in $B_{\bar{\delta}}[x_*]$. By the definition of the mapping $E_{f+g}(x, \cdot, z)$ in (2.2) and the triangular inequality, after some algebraic manipulations, we have

$$\begin{aligned} & \|E_{f+g}(x, p, z) - E_{f+g}(x_*, p, z)\| = \\ & \|f(x_*) + g(x_*) + (f'(x_*) + [z, x_*; g])(p - x_*) - f(x) - g(x) - (f'(x) + [z, x; g])(p - x)\| \leq \\ & \|f(x_*) - f(x) - f'(x)(x_* - x)\| + \|f'(x_*) - f'(x)\| \|p - x_*\| + \\ & \|g(x_*) - g(x) - [z, x; g](x_* - x)\| + \|[z, x_*; g](p - x_*) - [z, x; g](p - x_*)\|. \end{aligned} \quad (3.5)$$

The first inequality of this lemma together the Lipschitz condition in (3.3) imply that

$$\begin{aligned} & \|f(x_*) - f(x) - f'(x)(x_* - x)\| + \|f'(x) - f'(x_*)\| \|x_* - p\| \leq \\ & \frac{L}{2} \|x_* - x\|^2 + L \|x_* - x\| \|x_* - p\|. \end{aligned} \quad (3.6)$$

On the other hand, by Definition 2.2.1 we obtain

$$\|g(x_*) - g(x) - [z, x; g](x_* - x)\| = \|[z, x, x_*; g](x_* - z)(x_* - x)\|, \quad (3.7)$$

and again using Definition 2.2.1 we have

$$\begin{aligned} & \|[z, x_*; g](p - x_*) - [z, x; g](p - x_*)\| = \|([z, x_*; g] - [z, x; g])(p - x_*)\| \\ & = \|([x, z, x_*; g](x_* - x) + [x, z; g] - [z, x; g])(p - x_*)\| \\ & = \|[x, z, x_*; g](x_* - x)(p - x_*)\|. \end{aligned} \quad (3.8)$$

Combining (3.5) with (3.6), (3.7) and (3.8) and taking into account that $\|[x, y, z; g]\| \leq M$ for all x, y, z in a neighborhood of x_* , we conclude that

$$\begin{aligned} & \|E_{f+g}(x, p, z) - E_{f+g}(x_*, p, z)\| \\ & \leq \frac{L}{2} \|x_* - x\|^2 + L \|x_* - x\| \|x_* - p\| + M \|x_* - z\| \|x_* - x\| + M \|x_* - x\| \|p - x_*\|. \end{aligned}$$

Considering that $x, z, p \in B_{\bar{\delta}}[x_*]$, the last inequality implies

$$\|E_{f+g}(x, p, z) - E_{f+g}(x_*, p, z)\| \leq \frac{L}{2} \bar{\delta}^2 + L \bar{\delta}^2 + 2M \bar{\delta}^2 = \frac{3L + 4M}{2} \bar{\delta}^2.$$

Since (3.4) implies that $\bar{\delta} \leq \sqrt{2b/(3L + 4M)}$, the desired inequality follows from the last inequality. Therefore, the proof of the lemma is concluded. \blacksquare

For stating the next result, we define, for each fixed $x, z \in \mathbb{R}^n$, the following auxiliary set-valued mapping $\Phi_{(x,z)} : \Omega \rightrightarrows \mathbb{R}^n$:

$$\Phi_{(x,z)}(y) := L_{f+g+F}(x_*, E_{f+g}(x, y, z) - E_{f+g}(x_*, y, z), z)^{-1}, \quad (3.9)$$

where

$$\mathbb{R}^m \ni u \mapsto L_{f+g+F}(x_*, u, z)^{-1} := \{w \in \mathbb{R}^n : u \in L_{f+g+F}(x_*, w, z)\}$$

is the inverse of the partial linearization L_{f+g+F} defined in (2.3). By using (2.3) and (3.9), we obtain that, if $y \in \Phi_{(x,z)}(y)$, then x, y and z satisfy the inclusion

$$f(x) + g(x) + (f'(x) + [z, x; g])(y - x) + F(y) \ni 0.$$

Hence, a fixed point y of $\Phi_{(x,z)}$ is the next *secant iterate* from x . Our next task is to establish the existence of a fixed point to $\Phi_{(x,z)}$ for each x, z in a suitable neighborhood of x_* . Moreover, for a given fixed point y of $\Phi_{(x,z)}$, we also need to bound the distance between $x_* \in C$, a solution of (1.1), and the inexact projections of y onto the set C . The statement of this result is the following:

Lemma 3.1.4 *Let $\bar{\delta}$ and $\bar{\Theta}$ be as in (3.4). Assume that $x_* \in C$ is a solution of (1.1). Take $\theta \geq 0$ and $x, z \in C \cap B_{\bar{\delta}}(x_*) \setminus \{x_*\}$. Then, the mapping $\Phi_{(x,z)}$ has a fixed point $y \in \Omega$ and, for every $w \in P_C(y, x, \theta)$, holds*

$$\|x_* - w\| \leq \left[\left(1 + \sqrt{2\theta}\right) \frac{\kappa \left[\frac{L}{2} \|x_* - x\| + M \|x_* - z\| \right]}{1 - \kappa(L + M) \|x_* - x\|} + \sqrt{2\theta} \right] \|x_* - x\|. \quad (3.10)$$

In addition, if $\theta < 1/2$ then $P_C(y, x, \theta) \subset B_{\bar{\delta}}(x_) \cap C$.*

Proof. Using the definition of the set-valued mapping $\Phi_{(x,z)}$ given in (3.9) we arrive at

$$d(x_*, \Phi_{(x,z)}(x_*)) = d(x_*, L_{f+g+F}(x_*, E_{f+g}(x, x_*, z) - E_{f+g}(x_*, x_*, z), z)^{-1}). \quad (3.11)$$

Thus, applying the second part of Lemma 4.1.3 with $p = x_*$ we obtain

$$\|E_{f+g}(x, x_*, z) - E_{f+g}(x_*, x_*, z)\| < b,$$

and owing to $x_* \in B_a[x_*]$ and $0 \in L_{f+g+F}(x_*, x_*, z)$, we can use Definition 2.4.1 to conclude that

$$\begin{aligned} d(x_*, \Phi_{(x,z)}(x_*)) &\leq \kappa \|E_{f+g}(x, x_*, z) - E_{f+g}(x_*, x_*, z)\| \\ &= \kappa \|f(x_*) + g(x_*) - f(x) - g(x) - (f'(x) + [z, x; g])(x_* - x)\|. \end{aligned}$$

Since the first part of Lemma 4.1.3 implies $\|f(x_*) - f(x) - f'(x)(x_* - x)\| \leq (L/2)\|x_* - x\|^2$, using Definition 2.2.1 we arrive at

$$d(x_*, \Phi_{(x,z)}(x_*)) \leq \frac{\kappa L}{2} \|x_* - x\|^2 + \kappa \|[z, x, x^*; g](x^* - z)(x_* - x)\|.$$

Thus, due to $\|[z, x, x^*; g]\| \leq M$ for all $x, z, x_* \in \Omega$, the last inequality yields

$$d(x_*, \Phi_{(x,z)}(x_*)) \leq \frac{\kappa \left[\frac{L}{2} \|x_* - x\| + M \|x_* - z\| \right]}{1 - \kappa(L + M) \|x_* - x\|} \|x_* - x\| (1 - \kappa(L + M) \|x_* - x\|). \quad (3.12)$$

To simplify the notation, for every x, z in a neighborhood of x_* , we define the quantities

$$r := \frac{\kappa \left[\frac{L}{2} \|x_* - x\| + M \|x_* - z\| \right]}{1 - \kappa(L + M) \|x_* - x\|} \|x_* - x\|, \quad \lambda := \kappa(L + M) \|x_* - x\|. \quad (3.13)$$

From (3.4), $\bar{\delta} < \bar{\Theta}$. Thus, due to $\|x_* - x\| \leq \bar{\delta}$, we have $1 - \kappa(L + M) \|x_* - x\| > 0$. Hence, $r > 0$. Now, combining (3.12) with (3.13), we obtain

$$d(x_*, \Phi_{(x,z)}(x_*)) \leq r(1 - \lambda). \quad (3.14)$$

Take $p, q \in B_r[x_*]$. In order to apply Theorem 2.4.22, we will bound the excess $e(\Phi_{(x,z)}(p) \cap B_r[x_*], \Phi_{(x,z)}(q))$. Since $\bar{\delta} < \bar{\Theta}$, some algebraic manipulations show that $r \leq \bar{\delta}$. Thus, we also have $p, q \in B_{\bar{\delta}}[x_*]$. Considering that $x, z \in B_{\bar{\delta}}[x_*]$, it follows from Lemma 4.1.3 that

$$\|E_{f+g}(x, p, z) - E_{f+g}(x_*, p, z)\| < b, \quad \|E_{f+g}(x, q, z) - E_{f+g}(x_*, q, z)\| < b. \quad (3.15)$$

Since $e(\emptyset, \Phi_{(x,z)}(q)) = 0$, we can assume that $\Phi_{(x,z)}(p) \cap B_r[x_*] \neq \emptyset$, for all $p \in B_r[x_*]$. Let $v \in \Phi_{(x,z)}(p) \cap B_r[x_*]$, thus using Definition 2.4.1 with $u = E_{f+g}(x, q, z) - E_{f+g}(x_*, q, z)$, $\bar{x} = x_*$, $\bar{u} = 0$, $x = v$ and $G = L_{f+g+F}(x_*, \cdot, z)$, and using (3.9) and (3.15) we conclude that

$$d(v, \Phi_{(x,z)}(q)) \leq \kappa d(E_{f+g}(x, q, z) - E_{f+g}(x_*, q, z), L_{f+g+F}(x_*, v, z)).$$

As $v \in \Phi_{(x,z)}(p)$, from (3.9) it follows that $E_{f+g}(x, p, z) - E_{f+g}(x_*, p, z) \in L_{f+g+F}(x_*, v, z)$. Thus, by using the definition of distance given in (2.1) we obtain

$$d(E_{f+g}(x, q, z) - E_{f+g}(x_*, q, z), L_{f+g+F}(x_*, v, z)) \leq (\|f'(x) - f'(x_*)\| + \|[z, x; g] - [z, x_*; g]\|) \|p - q\|.$$

Hence, combining Definition 2.2.1 with the last two inequalities we conclude that

$$d(v, \Phi_{(x,z)}(q)) \leq \kappa \| [f'(x) - f'(x_*)](p - q) \| + \kappa \|[x_*, z, x; g](x - x_*)\| \|q - p\|,$$

so, considering that $\|[x_*, z, x; g]\| \leq M$ for all $x, z, x_* \in \Omega$ and using (3.3), we obtain

$$d(v, \Phi_{(x,z)}(q)) \leq \kappa(L + M)\|x - x_*\|\|q - p\|.$$

Taking the supremum with respect to $v \in \Phi_{(x,z)}(p) \cap B_r[x_*]$ in the last inequality, using (3.13) and the definition of excess given in (2.1), we have

$$e(\Phi_{(x,z)}(p) \cap B_r[x_*], \Phi_{(x,z)}(q)) \leq \lambda\|q - p\|. \quad (3.16)$$

Therefore, from (3.13), (3.14), (3.16) and Theorem 2.4.22 we conclude that there exists $y \in \Phi_{(x,z)}(y)$ with $y \in B_r[x_*]$. Hence, y satisfies the inequality

$$\|x_* - y\| \leq \frac{\kappa \left[\frac{L}{2}\|x_* - x\| + M\|x_* - z\| \right]}{1 - \kappa(L + M)\|x_* - x\|} \|x_* - x\|. \quad (3.17)$$

Furthermore, we conclude from (3.13) and (3.17) that $y \in B_r(x_*)$. Now, take $w \in P_C(y, x, \theta)$. Applying Lemma 2.5.3 with $\tilde{y} = x_*$ and $\tilde{x} = x_*$ we have

$$\begin{aligned} \|P_C(x_*, x_*, 0) - w\| &\leq \|x_* - y\| + \sqrt{2\theta}(\|x_* - x\| + \|x_* - y\|) \\ &= \left(1 + \sqrt{2\theta}\right) \|x_* - y\| + \sqrt{2\theta}\|x_* - x\|. \end{aligned}$$

Considering that y satisfies (3.17) and $P_C(x_*, x_*, 0) = x_*$, after some manipulations, the desired inequality (3.10) follows from the last inequality. Moreover, owing to $x, z \in B_{\bar{\delta}}(x_*) \setminus \{x_*\}$, using (3.4) together with $\theta < 1/2$ yields

$$\left[\left(1 + \sqrt{2\theta}\right) \frac{\kappa \left[\frac{L}{2}\|x_* - x\| + M\|x_* - z\| \right]}{1 - \kappa(L + M)\|x_* - x\|} + \sqrt{2\theta} \right] < 1.$$

Thus, it follows from (3.10) that $\|x_* - w\| < \|x_* - x\|$ for all $w \in P_C(y, x, \theta)$. Due to $\|x_* - x\| < \bar{\delta}$ we obtain $P_C(y, x, \theta) \subset B_{\bar{\delta}}(x_*)$. Because $P_C(y, x, \theta) \subset C$, the last statement of the lemma follows, which concludes the proof. \blacksquare

3.2 The Convergence Theorem

In the following, we state our main local convergence theorem for the sequence generated by the Secant-InexP method.

Theorem 3.2.1 *Let $\Omega \subset \mathbb{R}^n$ be an open set, $f : \Omega \rightarrow \mathbb{R}^n$ be a continuously differentiable function in Ω , $g : \Omega \rightarrow \mathbb{R}^n$ be a continuous function not necessarily differentiable and $F : \Omega \rightrightarrows \mathbb{R}^n$ be a set-valued mapping with closed graph. Assume that $C \subset \Omega$ is a closed convex set, $x_* \in C$, $f(x_*) + g(x_*) + F(x_*) \ni 0$, and assumptions **(A1)** and **(A2)** hold. Let*

L be as defined in **(A1)** and κ , a and b be constants as defined in **(A2)**. Let $\bar{\delta}$, $\bar{\Theta}$ and $\bar{\theta}$ be constants and $(\theta_k)_{k \in \mathbb{N}}$ be a sequence satisfying (3.4). Set $\gamma := \sup \{t \in \mathbb{R} : B_t(x_*) \subset \Omega\}$ and

$$0 < \delta < \min \{\gamma, a, \bar{\delta}\}. \quad (3.18)$$

If $x_{-1}, x_0 \in C \cap B_\delta(x_*) \setminus \{x_*\}$, then every sequence $(x_k)_{k \in \mathbb{N}}$ generated by Algorithm 3.1.1 belongs to $B_\delta(x_*) \cap C$. Moreover, $(x_k)_{k \in \mathbb{N}}$ converges to x_* and

$$\|x_* - x_{k+1}\| \leq \left(\left(1 + \sqrt{2\theta_k}\right) \frac{\kappa \left(\frac{L}{2}\|x_* - x_k\| + M\|x_* - x_{k-1}\|\right)}{1 - \kappa(L + M)\|x_* - x_k\|} + \sqrt{2\theta_k} \right) \|x_* - x_k\|. \quad (3.19)$$

As a consequence, if $\lim_{k \rightarrow +\infty} \theta_k = 0$, then $(x_k)_{k \in \mathbb{N}}$ converges to x_* superlinearly. In particular, if $\theta_k = 0$ for all $k = 0, 1, \dots$, then

$$\|x_* - x_{k+1}\| \leq \frac{\kappa(3L + 4M)}{2} \|x_* - x_k\| \max\{\|x_* - x_{k-1}\|, \|x_* - x_k\|\}, \quad k = 0, 1, \dots \quad (3.20)$$

Proof. First we will show by induction on k that, for any $x_{-1}, x_0 \in B_\delta(x_*) \cap C$, there exists a sequence $(x_k)_{k \in \mathbb{N}}$ satisfying (3.1) and (3.2) (i.e., $(x_k)_{k \in \mathbb{N}}$ is well-defined), (3.19) and $x_k \in B_\delta(x_*) \cap C$ for all $k \in \mathbb{N}$. Since $x_{-1}, x_0 \in B_\delta(x_*) \setminus \{x_*\} \cap C$, we can apply Lemma 3.1.4 with $x = x_0$ and $z = x_{-1}$ for concluding that there exists $y_0 \in \Phi_{(x_0, x_{-1})}(y_0)$ such that any $x_1 \in P_C(y_1, x_1, \theta_k) \subset B_\delta(x_*) \cap C$ satisfies (3.19) for $k = 0$. Thus, the last two previous conditions hold for $k = 0$. Assume by induction that them hold for $k = 1, 2, \dots, j - 1$. Since $x_{j-1}, x_j \in B_\delta(x_*) \setminus \{x_*\} \cap C$, we can apply again Lemma 3.1.4 with $x = x_j$ and $z = x_{j-1}$ to conclude that there exists $y_j \in \Phi_{(x_j, x_{j-1})}(y_j)$ such that any $x_{j+1} \in P_C(y_j, x_j, \theta_k) \subset B_\delta(x_*) \cap C$ satisfies (3.19) for $k = j + 1$, completing the induction step. In particular, (3.1) and (3.2) are satisfied for all $k = 1, 2, \dots$ and $(x_k)_{k \in \mathbb{N}}$ is well-defined.

We now proceed to prove that $(x_k)_{k \in \mathbb{N}}$ converge to x_* . Indeed, since $(x_k)_{k \in \mathbb{N}} \subset B_\delta(x_*)$, $\bar{\theta} = \sup_k \theta_k < 1/2$ and $\delta \leq [2(1 - \sqrt{2\bar{\theta}})] / [\kappa((3 - \sqrt{2\bar{\theta}})L + 4M)]$, we conclude from (3.19) that

$$\|x_* - x_{k+1}\| < \|x_* - x_k\|.$$

This implies that the sequence $(\|x_* - x_k\|)_{k \in \mathbb{N}}$ converges, let us say, $t_* = \lim_{k \rightarrow \infty} \|x_* - x_k\|$. Owing to $(x_k)_{k \in \mathbb{N}} \subset B_{r_*(x_*)} \cap C$ we conclude that $t_* \leq \delta$. On the other hand, letting k goes to $+\infty$ in (3.19) and using that $t_* = \lim_{k \rightarrow \infty} \|x_* - x_k\| \leq \delta$ we have

$$t_* \leq \left[\left(1 + \sqrt{2\bar{\theta}}\right) \frac{\kappa \left[\frac{L}{2} + M\right] \delta}{1 - \kappa(L + M)\delta} + \sqrt{2\bar{\theta}} \right] t_*.$$

If $t_* \neq 0$ we obtain from the last inequality that

$$\bar{\delta} < \frac{2(1 - \sqrt{2\bar{\theta}})}{\kappa \left[\left(3 - \sqrt{2\bar{\theta}}\right) L + 4M \right]} \leq \delta,$$

which contradicts (3.18). Hence, $t_* = 0$ and consequently the sequence $(x_k)_{k \in \mathbb{N}}$ converges to x_* . In particular, if $\lim_{k \rightarrow +\infty} \theta_k = 0$, taking the limit in (3.19) as k goes to $+\infty$ we obtain $\limsup_{k \rightarrow +\infty} [\|x_* - x_{k+1}\| / \|x_* - x_k\|] = 0$, i.e., the sequence $(x_k)_{k \in \mathbb{N}}$ converges to x_* superlinearly.

Finally, to prove (3.20), first let $\theta_k = 0$ for all $k = 0, 1, \dots$, in (3.19) to obtain

$$\|x_* - x_{k+1}\| \leq \left(\frac{\kappa \left(\frac{L}{2} + M \right)}{1 - \kappa(L + M)\|x_* - x_k\|} \right) \|x_* - x_k\| \max\{\|x_* - x_{k-1}\|, \|x_* - x_k\|\} \quad (3.21)$$

for all $k = 0, 1, \dots$. Moreover, $\theta_k = 0$ for all $k = 0, 1, \dots$ implies that $\bar{\theta} = 0$. Thus, (3.4) gives

$$\|x_* - x_k\| < \bar{\delta} < \frac{2}{\kappa(3L + 4M)},$$

which combined with (3.21) yields (3.20), concluding the proof. \blacksquare

Remark 3.2.2 Let us present relationships of Theorem 3.2.1 with some existing results in the literature:

- (i) If $C = \mathbb{R}^n$ then $\theta_k = 0$ for all $k = 0, 1, \dots$. In this case, letting $g = 0$ we conclude that Theorem 3.2.1 merges into [20, Theorem 1].
- (ii) Taking $C = \mathbb{R}^n$ and $\theta_k = 0$ for all $k = 0, 1, \dots$, after some adjustments Theorem 3.2.1 can be seen as version of [32, Theorem 3.1] for Euclidean setting.
- (iii) Let $\hat{\theta} = \limsup_{k \rightarrow +\infty} \theta_k$. Thus, (3.19) implies $\limsup_{k \rightarrow +\infty} [\|x_* - x_{k+1}\| / \|x_* - x_k\|] \leq \sqrt{2\hat{\theta}}$. Therefore, letting $g = 0$ and $F \equiv 0$, we obtain that Theorem 3.2.1 is related to [33, Theorem 6].
- (iv) Letting $g = 0$, we conclude that Theorem 3.2.1 merges into [18, Theorem 2].

3.2.1 The Uniqueness

Now, let us study the uniqueness of the solution for (1.1) in the neighborhood $B_\delta(x_*)$.

Theorem 3.2.3 *Let $\Omega \subset \mathbb{R}^n$ be an open set, $f : \Omega \rightarrow \mathbb{R}^n$ be a continuously differentiable function in Ω , $g : \Omega \rightarrow \mathbb{R}^n$ be a continuous function not necessarily differentiable and $F : \Omega \rightrightarrows \mathbb{R}^n$ be a set-valued mapping with closed graph. Assume that $C \subset \Omega$ is a closed convex set, $x_* \in C$, $f(x_*) + g(x_*) + F(x_*) \ni 0$. Let L be as defined in **(A1)**. Also, suppose that the set-valued mapping $\Omega \ni y \mapsto L_{f+g+F}(x_*, y, z)$ is strongly metrically regular at x_* for 0 with constants $\kappa > 0$, $a > 0$ and $b > 0$. Let $\bar{\delta}$, $\bar{\Theta}$ and $\bar{\theta}$ be constants and $(\theta_k)_{k \in \mathbb{N}}$ be a sequence satisfying (3.4). Set $\gamma := \sup\{t \in \mathbb{R} : B_t(x_*) \subset \Omega\}$ and $\delta > 0$ satisfying (3.18). Then x_* is the unique solution of (1.1) in $B_\delta(x_*)$.*

Proof. Let \hat{x} be a solution of (1.1) in $B_\delta(x_*)$ and assume that $\hat{x} \neq x_*$. In order to simplify the notations, for a given $z \in B_\delta(x_*)$ with $z \neq x_*$ and $z \neq \hat{x}$, we define

$$u(z, x_*, \hat{x}) := f(\hat{x}) + g(\hat{x}) - f(x_*) - g(x_*) - f'(x_*)(\hat{x} - x_*) - [z, x_*; g](\hat{x} - x_*). \quad (3.22)$$

Thus, using triangular inequality we conclude that

$$\|u(z, x_*, \hat{x})\| \leq \|f(\hat{x}) - f(x_*) - f'(x_*)(\hat{x} - x_*)\| + \|g(\hat{x}) - g(x_*) - [z, x_*; g](\hat{x} - x_*)\|. \quad (3.23)$$

Our first task is to bound the two terms in the right hand side of the last inequality. By using the first part of Lemma 4.1.3 and considering that $\hat{x} \in B_\delta(x_*)$, we bound the first term as follows

$$\|f(\hat{x}) - f(x_*) - f'(x_*)(\hat{x} - x_*)\| \leq \frac{L}{2} \|\hat{x} - x_*\|^2 < \frac{L}{2} \delta^2. \quad (3.24)$$

Thus, to bound the second term, we first take $z \in B_\delta(x_*)$ with $z \neq x_*$ and $z \neq \hat{x}$. Since $\hat{x} \neq x_*$, using the first part Definition 2.2.1 we have

$$\|g(\hat{x}) - g(x_*) - [z, x_*; g](\hat{x} - x_*)\| = \|[x_*, \hat{x}; g](\hat{x} - x_*) - [z, x_*; g](\hat{x} - x_*)\|.$$

Considering that $\|[x_*, \hat{x}; g](\hat{x} - x_*) - [z, x_*; g](\hat{x} - x_*)\| \leq \|[x_*, \hat{x}; g] - [z, x_*; g]\| \|\hat{x} - x_*\|$, and due $\hat{x} \neq x_*$, $z \neq x_*$ and $z \neq \hat{x}$, the second part of Definition 2.2.1 gives

$$\|g(\hat{x}) - g(x_*) - [z, x_*; g](\hat{x} - x_*)\| \leq \|[z, x_*, \hat{x}; g](\hat{x} - z)\| \|\hat{x} - x_*\|.$$

Since $\|[z, x_*, \hat{x}; g](\hat{x} - z)\| \leq \|[z, x_*, \hat{x}; g]\| \|\hat{x} - z\|$ and $\hat{x}, z \in B_\delta(x_*)$, the last inequality together with the second inequality in (3.3) imply that

$$\|g(\hat{x}) - g(x_*) - [z, x_*; g](\hat{x} - x_*)\| \leq M \|\hat{x} - z\| \|\hat{x} - x_*\| < 2M\delta^2. \quad (3.25)$$

On the other hand, using (3.4) and (3.18), it is straightforward to show that

$$\frac{L}{2} \delta^2 + 2M\delta^2 < \left(\frac{L}{2} + 2M\right) \frac{2b}{3L + 4M} = \frac{L + 4M}{2} \frac{2b}{3L + 4M} < b.$$

Thus, combining (3.23), (3.24) and (3.25) with the last inequality yields

$$\|u(z, x_*, \hat{x})\| < b. \quad (3.26)$$

Since, in particular, $L_{f+g+F}(x_*, \cdot, z)$ is metrically regular at x_* for $0, x_* \in B_\delta[x_*]$ and (3.26) holds, applying Definition 2.4.1 with $x = x_*$, $G(\cdot) = L_{f+g+F}(x_*, \cdot, z)$, $u = -u(z, x_*, \hat{x})$ and taking into account that $\delta < a$, we conclude that

$$d(x_*, L_{f+g+F}(x_*, -u(z, x_*, \hat{x}), z)^{-1}) \leq \kappa d(-u(z, x_*, \hat{x}), L_{f+g+F}(x_*, x_*, z)). \quad (3.27)$$

Due to $0 \in L_{f+g+F}(x_*, x_*, z)$ we have $d(-u(z, x_*, \hat{x}), L_{f+g+F}(x_*, x_*, z)) \leq \|u(z, x_*, \hat{x})\|$. Moreover, considering that (3.23) and first inequalities in (3.24) and (3.25) implies

$$\|u(z, x_*, \hat{x})\| \leq \frac{L}{2} \|\hat{x} - x_*\|^2 + M \|\hat{x} - z\| \|\hat{x} - x_*\|,$$

we obtain from (3.27) that

$$d(x_*, L_{f+g+F}(x_*, -u(z, x_*, \hat{x}), z)^{-1}) \leq \kappa \left(\frac{L}{2} \|\hat{x} - x_*\|^2 + M \|\hat{x} - z\| \|\hat{x} - x_*\| \right). \quad (3.28)$$

Considering that $L_{f+g+F}(x_*, \cdot, z)$ is strongly metrically regular at x_* for 0, the mapping

$$B_b[0] \ni u \mapsto L_{f+g+F}(x_*, u, z)^{-1} \cap B_a[x_*]$$

is single-valued. Thus, as \hat{x} is a solution of (1.1), it follows from (2.3) and (3.22) that

$$0 \in f(\hat{x}) + g(\hat{x}) + F(\hat{x}) = u(z, x_*, \hat{x}) + L_{f+g+F}(x_*, \hat{x}, z),$$

which implies $-u(z, x_*, \hat{x}) \in L_{f+g+F}(x_*, \hat{x}, z)$. Therefore, $\hat{x} = L_{f+g+F}(x_*, -u(z, x_*, \hat{x}), z)^{-1}$. Hence, (3.28) implies

$$\|\hat{x} - x_*\| \leq \kappa \left(\frac{L}{2} \|\hat{x} - x_*\| + M \|\hat{x} - z\| \right) \|\hat{x} - x_*\|,$$

and, as $\|\hat{x} - x_*\| \neq 0$, we have

$$\kappa \left(\frac{L}{2} \|\hat{x} - x_*\| + M \|\hat{x} - z\| \right) \geq 1.$$

Since $\hat{x}, z \in B_\delta(x_*)$ we have $\|\hat{x} - z\| \leq \|\hat{x} - x_*\| + |x_* - z|$. The last inequality implies that $\kappa(L + 2M)\delta > 1$, or equivalently,

$$\delta > \frac{1}{\kappa(L + 2M)}. \quad (3.29)$$

On the other hand, (3.4) implies that

$$\delta < \frac{2(1 - \sqrt{2\bar{\theta}})}{\kappa \left[(3 - \sqrt{2\bar{\theta}})L + 4M \right]}, \quad 0 \leq \bar{\theta} < \frac{1}{2},$$

which in turn implies that $\delta < 1/(\kappa(L + 2M))$. This inequality contradicts (3.29). Therefore, $\|\hat{x} - x_*\| = 0$, that is $\hat{x} = x_*$, and the proof is concluded. \blacksquare

3.2.2 Application

In this section, we present a particular instance of Theorems 3.2.1 and 3.2.3.

Theorem 3.2.4 *Let $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuously differentiable function, $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuous function, $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximal monotone mapping and C be a convex and closed set in \mathbb{R}^n . Assume that $x_* \in C$, $h(x_*) + g(x_*) + F(x_*) \ni 0$, and that there exist $\alpha > 0$, $\beta > 0$, $L > 0$ and $M > 0$ such that*

$$\langle h'(x_*)p, p \rangle \geq \beta \|p\|^2; \quad \langle [z, x_*; g]q, q \rangle \geq \alpha \|q\|^2;$$

$$\|h'(x) - h'(y)\| \leq L\|x - y\|; \quad \|[x, y, z; g]\| \leq M \quad \forall p, q, x, y, z \in \mathbb{R}^n.$$

Let $(\theta_k)_{k \in \mathbb{N}} \subset [0, 1/2)$ be a sequence such that $\tilde{\theta} := \sup_k \theta_k < 1/2$ and take $\delta > 0$ such that

$$\delta < \frac{2(1 - \sqrt{2\tilde{\theta}})(\alpha + \beta)}{(3 - \sqrt{2\tilde{\theta}})L + 4M}.$$

Then x_* is the unique solution of (1.1) in $B_\delta(x_*)$. If $x_{-1}, x_0 \in C \cap B_\delta(x_*) \setminus \{x_*\}$, then every sequence $(x_k)_{k \in \mathbb{N}}$ generated by Algorithm 3.1.1 belongs to the set $B_\delta(x_*) \cap C$. Moreover, $(x_k)_{k \in \mathbb{N}}$ converges to x_* and the convergence rate is as follows:

$$\|x_* - x_{k+1}\| \leq \left[\left(1 + \sqrt{2\theta_k}\right) \frac{L\|x_* - x_k\| + 2M\|x_* - x_{k-1}\|}{2((\alpha + \beta) - (L + M)\|x_* - x_k\|)} + \sqrt{2\theta_k} \right] \|x_* - x_k\|, \quad k = 0, 1, \dots$$

As a consequence, if $\lim_{k \rightarrow +\infty} \theta_k = 0$, then $(x_k)_{k \in \mathbb{N}}$ converges to x_* superlinearly. In particular, if $\theta_k = 0$, for all $k = 0, 1, \dots$, then

$$\|x_* - x_{k+1}\| \leq \frac{(3L + 4M)}{2(\alpha + \beta)(L + 2M)} (L\|x_* - x_k\| + 2M\|x_* - x_{k-1}\|) \|x_* - x_k\|, \quad k = 0, 1, \dots$$

Proof. Applying Proposition 2.4.13 we obtain that, for all $z \in \mathbb{R}^n$, $L_{h+g+F}(x_*, \cdot, z) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is strongly metrically regular at $x_* \in \mathbb{R}^n$ for $0 \in \mathbb{R}^n$, with constants $\kappa = 1/(\alpha + \beta)$, $a = +\infty$ and $b = +\infty$. Moreover, as $x_* \in C$ satisfies $h(x_*) + g(x_*) + F(x_*) \ni 0$, we conclude that h , g and F fulfill all assumptions of Theorems 3.2.1 and 3.2.3. Therefore, the statement follows. \blacksquare

Remark 3.2.5 It is worth noting that letting $g = 0$ and $F = N_D$, where N_D denotes the normal cone, Theorem 3.2.4 merges into [18, Theorem 3].

We end this section by presenting some particular cases of constrained mixed generalized equations. We begin by presenting the so-called *Mixed Constrained Variational Inequality Problem* (MCVIP); see a particular instance for example in [15]. Let U and Ω be closed

convex sets in \mathbb{R}^n , $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuously differentiable function and $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuous function. The MCVIP is defined as

$$\text{find } x_* \in U \cap \Omega \quad \text{such that} \quad \langle f(x_*) + g(x_*), x - x_* \rangle \geq 0, \quad \forall x \in U. \quad (3.30)$$

Let $N_U : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be the normal cone mapping of a closed convex set U defined by

$$N_U(x) = \{d \in \mathbb{R}^n : \langle d, x - \bar{x} \rangle \leq 0, \forall x \in U\}.$$

Problem (3.30) can be rewritten equivalently as follows:

$$\text{find } x_* \in U \cap \Omega \quad \text{such that} \quad f(x_*) + g(x_*) + N_U(x_*) \ni 0. \quad (3.31)$$

Then, (3.30) can be seen as a special instance the constrained generalized equation (1.1). Observe that the classical mixed variational inequality problem is not equivalent to the above MCVIP, since in (3.30) the point x_* must belong to $U \cap \Omega$.

Next we describe the *Split Mixed Variational Inequality Problem* (SMVIP), which can be rewritten as special case of MCVIP; see for example [14, 36] for an extensive discussion on this problem. Let $U \subset \mathbb{R}^n$ and $\Omega \subset \mathbb{R}^m$ be nonempty, closed and convex sets, and $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear operator. Let $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $g_1 : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be continuously differentiable functions and $f_2 : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $g_2 : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be continuous functions. The SMVIP is formulated as follows: find a point $x_* \in U$ such that

$$\langle f_1(x_*) + f_2(x_*), x - x_* \rangle \geq 0, \quad \forall x \in U,$$

and such that the point $y_* = Ax_* \in \Omega$ satisfies

$$\langle g_1(y_*) + g_2(y_*), y - y_* \rangle \geq 0, \quad \forall y \in \Omega.$$

By taking $\mathbb{R}^{nm} := \mathbb{R}^n \times \mathbb{R}^m$, $D := U \times \Omega$ and $V := \{w = (x, y) \in \mathbb{R}^{nm} : Ax = y\}$ the SMVIP is equivalent to the following MCVIP:

$$\text{find } w_* \in D \cap V \quad \text{such that} \quad \langle f(w_*) + g(w_*), w - w_* \rangle \geq 0, \quad \forall w \in D, \quad (3.32)$$

where $w = (x, y)$ and the functions $f, g : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{nm}$ are defined by $f(x, y) := (f_1(x), g_1(y))$ and $g(x, y) := (f_2(x), g_2(y))$. See [14, Lemma 5.1] for a particular instance of this problem. Therefore, from (3.31) and (3.32), SMVIP is equivalent to the following constrained generalized equation: find $w_* \in D \cap V$ such that $f(w_*) + g(w_*) + N_D(w_*) \ni 0$, where $f, g : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{nm}$ are defined by $f(x, y) := (f_1(x), g_1(y))$ and $g(x, y) := (f_2(x), g_2(y))$, $D := U \times \Omega$ and $V := \{w = (x, y) \in \mathbb{R}^{nm} : Ax = y\}$.

Remark 3.2.6 It is worth noting that SMVIP is quite general and includes several problems as special cases. For instance, *Split Minimization Problem* and *Common Solutions to Variational Inequalities Problem*; see for example [1, 14, 15, 49].

Chapter 4

Broyden quasi-Newton secant-type method with feasible inexact projection for solving constrained mixed generalized equations

In this chapter, we are going to propose a version of the Secant-type method for solving constrained generalized equations, i.e., for solving the problem (1.1). Basically, the proposed method can be seen as a combination of the Secant-type method studied in the last chapter and Broyden-type quasi-Newton method. The local convergence of general quasi-Newton approaches is established under a bounded deterioration of the update matrix and Lipschitz continuity hypotheses.

Broyden quasi-Newton secant-type method to generalized equations, which has its origin in the works of N. H. Josephy, see [42], have the following formulation: for the current iterates $x_{k-1}, x_k \in \mathbb{R}^n$, the next iterate x_{k+1} is computed as a point satisfying the following inclusion

$$f(x_k) + g(x_k) + (B_k + [x_{k-1}, x_k; g])(y_k - x_k) + F(y_k) \ni 0., \quad k = 0, 1, \dots, \quad (4.1)$$

where $\{B_k\}$ is a sequence of bounded linear mappings between \mathbb{R}^n satisfying the classical Broyden update rule. and $[x_{k-1}, x_k; g]$ is the first order divided difference of the function $g : \Omega \rightarrow \mathbb{R}^n$ on the points x, y . The specific choice of how B_k is constructed dictates the particular quasi-Newton method being used, such as Broyden, BFGS, SR1, and others. Typically, the initial mapping B_0 is chosen to be a good approximation of $f'(\bar{x})$. Note that at each iteration, a partially linearized inclusion at the current iterate needs to be solved. The method described by (4.1) serves as a framework for various iterative

procedures in numerical nonlinear programming. For example, when $F \equiv \{0\}$, this method reduces to the standard Newton's method for solving systems of nonlinear equations. If F represents the product of the negative orthant in \mathbb{R}^s with the origin in \mathbb{R}^{m-s} , i.e., $F = \mathbb{R}_-^s \times \{0\}^{m-s}$, then (4.1) corresponds to Newton's method for solving systems of nonlinear equalities and inequalities [17]. Conversely, if the problem (1.1) with $C = \Omega$ represents the Karush-Kuhn-Tucker optimality conditions for a nonlinear programming problem, then (4.1) describes the well-known sequential quadratic programming method [24, p. 384]; see also [21, 39].

For unconstrained problems, where $C = \mathbb{R}^n$ or $C = \Omega$, we can simply set $x_{k+1} = y_k$. The iterative scheme described in (4.1) was initially proposed by [16] for the specific case where $F \equiv 0$ and $C = \mathbb{R}^n$. This was followed by a local convergence analysis and, later, a semi-local analysis by [37]. [32] extended this method to the case where $C = \mathbb{R}^n$ and $F \neq 0$, providing a local convergence analysis, with a semi-local analysis appearing subsequently in [38]. Other significant variants of this method have been explored, for instance, in [40, 50]. Since the iterate y_k in (4.1) can sometimes be infeasible with respect to the constraint set C , a strategy is needed to ensure the feasibility of the sequence generated. To address this, we employ the conditional gradient (CondG) method, also known as the Frank-Wolfe method. This method, one of the earliest first-order techniques, was originally proposed for quadratic problems over a polyhedron (see [31]) and has since been generalized to other problems [19, 25–27, 46]. The CondG method has garnered significant attention in recent years [35, 45, 47], motivating its use in conjunction with the Broyden quasi-Newton secant-type method for solving problem (1.1). From a theoretical perspective, we establish that the Broyden quasi-Newton secant-type method (4.1) is well-defined and locally convergent, with convergence rate that is q-linear. The key assumptions required for these results include metric regularity, Lipschitz continuity of the derivative f' , Approximation of $f'(x_k)$ - sequence $\{B_k\}$ of bounded linear mappings between \mathbb{R}^n satisfying the classical Broyden update rule. and boundedness of the second-order divided difference of g . For a detailed discussion on the metric regularity, see [23].

4.1 The Inexact Broyden quasi-Newton secant-type Algorithm

The conceptual secant method with a feasible inexact projection, named the *inexact Broyden quasi-Newton secant-type* method for solving the mixed generalized equation (1.1) and with $x_{-1}, x_0 \in C$, $(\theta_k)_{k \in \mathbb{N}} \subset [0, +\infty)$ and $B_0 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ as the input data, is a method that utilizes the CondG procedure to handle constraints and employs a Broyden update to approximate the Jacobian of f . The algorithm is formally described below.

Algorithm 4.1.1 Inexact Broyden quasi-Newton secant-type

Step 0 Initialize with $x_{-1}, x_0 \in C$, a sequence $(\theta_k)_{k \in \mathbb{N}} \subset [0, +\infty)$, and $B_0 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. Set $k = 0$.

Step 1 If $0 \in f(x_k) + g(x_k) + F(x_k)$, **stop**; a solution has been found. Otherwise, compute $y_k \in \mathbb{R}^n$ satisfying

$$0 \in f(x_k) + g(x_k) + (B_k + [x_{k-1}, x_k; g])(y_k - x_k) + F(y_k). \quad (4.2)$$

Update

$$s_k := y_k - x_k, \quad z_k := f(y_k) - f(x_k),$$

and compute

$$B_{k+1} := B_k + \frac{(z_k - B_k s_k) s_k^\top}{\|s_k\|^2}. \quad (4.3)$$

Step 2 If $y_k \in C$, set $x_{k+1} = y_k$. Otherwise, compute $x_{k+1} \in P_C(y_k, x_k, \theta_k)$ using the CondG procedure.

Step 3 Set $k \leftarrow k + 1$ and return to **Step 1**.

In **Step 1** of Algorithm 4.1.1, we first check if the current iterate x_k is a solution of (1.1). If x_k is not a solution of (1.1), we compute y_k satisfying the inclusion (3.1). Note that $y_k \neq x_k$. Since y_k in **Step 1** may not belong to the constraint set C , the Secant-InexP method uses Algorithm 2.5.4 to compute x_{k+1} in C . Because $y_k \neq x_k$, we also have $x_k \neq x_{k+1}$. Hence, Algorithm 4.1.1 either generates a finite sequence with the last iterate being a solution of (1.1) or an infinite sequence. Therefore, from now on, we assume that Algorithm 4.1.1 generates an infinite sequence $(x_k)_{k \in \mathbb{N}} \subset C$ with $x_k \neq x_{k+1}$ for all $k = 0, 1, \dots$. Finally, it is worth noting that x_{k+1} obtained in (3.2) is an approximate solution for the projection subproblem $\min_{p \in C} \|p - y_k\|^2/2$, satisfying the stopping criterion $\langle y_k - x_{k+1}, p - x_{k+1} \rangle \leq \theta_k \|y_k - x_k\|^2$ for any $p \in C$.

Remark 4.1.2 Whenever $g = 0$ and $F \equiv \{0\}$ the above algorithm merges into the one proposed in [33]. On the other hand, if $C = \mathbb{R}^n$ then Algorithm 4.1.1 merges into the finite dimensional version of the secant method proposed in [32].

To proceed with the analysis of Algorithm 4.1.1, we assume throughout our presentation that the following assumptions hold:

(A0) The derivative f' is Lipschitz continuous with constant $L > 0$ on Ω , i.e.,

$$\|f'(x) - f'(y)\| \leq L \|x - y\|, \quad \forall x, y \in \Omega. \quad (4.4)$$

(A1) There exists $M > 0$ such that

$$\|[x, y, z; g]\| \leq M, \quad \forall x, y, z \in \Omega, \quad (4.5)$$

where $[x, y, z; g]$ denotes the second-order divided difference of g .

(A2) The mapping $f + g + F$ is metrically regular at x_* for 0 with constants $\kappa > 0$, $a > 0$, and $b > 0$.

We begin by establishing a technical result that will be useful in our analysis.

Lemma 4.1.3 *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuously differentiable function such that f' satisfies (A0). Let $r > 0$ and $\epsilon > 0$ be such that $2Lr < \epsilon$. Then, for all $x, y \in B_r(x_*)$, the following inequalities hold:*

$$\|f(y) - f(x) - f'(x)(y - x)\| \leq \frac{L}{2}\|y - x\|^2, \quad (4.6)$$

and

$$\|f(y) - f(x) - f'(x_*)(y - x)\| \leq \epsilon\|y - x\|. \quad (4.7)$$

Proof. Since f' is Lipschitz continuous with constant L , by applying Taylor's theorem with integral the equality (4.6) follows. For prove (4.7), we have

$$\begin{aligned} \|f(y) - f(x) - f'(x_*)(y - x)\| &\leq \|f(y) - f(x) - f'(x)(y - x)\| + \|[f'(x) - f'(x_*)](y - x)\| \\ &\leq ((L/2)\|y - x\| + L\|x - x_*\|)\|y - x\| \\ &\leq \epsilon\|y - x\|, \end{aligned}$$

where the last inequality holds because $\|y - x\| \leq 2r$ and $2Lr < \epsilon$. ■

4.2 The Convergence Theorem

With these preliminary results in hand, we now state and prove our main convergence theorem for Algorithm 4.1.1.

Theorem 4.2.1 *Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $C \subseteq \Omega$ be a closed convex set, and let $f : \Omega \rightarrow \mathbb{R}^n$ be a continuously differentiable function. Let $g : \Omega \rightarrow \mathbb{R}^n$ be a continuous function (not necessarily differentiable), and let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping with closed graph. Suppose that x_* is a solution to the generalized equation*

$$f(x) + g(x) + F(x) \ni 0.$$

Assume that conditions **(A0)**, **(A1)**, and **(A2)** hold. Take the initial approximation B_0 satisfying

$$\delta := \|B_0 - f'(x_*)\| < \frac{1}{2\kappa}. \quad (4.8)$$

Let $(\theta_k)_{k \in \mathbb{N}} \subset [0, +\infty)$ be a sequence with $\hat{\theta} = \sup_{k \in \mathbb{N}} \theta_k < \frac{1}{2}$. Then, there exists $\tau > 0$ such that, for any initial points $x_{-1}, x_0 \in B_\tau[x_*] \cap C$, there exists a sequence $(x_k)_{k \in \mathbb{N}} \subset B_\tau[x_*] \cap C$ generated by Algorithm 2 with starting points x_{-1}, x_0 converges q -linearly to x_* . If, in addition, $f + g + F$ is strongly metrically regular at x_* for 0, then the sequence $(x_k)_{k \in \mathbb{N}}$ starting from x_{-1}, x_0 is unique and converges q -linearly to x_* .

Proof. Let $L > 0$ satisfying **(A0)**, $M > 0$ satisfying **(A1)** and $\kappa \geq 0$, $a > 0$, and $b > 0$ satisfying **(A2)**. Choose $\kappa' > \kappa$ and $\gamma > 0$ such that

$$\frac{\kappa'\delta}{1 - \kappa\delta} < 1, \quad \frac{\kappa'\delta}{1 - \kappa\delta} < \gamma < 1.$$

To proceed with our analysis it is convenient to note the function $(0, +\infty) \ni t \mapsto \kappa't/(1 - \kappa t)$ is strictly increasing. Then, we can choose a sufficiently small $\epsilon > 0$ that satisfies the following two conditions

$$\epsilon < \delta, \quad \frac{\kappa'}{1 - \kappa(\delta + \epsilon)}(\delta + \epsilon) < \gamma. \quad (4.9)$$

In addition, considering that $\delta > 0$ satisfies (4.8), without loss of generality, we can also assume that $a > 0$ and $\tau > 0$ satisfy the following three conditions

$$a < \frac{\epsilon}{2L}, \quad \kappa \left(\delta + \frac{(L/2)a}{1 - \gamma} + M\tau \right) < 1, \quad (4.10)$$

$$\frac{\kappa'}{1 - \kappa \left(\delta + \frac{(L/2)a}{1 - \gamma} + M\tau \right)} \left(\delta + \epsilon + \frac{(L/2)a}{1 - \gamma} + M\tau \right) < \gamma. \quad (4.11)$$

It is also convenient to define the constant

$$\mu := \delta + \frac{(L/2)a}{1 - \gamma} + M\tau. \quad (4.12)$$

Choose $\alpha > 0$ and $\beta > 0$ satisfying (2.21) in Theorem 2.4.23, i.e., satisfying following inequalities

$$\epsilon + 3\alpha M \leq \mu, \quad 0 < \alpha \leq \frac{a}{2}, \quad \mu\alpha + 2\beta \leq b, \quad 2\kappa'\beta \leq \alpha(1 - \kappa\mu). \quad (4.13)$$

Finally, taking into account that $\beta > 0$, we can again make $\tau > 0$ sufficiently small such that

$$0 < \tau < \frac{1}{M} \left[\frac{(\gamma - \sqrt{2\theta_0})(1 - \kappa\mu)}{(1 + \sqrt{2\theta_0})\kappa'} - (\epsilon + \delta) \right], \quad \left(\epsilon + \delta + \frac{(L/2)a}{1 - \gamma} + M\tau \right) \tau \leq \beta. \quad (4.14)$$

Before continuing with our analysis let us note that the first two inequalities in (4.14) implies that

$$\frac{\kappa'}{1 - \kappa\mu} \left(\epsilon + \delta + \frac{(L/2)a}{1 - \gamma} + M\tau \right) \left(1 + \sqrt{2\theta_0} \right) + \sqrt{2\theta_0} \leq \gamma. \quad (4.15)$$

Consider

$$h(x) = f(x_*) + g(x_*) + [f'(x_*) + [z, x_*; g]](x - x_*) - f(x) - g(x). \quad (4.16)$$

Note that $h(x_*) = 0$ and, for all $x, x', z \in B_\alpha[x_*]$, we have

$$\begin{aligned} \|h(x) - h(x')\| &= \|f(x') + g(x') - f(x) - g(x) + (f'(x_*) + [z, x_*; g])(x - x')\| \\ &\leq \|f(x') - f(x) - f'(x_*)(x' - x)\| + \|g(x') - g(x) - [z, x_*; g](x' - x)\|. \end{aligned} \quad (4.17)$$

To bound the first term in (4.17) we note that (4.7) in Lemma 4.1.3 implies that

$$\|f(x') - f(x) - f'(x_*)(x' - x)\| \leq \epsilon \|x' - x\|. \quad (4.18)$$

Now, we are going to bound the second term (4.17). For that, first note that

$$\begin{aligned} \|g(x') - g(x) - [z, x_*; g](x' - x)\| &= \|[x, x'; g](x' - x) - [z, x_*; g](x' - x)\| \\ &= \|([x, x'; g] - [z, x_*; g])(x' - x)\| \\ &= \|([x, x'; g] - [z, x; g] - ([z, x_*; g] - [z, x; g]))(x' - x)\|. \end{aligned}$$

Considering that $[z, x_*; g] - [z, x; g] = [z, x_*; g] - [x, z; g]$, the last inequality implies that

$$\|g(x') - g(x) - [z, x_*; g](x' - x)\| = \|([x, x'; g] - [z, x; g] - ([z, x_*; g] - [x, z; g]))(x' - x)\|.$$

Thus, using Definition 2.2.1 and **A1**, and taking into account that $x, x', z \in B_\alpha[x_*]$, we conclude that

$$\begin{aligned} \|g(x') - g(x) - [z, x_*; g](x' - x)\| &= \|([z, x, x'; g](x' - z) - [x, z, x_*; g](x_* - x))(x' - x)\| \\ &\leq (\|[z, x, x'; g]\| \|x' - z\| + \|[x, z, x_*; g]\| \|x_* - x\|) \|x' - x\| \\ &\leq 3\alpha M \|x' - x\|. \end{aligned}$$

Hence, using (4.18) and the last inequality we obtain from (4.17) and the first inequality in (4.13) that

$$\|h(x) - h(x')\| \leq (\epsilon + 3\alpha M) \|x' - x\| \leq \mu \|x' - x\|, \quad \forall x, x' \in B_\alpha[x_*]. \quad (4.19)$$

To simplify the notation define the mapping

$$G(x) = f(x_*) + g(x_*) + [f'(x_*) + [z, x_*; g]](x - x_*) + F(x). \quad (4.20)$$

It follows from (4.16) and (4.20) that $G = H + h$, where $H = f + g + F$. It follows from **(A2)** that H is metrically regular at x_* for 0 with constants $\kappa > 0$, $a > 0$, and $b > 0$. Considering that $h(x_*) = 0$, using (4.13) and (4.19) we can apply Theorem 2.4.23 with \bar{x} we conclude that G is metrically regular at x_* for 0 with constant κ . Therefore, there exist positive constants a and b such that

$$d(x, G^{-1}(y)) \leq \kappa d(y, G(x)), \quad \forall x \in B_a[x_*], \quad \forall y \in B_b[0].$$

Let $O := B_\tau[x_*]$, and choose any points $x_0, z \in O \cap C$ with $x_0 \neq z$, $x_0 \neq x_*$ and $z \neq x_*$. Define

$$h_0(x) := f(x_0) + g(x_0) + (B_0 + [z, x_0; g])(x - x_0) - f(x_*) - g(x_*) - (f'(x_*) + [z, x_*; g])(x - x_*), \quad (4.21)$$

and $w_0 := f(x_0) + g(x_0) - f(x_*) - g(x_*) + (B_0 + [z, x_0; g])(x_* - x_0)$. Then, we have

$$w_0 = h_0(x_*).$$

We proceed to estimate $\|h_0(x_*)\|$. For that, using (4.21), some manipulations show that

$$\begin{aligned} \|h_0(x_*)\| &\leq \|f(x_0) - f(x_*) - f'(x_*)(x_0 - x_*)\| + \|(B_0 - f'(x_*))(x_* - x_0)\| + \\ &\quad \|g(x_0) - g(x_*) + [z, x_0; g](x_* - x_0)\|. \end{aligned} \quad (4.22)$$

Let us first bound the right hand side of the last inequality. For the first term, using (4.7) in Lemma 4.1.3, we have

$$\|f(x_0) - f(x_*) - f'(x_*)(x_0 - x_*)\| \leq \epsilon \|x_0 - x_*\|.$$

We bound the second term, by using definition of δ in (4.8), which gives

$$\|(B_0 - f'(x_*))(x_* - x_0)\| \leq \delta \|x_0 - x_*\|.$$

For the term involving g , we use the definition of the second-order divided difference and assumption **(A1)**. Thus, we have

$$\begin{aligned} \|g(x_0) - g(x_*) + [z, x_0; g](x_* - x_0)\| &= \|[x_0, x_*; g](x_* - x_0) - [z, x_0; g](x_* - x_0)\| \\ &= \|[z, x_0, x_*; g](x_* - z)(x_* - x_0)\| \\ &\leq M \|x_* - z\| \|x_0 - x_*\| \leq M\tau \|x_0 - x_*\|. \end{aligned}$$

Combining these estimates and taking into account that $x_0 \in O \cap C$, the inequality (4.22) yields

$$\|h_0(x_*)\| \leq (\epsilon + \delta + M\tau) \|x_0 - x_*\| \leq (\epsilon + \delta + M\tau) \tau \leq \left(\epsilon + \delta + \frac{(L/2)a}{1 - \gamma} + M\tau \right) \tau. \quad (4.23)$$

Thus, it follows from (4.14) that

$$\|h_0(x_*)\| \leq \beta.$$

Hence $w_0 = h_0(x_*) \in B_\beta[0]$. Similarly, give $x, x' \in B_\alpha(x_*)$ and considering $z = x_{-1}$, by using definition (4.21), some algebraic manipulations imply that

$$\begin{aligned} \|h_0(x) - h_0(x')\| &= \|(B_0 + [z, x_0; g])(x - x') - (f'(x_*) + [z, x_*; g])(x - x')\| \\ &\leq \|(B_0 - f'(x_*))(x - x')\| + \|[z, x_0; g] - [z, x_*; g]\| \|x - x'\|. \end{aligned}$$

Thus, after some calculations and using Definition 2.2.1 we conclude that

$$\begin{aligned} \|h_0(x) - h_0(x')\| &\leq \|(B_0 - f'(x_*))(x - x')\| + \|[z, x_0; g] - [z, x_*; g]\| \|x - x'\| \\ &= \|(B_0 - f'(x_*))(x - x')\| + \|[z, x_0; g] - [x_*, z; g] + [x_*, z; g] - [z, x_*; g]\| \|x - x'\| \\ &= \|(B_0 - f'(x_*))(x - x')\| + \|[x_*, z, x_0; g](x_0 - x_*) + [z, x_*, z; g](z - z)\| \|x - x'\| \\ &\leq \|(B_0 - f'(x_*))(x - x')\| + \|[x_*, z, x_0; g]\| \|x_0 - x_*\| \|x - x'\|. \end{aligned}$$

Since $[z, x_*, z; g](z - z) = 0$, using properties of the norm we obtain from the last inequality that

$$\|h_0(x) - h_0(x')\| \leq \|(B_0 - f'(x_*))(x - x')\| + \|[x_*, z, x_0; g]\| \|x_0 - x_*\| \|x - x'\|.$$

Finally, using (4.8), assumption **(A1)** and definition of μ in (4.12), we conclude that

$$\|h_0(x) - h_0(x')\| \leq \delta \|x - x'\| + M\tau \|x - x'\| \leq \mu \|x - x'\|, \quad \forall x, x' \in B_\alpha(x_*).$$

By Theorem 2.4.23, with $\bar{x} = x_*$, $h = h_0$, $H = G$, $y = w_0$ and $x = x_*$, there exists $y_0 \in (h + G)^{-1}(0)$ satisfying (4.2), for $k = 0$, and

$$\|y_0 - x_*\| \leq \frac{\kappa'}{1 - \kappa\mu} \|h_0(x_*)\|.$$

It follows from the first inequality in (4.23) and the last inequality, by taking into account (4.11) and (4.12), that

$$\|y_0 - x_*\| \leq \frac{\kappa'}{1 - \kappa\mu} (\epsilon + \delta + M\tau) \|x_0 - x_*\| \leq \gamma \|x_0 - x_*\|. \quad (4.24)$$

Therefore, considering that $\gamma < 1$, we conclude that $y_0 \in O$. Next, applying Lemma 2.5.3 with $u = x_1$, $y = y_0$, $x = x_0$, $\tilde{y} = x_*$, $\tilde{x} = x_*$, and $\theta = \theta_0$, we obtain that

$$\begin{aligned} \|x_1 - x_*\| &\leq \|y_0 - x_*\| + \sqrt{2\theta_0} \|y_0 - x_0\| \\ &\leq \|y_0 - x_*\| + \sqrt{2\theta_0} (\|y_0 - x_*\| + \|x_* - x_0\|) \\ &\leq (1 + \sqrt{2\theta_0}) \|y_0 - x_*\| + \sqrt{2\theta_0} \|x_0 - x_*\|. \end{aligned}$$

Using inequality (4.15) and (4.24) we conclude that

$$\|x_1 - x_*\| \leq \left[\frac{\kappa'}{1 - \kappa\mu} (\epsilon + \delta + M\tau)(1 + \sqrt{2\theta_0}) + \sqrt{2\theta_0} \right] \|x_0 - x_*\| \leq \gamma \|x_0 - x_*\|.$$

Therefore, since $\gamma < 1$, we obtain that $x_1 \in O \cap C$. By induction, suppose that for some $n \geq 1$, we have points $x_k \in O \cap C$, $y_k \in O$, satisfying

$$\|y_{k-1} - x_*\| \leq \gamma \|x_{k-1} - x_*\|, \quad \|x_k - x_*\| \leq \gamma \|x_{k-1} - x_*\|, \quad k = 1, \dots, n.$$

Thus, it follows from Proposition 2.6.5 together with the two last inequalities, that

$$\begin{aligned} \|B_n - f'(x_*)\| &\leq \|B_0 - f'(x_*)\| + \frac{L}{2} \left(\|x_0 - x_*\| + \sum_{k=1}^n (\|x_k - x_*\|) + \sum_{k=1}^n \|y_{k-1} - x_*\| \right) \\ &\leq \delta + \frac{L}{2} \|x_0 - x_*\| + L \sum_{k=1}^n \gamma^k \|x_0 - x_*\| \leq \delta + L \sum_{k=0}^n \gamma^k \|x_0 - x_*\| \\ &\leq \delta + \frac{L \|x_0 - x_*\|}{1 - \gamma}. \end{aligned}$$

Hence, using (4.13), we conclude that

$$\|B_n - f'(x_*)\| \leq \delta + \frac{(L/2)a}{1 - \gamma}.$$

Consider $z = x_{-1}$ and define

$$h_n(x) := f(x_n) + g(x_n) + (B_n + [z, x_n; g])(x - x_n) - f(x_*) - g(x_*) - (f'(x_*) + [z, x_*; g])(x - x_*), \quad (4.25)$$

$$w_n := f(x_n) + g(x_n) - f(x_*) - g(x_*) + (B_n + [z, x_n; g])(x_* - x_n) = h_n(x_*).$$

Using similar estimates as before, to obtain that (4.23), we can also obtain that

$$\|w_n\| \leq \left(\epsilon + \delta + \frac{(L/2)a}{1 - \gamma} + M\tau \right) \|x_n - x_*\| \leq \beta.$$

Hence, $\|h_n(x_*)\| = \|w_n\| \leq \beta$. In addition, also we can obtain that

$$\begin{aligned} \|h_n(x) - h_n(x')\| &= \|(B_n + [z, x_n; g])(x - x') - (f'(x_*) + [z, x_*; g])(x - x')\| \\ &\leq \|(B_n - f'(x_*))(x - x')\| + \|[z, x_n; g] - [z, x_*; g]\| \|x - x'\| \\ &\leq \left(\delta + \frac{(L/2)a}{1 - \gamma} \right) \|x - x'\| + M\tau \|x - x'\| \\ &= \mu \|x - x'\|, \quad \forall x, x' \in B_\alpha[x_*]. \end{aligned}$$

By applying Theorem 2.4.23, with $\bar{x} = x_*$, $h = h_n$, $H = G$, $y = w_n$ and $x = x_*$, there exists $y_n \in (h_n + G)^{-1}(0)$ satisfying (4.2) for $k = n$, and

$$\|y_n - x_*\| \leq \frac{\kappa'}{1 - \kappa\mu} \|w_n\| \leq \frac{\kappa'}{1 - \kappa\mu} \left(\epsilon + \delta + \frac{(L/2)a}{1 - \gamma} + M\tau \right) \|x_n - x_*\| \leq \gamma \|x_n - x_*\|. \quad (4.26)$$

Again, considering that $\gamma < 1$, we conclude that $y_n \in O$. Next, applying Lemma 2.5.3 with $u = x_{n+1}$, $y = y_n$, $x = x_n$, $\tilde{y} = x_*$, $\tilde{x} = x_*$, and $\theta = \theta_0$, and the inequality (4.15), we obtain

$$\begin{aligned} \|x_{n+1} - x_*\| &\leq \|y_n - x_*\| + \sqrt{2\theta_0} \|y_n - x_n\| \\ &\leq \|y_n - x_*\| + \sqrt{2\theta_0} (\|y_n - x_*\| + \|x_* - x_n\|) \\ &\leq (1 + \sqrt{2\theta_0}) \|y_n - x_*\| + \sqrt{2\theta_0} \|x_n - x_*\|. \end{aligned}$$

Using inequality (4.15) and (4.26) we conclude that

$$\|x_{n+1} - x_*\| \leq \left(\frac{\kappa'}{1 - \kappa\mu} (\epsilon + \delta + \frac{(L/2)a}{1 - \gamma} + M\tau) (1 + \sqrt{2\theta_0}) + \sqrt{2\theta_0} \right) \|x_n - x_*\| \leq \gamma \|x_n - x_*\|.$$

Hence, the induction is completed. Moreover, taking into account that $\gamma < 1$, we conclude that $x_{n+1} \in O \cap C$. Therefore,

$$\|x_k - x_*\| \leq \gamma \|x_{k-1} - x_*\|, \quad k = 1, 2, \dots$$

Consequently, the sequence $(x_k)_{k \in \mathbb{N}} \subset B_\tau[x_*] \cap C$ and converges q-linearly to x_* .

If $H = f + g + F$ is strongly metrically regular at x_* for 0, then the mapping G is strongly metrically regular, and by the uniqueness part of Theorem 2.4.23, the sequence $(x_k)_{k \in \mathbb{N}} \subset B_\tau[x_*] \cap C$ is unique and converges q-linearly to x_* . This completes the proof. \blacksquare

The proposed Broyden quasi-Newton secant-type method extends existing techniques for solving constrained generalized equations by incorporating elements of the Conditional Gradient method. The method achieves local Q-linear convergence under standard assumptions, such as Lipschitz continuity and metric regularity. Future work could explore extensions to global convergence and performance improvements for large-scale applications, as well as adaptations for specific problem classes, such as variational inequalities and complementarity problems.

Chapter 5

Conclusion

5.1 Summary of Contributions

This thesis has embarked on the development and analysis of novel secant-type methods tailored for solving constrained generalized equations. The primary contributions can be succinctly summarized as follows:

- **Hybrid Secant-Type Method:** We introduced a new secant-type iterative scheme that effectively combines the traditional secant method with the conditional gradient (Frank-Wolfe) method. This hybrid approach addresses the inherent challenges posed by constraints in generalized equations, ensuring that all iterates remain feasible within the convex set C .
- **Quasi-Newton Enhancements:** Building upon the foundational work of Broyden and inspired by Andreani et al. [5], we integrated a Broyden-type quasi-Newton update into the secant-type framework. This integration enhances the approximation of the Jacobian matrix, thereby improving the convergence properties of the algorithm without incurring the computational burden of exact derivative evaluations.
- **Theoretical Convergence Analysis:** Utilizing the contraction mapping principle, we established rigorous convergence results for the proposed methods. Under the assumptions of Lipschitz continuity of the gradient and metric regularity of the underlying mappings, we demonstrated that the generated sequences are well-defined and converge locally to a solution with either linear or superlinear rates.
- **Inexact Projection Mechanism:** Recognizing that intermediate iterates may violate feasibility constraints, we employed the conditional gradient method to perform inexact

projections onto the convex set C . This strategy balances computational efficiency with the necessity of maintaining feasibility, ensuring the practicality of the algorithm in large-scale or complex constrained environments.

5.2 Implications and Significance

The advancements presented in this thesis hold significant implications for both theoretical research and practical applications in numerical optimization and equilibrium analysis:

- **Enhanced Computational Efficiency:** By approximating the Jacobian matrix iteratively through quasi-Newton updates, the proposed methods reduce the computational overhead associated with exact derivative computations. This efficiency is particularly beneficial in high-dimensional or large-scale problems where derivative evaluations are costly.
- **Robust Handling of Constraints:** The integration of the conditional gradient method ensures that all iterates remain within the feasible region defined by the convex set C . This robust handling of constraints expands the applicability of secant-type methods to a broader class of problems, including those with complex or non-linear constraints.
- **Theoretical Rigor and Practical Reliability:** The convergence analyses grounded in metric regularity and Lipschitz continuity provide a solid theoretical foundation for the proposed algorithms. This rigor translates to practical reliability, assuring practitioners of the methods' effectiveness under well-defined conditions.
- **Versatility Across Problem Domains:** Generalized equations encapsulate a wide array of problem types, from systems of nonlinear equations to variational inequalities and complementarity problems. The methods developed herein are thus versatile tools that can be adapted to diverse applications in engineering, economics, machine learning, and beyond.

5.3 Future Research Directions

While this thesis has made significant strides in advancing secant-type methods for constrained generalized equations, several avenues remain ripe for exploration:

- **Extension to Non-Convex Constraints:** The current framework assumes that the constraint set C is convex. Extending these methods to handle non-convex constraints would broaden their applicability, albeit with additional theoretical challenges.
- **Incorporation of Stochastic Elements:** In many real-world applications, data or model parameters may be subject to uncertainty. Developing stochastic variants of the proposed methods could enhance their robustness and applicability in such settings.
- **Adaptive Parameter Selection:** The convergence and performance of the algorithms are influenced by parameters such as the Lipschitz constants and metric regularity parameters. Investigating adaptive strategies for selecting these parameters dynamically could improve convergence rates and stability.
- **Parallel and Distributed Implementations:** Leveraging modern computational architectures, such as parallel and distributed systems, to implement the proposed methods can significantly reduce computation times, especially for large-scale problems.
- **Application to Machine Learning and Data Science:** Given the prevalence of constrained optimization problems in machine learning, exploring the application of these methods in training models, especially those with complex constraints, could yield impactful results.
- **Enhanced Quasi-Newton Updates:** Further refining the quasi-Newton update rules, possibly incorporating second-order information or leveraging limited-memory techniques, could enhance the efficiency and convergence properties of the algorithms.

5.4 Final Remarks

The proposed Broyden quasi-Newton secant-type method extends existing techniques for solving constrained generalized equations by incorporating elements of the Conditional Gradient method. The method achieves local Q-linear convergence under standard assumptions, such as Lipschitz continuity and metric regularity. Future work could explore extensions to global convergence and performance improvements for large-scale applications, as well as adaptations for specific problem classes, such as variational inequalities and complementarity problems. This thesis has contributed to the evolving landscape of numerical methods for solving constrained generalized equations by introducing and analyzing hybrid secant-type methods augmented with quasi-Newton updates and inexact projection mechanisms. The convergence guarantees and empirical validations underscore

the potential of these methods to address complex, constrained problems efficiently and reliably. As optimization and equilibrium modeling continue to underpin advances in various scientific and engineering disciplines, the methodologies developed herein are poised to play a pivotal role in facilitating these endeavors.

Bibliography

- [1] M. Abbas, M. AlShahrani, Q. H. Ansari, O. S. Iyiola, and Y. Shehu. Iterative methods for solving proximal split minimization problems. *Numer. Algorithms*, 78(1):193–215, May 2018.
- [2] S. Adly and H. Ngai. Quasi-newton methods for solving nonsmooth equations: generalized dennis–moré theorem and broyden’s update. *J. Convex Anal.*, 25(4):1075–1104, 2018.
- [3] S. Adly, H. Van Ngai, and V. V. Nguyen. Newton’s method for solving generalized equations: Kantorovich’s and Smale’s approaches. *J. Math. Anal. Appl.*, 439(1):396–418, 2016.
- [4] R. Andreani, R. M. Carvalho, L. D. Secchin, and G. N. Silva. Convergence of quasi-Newton methods for solving constrained generalized equations. *ESAIM Control Optim. Calc. Var.*, 28:Paper No. 32, 24, 2022.
- [5] R. Andreani, R. M. Carvalho, L. D. Secchin, and G. N. Silva. Convergence of quasi-newton methods for solving constrained generalized equations. *ESAIM: Control, Optimisation and Calculus of Variations*, 28:32, 2022.
- [6] F. J. Aragón Artacho, A. Belyakov, A. L. Dontchev, and M. López. Local convergence of quasi-Newton methods under metric regularity. *Comput. Optim. Appl.*, 58(1):225–247, 2014.
- [7] F. J. Aragón Artacho, A. L. Dontchev, M. Gaydu, M. H. Geoffroy, and V. M. Veliov. Metric regularity of Newton’s iteration. *SIAM J. Control Optim.*, 49(2):339–362, 2011.
- [8] A. Auslender and M. Teboulle. *Asymptotic cones and functions in optimization and variational inequalities*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.
- [9] D. Azé and C. Chou. On a newton type iterative method for solving inclusions. *Mathematics of operations research*, 20(4):790–800, 1995.

- [10] A. Beck and M. Teboulle. A conditional gradient method with linear rate of convergence for solving convex linear systems. *Math. Methods Oper. Res.*, 2004.
- [11] S. Bellavia and B. Morini. Subspace trust-region methods for large bound-constrained nonlinear equations. *SIAM J. Numer. Anal.*, 44(4):1535–1555, 2006.
- [12] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific Optimization and Computation Series. Athena Scientific, Belmont, MA, second edition, 1999.
- [13] J. F. Bonnans. Local analysis of Newton-type methods for variational inequalities and nonlinear programming. *Appl. Math. Optim.*, 29(2):161–186, 1994.
- [14] Y. Censor, A. Gibali, and S. Reich. Algorithms for the split variational inequality problem. *Numer. Algorithms*, 59(2):301–323, Feb 2012.
- [15] Y. Censor, A. Gibali, S. Reich, and S. Sabach. Common solutions to variational inequalities. *Set-Valued Var. Anal.*, 20(2):229–247, 2012.
- [16] E. Căţinaş. On some iterative methods for solving nonlinear equations. *Rev. Anal. Numér. Théor. Approx.*, 23(1):47–53, 1994.
- [17] J. W. Daniel. Newton’s method for nonlinear inequalities. *Numer. Math.*, 21:381–387, 1973.
- [18] F. R. de Oliveira, O. P. Ferreira, and G. N. Silva. Newton’s method with feasible inexact projections for solving constrained generalized equations. *Comput. Optim. Appl.*, 72(1):159–177, 2019.
- [19] V. F. Demyanov and A. M. Rubinov. *Approximate Methods in Optimization Problems*. American Elsevier Publishing Company, New York, 1970.
- [20] A. L. Dontchev. Local analysis of a Newton-type method based on partial linearization. In *The mathematics of numerical analysis (Park City, UT, 1995)*, volume 32 of *Lectures in Appl. Math.*, pages 295–306. Amer. Math. Soc., Providence, RI, 1996.
- [21] A. L. Dontchev. Uniform convergence of the Newton method for Aubin continuous maps. *Serdica Math. J.*, 22(3):283–296, 1996.
- [22] A. L. Dontchev. On existence of solutions of parametrized generalized equations. *Set-Valued Var. Anal.*, 29(3):735–744, 2021.
- [23] A. L. Dontchev and R. T. Rockafellar. Convergence of inexact Newton methods for generalized equations. *Math. Program.*, 139(1-2, Ser. B):115–137, 2013.

- [24] A. L. Dontchev and R. T. Rockafellar. *Implicit functions and solution mappings*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2014. A view from variational analysis.
- [25] J. C. Dunn. Rates of convergence for conditional gradient algorithms near singular and nonsingular extremals. *SIAM J. Control Optim.*, 17:187–211, 1979.
- [26] J. C. Dunn. Convergence rates for conditional gradient sequences generated by implicit step length rules. *SIAM J. Control Optim.*, 18(5):473–487, 1980.
- [27] J. C. Dunn and S. Harshbarger. Conditional gradient algorithms with open loop step size rules. *J. Math. Anal. Appl.*, 62(2):432–444, 1978.
- [28] O. P. Ferreira. A robust semi-local convergence analysis of Newton’s method for cone inclusion problems in Banach spaces under affine invariant majorant condition. *J. Comput. Appl. Math.*, 279:318–335, 2015.
- [29] O. P. Ferreira and G. N. Silva. Kantorovich’s theorem on Newton’s method for solving strongly regular generalized equation. *SIAM J. Optim.*, 27(2):910–926, 2017.
- [30] O. P. Ferreira and G. N. Silva. Local convergence analysis of Newton’s method for solving strongly regular generalized equations. *J. Math. Anal. Appl.*, 458(1):481–496, 2018.
- [31] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Nav. Res. Log.*, pages 95–110, 1956.
- [32] M. H. Geoffroy and A. Piétrus. Local convergence of some iterative methods for generalized equations. *J. Math. Anal. Appl.*, 290(2):497–505, 2004.
- [33] M. L. N. Gonçalves and J. G. Melo. A Newton conditional gradient method for constrained nonlinear systems. *J. Comput. Appl. Math.*, 311:473–483, 2017.
- [34] M. L. N. Gonçalves and F. R. Oliveira. An inexact Newton-like conditional gradient method for constrained nonlinear systems. *preprint arXiv:1705.07684*, 2017.
- [35] Z. Harchaoui, A. Juditsky, and A. Nemirovski. Conditional gradient algorithms for norm-regularized smooth convex optimization. *Math. Program.*, 152(1-2, Ser. A):75–112, 2015.
- [36] H. He, C. Ling, and H.-K. Xu. A relaxed projection method for split variational inequalities. *J. Optim. Theory Appl.*, 166(1):213–233, 2015.

- [37] M. A. Hernández and M. J. Rubio. Semilocal convergence of the secant method under mild convergence conditions of differentiability. *Comput. Math. Appl.*, 44(3-4):277–285, 2002.
- [38] S. Hilout and A. Piétrus. A semilocal convergence of a secant-type method for solving generalized equations. *Positivity*, 10(4):693–700, 2006.
- [39] A. F. Izmailov and M. V. Solodov. Inexact Josephy-Newton framework for generalized equations and its applications to local analysis of Newtonian methods for constrained optimization. *Comput. Optim. Appl.*, 46(2):347–368, 2010.
- [40] C. Jean-Alexis and A. Piétrus. On the convergence of some methods for variational inclusions. *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM*, 102(2):355–361, 2008.
- [41] N. H. Josephy. Newton’s method for generalized equations. Technical report, DTIC Document, 1979.
- [42] N. H. Josephy. *Newton’s method for generalized equations and the PIES energy model*. PhD thesis, Department of Industrial Engineering, University of Wisconsin–Madison, 1979.
- [43] D. Klatte and B. Kummer. Approximations and generalized newton methods. *Mathematical Programming*, 168:673–716, 2018.
- [44] W. La Cruz. A projected derivative-free algorithm for nonlinear equations with convex constraints. *Optim. Methods Softw.*, 29(1):24–41, 2014.
- [45] G. Lan and Y. Zhou. Conditional gradient sliding for convex optimization. *SIAM J. Optim.*, 26(2):1379–1409, 2016.
- [46] E. S. Levitin and B. T. Polyak. Constrained minimization methods. *USSR Comput. Math. Math. Phys.*, 6(5):1–50, 1966.
- [47] R. Luss and M. Teboulle. Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint. *SIAM Rev.*, 55(1):65–98, 2013.
- [48] L. Marini, B. Morini, and M. Porcelli. Quasi-Newton methods for constrained nonlinear systems: complexity analysis and applications. *Available on <http://www.optimization-online.org/>*, 2017.
- [49] A. Moudafi. Split monotone variational inclusions. *J. Optim. Theory Appl.*, 150(2):275–283, Aug 2011.

- [50] M. H. Rashid, J. H. Wang, and C. Li. Convergence analysis of a method for variational inclusions. *Appl. Anal.*, 91(10):1943–1956, 2012.
- [51] M. H. Rashid, S. H. Yu, C. Li, and S. Y. Wu. Convergence analysis of the Gauss-Newton-type method for Lipschitz-like mappings. *J. Optim. Theory Appl.*, 158(1):216–233, 2013.
- [52] S. M. Robinson. Extension of Newton’s method to nonlinear functions with values in a cone. *Numer. Math.*, 19:341–347, 1972.
- [53] S. M. Robinson. Normed convex processes. *Trans. Amer. Math. Soc.*, 174:127–140, 1972.
- [54] S. M. Robinson. Generalized equations and their solutions, Part I: Basic theory. *Math. Program. Stud.*, 10:128–141, 1979.
- [55] S. M. Robinson. Strongly regular generalized equations. *Math. Oper. Res.*, 5(1):43–62, 1980.
- [56] S. M. Robinson. Generalized equations. In *Mathematical programming: the state of the art (Bonn, 1982)*, pages 346–367. Springer, Berlin, 1983.