NEWTON’S METHOD FOR SOLVING STRONGLY REGULAR GENERALIZED EQUATIONS

DOCTORAL THESIS BY
Gilson do Nascimento Silva

SUPERVISED BY
Prof. Dr. Orison Pereira Ferreira

FUNDED BY
CAPES

IME - INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE FEDERAL DE GOIÁS
GOIÂNIA, GOIÁS, BRAZIL
MARÇO 2017
TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizou 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 nº 9610/98, 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.

1. Identificação do material bibliográfico:  [ ] Dissertação    [x] Tese

2. Identificação da Tese ou Dissertação

Nome completo do autor: Gilson do Nascimento Silva

Título do trabalho: Newton's method for solving strongly regular generalized equations

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [ x ] SIM    [ ] NÃO¹

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Gilson do Nascimento Silva
Assinatura do (a) autor (a)

Data: 21/03/2017

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.
NEWTON’S METHOD FOR SOLVING STRONGLY REGULAR GENERALIZED EQUATIONS

Tese apresentada ao Programa de Pós-Graduação 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

Goiânia
2017
Silva, Gilson do Nascimento
NEWTON'S METHOD FOR SOLVING STRONGLY REGULAR
GENERALIZED EQUATIONS [manuscrito] / Gilson do Nascimento
Silva. - 2017.
66 f.

Orientador: Prof. Dr. Orizon Pereira Ferreira.
Tese (Doutorado) - Universidade Federal de Goiás, Instituto de
Matemática e Estatística (IME), Programa de Pós-Graduação em
Matemática, Goiânia, 2017.
Bibliografia.

1. Generalized equations. 2. Newton's method. 3. Majoran
condition. 4. Strong regularity. I. Ferreira, Orizon Pereira, orient. II.
Título.

CDU 51
ATA DA REUNIÃO DA BANCA EXAMINADORA DA DEFESA DE TESE DE GILSON DO NASCIMENTO SILVA – Aos treze dias do mês de março do ano de dois mil e dezessete (13/03/2017), às 10:00 horas, reuniram-se os componentes da Banca Examinadora: Prof. Orizont Pereira Ferreira - Orientador, Prof. Max Leandro Nobre Gonçalves, Prof. Jefferson Divino Gonçalves de Melo, Prof. Paulo José da Silva e Silva, e Profa. Elizabeth Wegner Karas, sob a presidência do primeiro, e em sessão pública realizada na sala de multimídia do LEMAT do Instituto de Matemática e Estatística, procederem a avaliação da defesa de tese intitulada: “NEWTON’S METHOD FOR SOLVING STRONGLY REGULAR GENERALIZED EQUATIONS”, em nível de Doutorado, área de concentração em Optimização, de autoria de Gilson do Nascimento Silva, discente do Programa de Pós-Graduação em Matemática da Universidade Federal de Goiás. A sessão foi aberta pelo Presidente da Banca, Prof. Orizont Pereira Ferreira que fez a apresentação formal dos membros da Banca. A seguir, a palavra foi concedida ao autor da tese que, em 45 minutos apresentou a apresentação de seu trabalho. Terminada a apresentação, cada membro da Banca arguiu o examinando, tendo-se adotado o sistema de diaólogos sequenciados. Terminada a fase de arguição, procedeu-se a avaliação da defesa. Tendo-se em vista o que consta na Resolução nº. 1068 do Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC), que regulamenta o Programa de Pós-Graduação em Matemática e procedidas as correções recomendadas, a tese foi APROVADA por unanimidade, considerando-se integralmente cumprido este requisito para fins de obtenção do título de DOCTOR EM MATEMÁTICA, na área de concentração em Optimização, pela Universidade Federal de Goiás. A conclusão do curso dar-se-á quando da entrega na secretaria do PPGM da versão definitiva da tese, com as devidas correções supervisionadas e aprovadas pelo orientador. Cumpridas as formalidades de pauta, às 12:20 horas a presidência da mesa encerrou esta sessão de defesa de tese e para constar eu, Ulisses José Gabry, secretário do PPGM, lavrei a presente Ata que, depois de lida e aprovada, será assinada pelos membros da Banca Examinadora em quatro vias de igual teor.

Prof. Dr. Orizont Pereira Ferreira  
Presidente - IME/UFG

Prof. Dr. Max Leandro Nobre Gonçalves  
Membro – IME/UFG

Prof. Dr. Jefferson Divino Gonçalves de Melo  
Membro – IME/UFG

Prof. Dr. Paulo José da Silva e Silva  
Membro – IMECC/UNICAMP

Prof. Dra. Elizabeth Wegner Karas  
Membro – DMAT/UFPR
Dedicado a:

Meus pais
Meus irmãos
Meus amigos.
Agradecimentos

Primeiramente, agradeço a Deus por estar sempre presente e por ter me dado forças para não desistir no meio do caminho.

Agradeço ao Professor Orizon Pereira Ferreira, meu orientador, por ter aceitado me orientar, pela enorme paciência e por ajudar a desenvolver todo este trabalho.

Agradeço aos meus familiares e amigos pelo apoio e torcida, principalmente meus pais e meus irmãos que sempre acreditaram em mim.

Agradeço aos professores do IME/UFG que contribuíram na minha formação: Luís Roman, Ronaldo Garcia, Levi Rosa, Jefferson Melo, Glaydston Bento e Orizon Ferreira. Também agradeço a todos os professores e colegas do grupo de otimização deste instituto.

Agradeço aos Professores Paulo Silva, Elizabeth Karas, Max Leandro Gonçalves e Jefferson Melo por participarem da comissão examinadora.

Agradeço aos amigos que conviveram comigo durante esse doutorado: Valdinês Júnior, Yuri Rafael, Lucas Vidal, Edvaldo Elias e Luama Dalria. Agradeço também a todos meus amigos de Goiânia pela amizade e excelente hospitalidade.

Agradeço à CAPES pelo apoio financeiro.
Abstract

We consider Newton’s method for solving a generalized equation of the form

$$f(x) + F(x) \ni 0,$$

where $f : \Omega \to \mathbb{Y}$ is continuously differentiable, $\mathbb{X}$ and $\mathbb{Y}$ are Banach spaces, $\Omega \subseteq \mathbb{X}$ is open and $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ has nonempty closed graph. Assuming strong regularity of the equation and that the starting point satisfies Kantorovich’s conditions, we show that the method is quadratically convergent to a solution, which is unique in a suitable neighborhood of the starting point. In addition, a local convergence analysis of this method is presented. Moreover, using convex optimization techniques introduced by S. M. Robinson (Numer. Math., Vol. 19, 1972, pp. 341-347), we prove a robust convergence theorem for inexact Newton’s method for solving nonlinear inclusion problems in Banach space, i.e., when $F(x) = -C$ and $C$ is a closed convex set. Our analysis, which is based on Kantorovich’s majorant technique, enables us to obtain convergence results under Lipschitz, Smale’s and Nesterov-Nemirovskii’s self-concordant conditions.

**Keywords:** Generalized equation, Newton’s method, strong regularity, majorant condition, semi-local convergence, inclusion problems, inexact Newton method.
Resumo

Nós consideraremos o método de Newton para resolver uma equação generalizada da forma

\[ f(x) + F(x) \geq 0, \]

onde \( f : \Omega \to Y \) é continuamente diferenciável, \( X \) e \( Y \) são espaços de Banach, \( \Omega \subseteq X \) é aberto e \( F : X \rightrightarrows Y \) tem gráfico fechado não-vazio. Supondo regularidade forte da equação e que o ponto inicial satisfaz as hipóteses de Kantorovich, mostraremos que o método é quadraticamente convergente para uma solução, a qual é única em uma vizinhança do ponto inicial. Uma análise de convergência local deste método também é apresentada. Além disso, usando técnicas de otimização convexa introduzida por S. M. Robinson (Numer. Math., Vol. 19, 1972, pp. 341-347), provaremos um robusto teorema de convergência para o método de Newton inexato para resolver problemas de inclusão não-linear em espaços de Banach, i.e., quando \( F(x) = -C \) e \( C \) é um conjunto convexo fechado. Nossa análise, a qual é baseada na técnica majorante de Kantorovich, nos permite obter resultados de convergência sob as condições Lipschitz, Smale e Nesterov-Nemirovskii auto-concordante.

\textit{Palavras-chave:} Equação generalizada, método de Newton, regularidade forte, condição majorante, convergência semi-local, problemas de inclusão, método de Newton inexato.
Basic notation and terminology

\(B(x, \delta):\) the open ball at \(x\) with radius \(\delta > 0,\)
\(B[x, \delta]:\) the closed ball at \(x\) with radius \(\delta \geq 0,\)
\(\mathbb{X}, \mathbb{Y}:\) Banach spaces,
\(\mathbb{X}^*:\) the dual of \(\mathbb{X},\)
\(L(\mathbb{X}, \mathbb{Y}):\) the space consisting of all continuous linear mappings \(A : \mathbb{X} \to \mathbb{Y},\)
\(F : \mathbb{X} \rightrightarrows \mathbb{Y}:\) a set-valued mapping with nonempty closed graph,
\(N_C:\) the normal cone mapping of a convex set \(C,\)
\(T_{x_0}:\) the convex process given by \(T_{x_0}d := f'(x_0)d - C, \quad d \in \mathbb{X},\)
Contents

1 Preliminaries 9
  1.1 Strong Regularity .................................................. 9
  1.2 Majorant condition .................................................. 13

2 Local convergence analysis of Newton’s method for solving strongly regular generalized equations 18
  2.1 Local analysis of Newton’s method ................................. 18
  2.2 Preliminary results .................................................... 19
    2.2.1 Proof of Theorem 2.1.1 ........................................ 24
  2.3 Particular cases ...................................................... 26
    2.3.1 Under Hölder-type condition .................................. 26
    2.3.2 Under Smale’s-type condition .................................. 27
    2.3.3 Under Nesterov-Nemirovskii’s condition .................... 29

3 Kantorovich’s theorem on Newton’s method for solving strongly regular generalized equation 30
  3.1 Kantorovich’s theorem .............................................. 30
    3.1.1 Basic results .................................................... 31
    3.1.2 Convergence analysis .......................................... 36
  3.2 Special cases ....................................................... 39

4 Inexact Newton method for nonlinear functions with values in a cone 42
  4.1 Convex Process ....................................................... 43
  4.2 Inexact Newton method ............................................. 44
4.2.1 Preliminary results .............................................. 46
4.3 Convergence analysis of the inexact Newton Method ......................... 49
  4.3.1 The inexact Newton iteration .................................. 49
  4.3.2 Convergence analysis ........................................... 52
4.4 Special cases ................................................................ 56
  4.4.1 Under affine invariant Lipschitz condition ....................... 56
  4.4.2 Under affine invariant Smale’s condition ....................... 57

5 Final remarks ............................................................. 59
Introduction

It is well-known that the classical Newton method and its generalizations are among the most effective methods for numerically solving the nonlinear equation

\[ f(x) = 0, \]  

for a given function \( f : \mathbb{X} \rightarrow \mathbb{Y} \). Its effectiveness is due to the quadratic rate of convergence under suitable assumptions on the problem data and the choice of the initial point. The classical convergence results state that Newton’s method is only locally convergent. More precisely, if the function \( f \) is sufficiently smooth and its Jacobian \( f'(x^*) \) is nonsingular at the solution \( x^* \), then, by choosing an initial point \( x_0 \) in a neighborhood of this solution, the sequence generated by Newton’s method converges and the rate of convergence is at least quadratic. For a detailed discussion about the classical Newton method see, for instance, the book by Ortega and Rheinboldt [46].

In this work, we consider Newton’s method for solving the generalized equation

\[ f(x) + F(x) \ni 0, \]  

where \( f : \Omega \rightarrow \mathbb{Y} \) is a continuously differentiable function, \( \mathbb{X} \) and \( \mathbb{Y} \) are Banach spaces, \( \Omega \subseteq \mathbb{X} \) is an open set and \( F : \mathbb{X} \rightrightarrows \mathbb{Y} \) is a set-valued mapping with nonempty closed graph. As is well-known, (2) is an abstract model for a wide range of problems in mathematical programming. See, for instance, [3, 4, 13, 16–18, 35, 36, 51] as part of a whole. In the case \( F \equiv \{0\} \), (2) becomes (1). If \( \mathbb{X} = \mathbb{R}^n \), \( \mathbb{Y} = \mathbb{R}^{p+q} \) and \( F = \mathbb{R}^p \times \{0\} \) is the product of the non-positive orthant in \( \mathbb{R}^p \) with the origin at \( \mathbb{R}^q \), then (2) describes a system of equalities and inequalities. A particular case of problem (2) is when \( F = -C \), where \( C \subset \mathbb{Y} \) is a nonempty closed convex cone. Thus, problem (2) becomes

\[ f(x) \in C. \]  

If \( \mathbb{Y} \) is the dual \( \mathbb{X}^* \) of \( \mathbb{X} \) and \( F \) is the normal cone mapping \( N_C \) of a closed convex set \( C \subset \mathbb{X} \), then the inclusion (2) is the variational inequality problem

\[ u \in C, \quad \langle f(u), v - u \rangle \geq 0 \quad \forall \, v \in C. \]
This, in particular, includes the Karush-Kuhn-Tucker (KKT) optimality conditions via the following construction: Consider the problem

$$\min \phi(x) \quad \text{subject to} \quad h(x) = 0, \quad g(x) \leq 0,$$

where the objective function $\phi : \mathbb{R}^n \to \mathbb{R}$ and the constraints mappings $h : \mathbb{R}^n \to \mathbb{R}^p$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable. The stationary points of problem (5) and the associated Lagrange multipliers are characterized by the KKT optimality system

$$\nabla_x (x, \lambda, \mu) = 0, \quad h(x) = 0, \quad \mu \geq 0, \quad g(x) \leq 0, \quad \langle \mu, g(x) \rangle = 0,$$

where $L : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}$ is the Lagrangian of problem (5)

$$L(x, \lambda, \mu) = \phi(x) + \langle \lambda, h(x) \rangle + \langle \mu, g(x) \rangle.$$ 

Then the KKT system (6) is a particular instance of the generalized equation with the mapping $f : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m$ given by

$$f(x, \lambda, \mu) = (\nabla_x (x, \lambda, \mu), -h(x), -g(x))$$

and with

$$F(x) = N_C(x), \quad C = \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m_+.$$ 

Additional comments on problem (2) can be found in [3, 4, 13, 16–19, 36, 51].

Newton method for solving the generalized equation (2) for an initial point $x_0$ is defined as follows:

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

Studies concerned with this method include [3, 4, 12–14, 18, 36]; see also [16, Section 6C], where an interesting discussion about iterative methods for solving generalized equations is presented. When $F \equiv \{0\}$, (9) becomes the standard Newton method for solving $f(x) = 0,

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) = 0, \quad k = 0, 1, \ldots.$$  

If $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ and $F = \mathbb{R}_-^s \times \{0\}^{m-s}$, then (9) becomes Newton method for solving a system of equalities and inequalities; see [9]. In particular, if (2) represents the Karush-Kuhn-Tucker optimality conditions for a nonlinear programming problem, then (9) describes a well-known sequential quadratic programming method; see, for example, [16, p. 334]. By contrast, if $F = -C$, in order to solve (3), S. M. Robinson in [51], based on the idea of convex process introduced by Rockafellar [55], proposed the following Newton-type method:

$$x_{k+1} = x_k + d_k, \quad d_k \in \arg\min_{d \in X} \{\|d\| : f(x_k) + f'(x_k)d \in C\}, \quad k = 0, 1, \ldots.$$  

5
We remark that if $C \equiv \{0\}$, then the Newton-type method (11) becomes the standard
Newton method (10). For more details and extensions of (11), see, for instance, [8,19,24,
40–42,51]. If $\mathcal{Y}$ is the dual $\mathcal{X}^*$ of $\mathcal{X}$ and $F$ is the normal cone mapping $N_C$ of a closed convex
set $C \subset \mathcal{X}$, then (9) is the version of Newton method for solving variational inequalities;
see [12,36].

The classical local analysis of Newton’s method for solving $f(x) = 0$ requires invertibility of
$f'$ at a solution $\bar{x}$ to ensure uniqueness of the solution to the linearization $f(\hat{x})+f'(\hat{x})(x-\hat{x}) =
0$ at $\hat{x}$, for all $\hat{x}$ in a neighborhood of $\bar{x}$. This, in turn, ensures that the method is well-defined.
Therefore, for the local as well as the semi-local analysis of Newton’s method (9), a similar
concept will be needed. L. V. Kantorovich in [37] (see also [38,48]) was the first to prove
a convergence result for Newton's method for solving $f(x) = 0$, by imposing conditions on
the starting point $x_0$ rather than on the unknown solution. Using suitable conditions on $x_0$,
that is, $f'(x_0)^{-1}$ exists and $\|f'(x_0)^{-1}f(x_0)\|$ is small enough, Kantorovich established
that the sequence generated by Newton’s method is well defined. Moreover, he proved
that it converges quadratically to a solution and that this solution is unique in a suitable
neighborhood. The proof of convergence is based on the technique of majorization, which
involves bounding Newton’s sequence by a scalar sequence. This technique has been used and
extended in [8,19,24,27,32,49,61,63]. In his Ph.D. thesis, N. H. Josephy [36] studied Newton’s
method for solving $f(x) + N_C(x) \ni 0$, where $f: \Omega \to \mathbb{R}^m$ is continuously differentiable,
$\Omega \subset \mathbb{R}^n$ is an open set, and $C \subset \mathbb{R}^m$ is a convex set. To validate the definition of the
sequence generated by the method, the strong regularity property of $f + N_C$, a concept
introduced by Robinson in [53], was used. If $\mathcal{X} = \mathcal{Y}$ and $N_C = \{0\}$, then strong regularity at
$x$ is equivalent to $f'(x)^{-1}$ being a continuous linear operator. An important case is when (2)
represents Karush–Kuhn–Tucker systems for the standard nonlinear programming problem
with a strict local minimizer; see [16, p. 232]. In this case, strong regularity of this system is
equivalent to the linear independence of the gradients of the active constraints and a strong
form of the second-order sufficient optimality condition: one has

$$\langle x', \nabla_{xx}^2 L(x, \lambda, \mu)x' \rangle > 0, \text{ for all } x' \neq 0 \text{ in the subspace }$$

\[ M = \{ x' : \langle x', \nabla_x h(x) \rangle = 0, \; \langle x', \nabla_x g(x) \rangle = 0 \}, \]

for details see [15, Theorem 6].

In general, algorithm (11) may fail to converge and may even fail to be well defined. To
ensure that the method is well defined and converges to a solution of a given nonlinear
inclusion, S. M. Robinson in [51] made two important assumptions:

**H1.** There exists $x_0 \in \mathcal{X}$ such that \( \text{rge } T_{x_0} = \mathcal{Y} \), where $T_{x_0}: \mathcal{X} \rightrightarrows \mathcal{Y}$ is the convex process
given by

$$T_{x_0}d := f'(x_0)d - C, \quad d \in \mathcal{X}$$

6
and \( \text{rge } T_{x_0} = \{ y \in \mathbb{Y} : y \in T_{x_0}(x) \text{ for some } x \in \mathbb{X} \} \); see [16] for details.

**H2.** \( f' \) is Lipschitz continuous with modulus \( L \), i.e., \( \| f'(x) - f'(y) \| \leq L \| x - y \| \), for all \( x, y \in \mathbb{X} \).

Under these assumptions, it was proved in [51] that the sequence \( \{ x_k \} \) generated by (11) is well defined and converges to an \( x_* \) satisfying \( f(x_*) \in C \), provided that the following convergence criterion is satisfied:

\[
\| x_1 - x_0 \| \leq \frac{1}{2L\| T_{x_0}^{-1} \|}.
\]

The first affine invariant version of this result was presented by Li and Ng in [41]. In [42] Li and Ng introduced the weak-Robinson condition for convex processes and presented an extension of the results of [41] under an \( L \)-average Lipschitz condition. Two special cases were considered, namely, the convergence result for the method under a Lipschitz condition and under Smale’s condition. In [19], under an affine majorant condition, a robust analysis of this method was established. As in [41], the analysis assuming a Lipschitz condition and Smale’s condition was obtained as a special case; see also [10, 23].

As is well-known, the hypothesis usually used to obtain quadratic convergence of (9) is the Lipschitz continuity of \( f' \) in a neighborhood of an initial point; see [8, 12, 13, 19, 24, 27, 36]. Indeed, maintaining control of \( f' \) is an important consideration in the analysis of (9). However, certain authors have been concerned with the issue of convergence analysis of (9) for \( F \equiv \{0\} \) by relaxing the Lipschitz continuity of \( f' \); see, for example, [27, 29–31, 57, 58, 61, 62]. The conditions in these works are equivalent to the one introduced in [61]. The advantage of using a majorant condition, relaxing the Lipschitz continuity of \( f' \), lies in the fact that it allows us to unify several convergence results pertaining to Newton’s method; see [27, 61]. In this thesis, we rephrase the majorant condition introduced in [27], in order to study the properties of Newton’s method (9). The thesis is organized as follows:

**Chapter 1:** Therein, we recall some notations and results that are used throughout the thesis. In particular, we present a version of the Banach Perturbation Lemma involving a general set-valued mapping. Moreover, we prove a corollary of this result which will play an important role in subsequent chapters. The results of this chapter are from the paper [26].

**Chapter 2:** This is devoted to the local convergence analysis of Newton’s method for solving the generalized equation (2). Accordingly, we assume that \( f + F \) is strongly regular at \( \bar{x} \) for 0 with modulus \( \lambda > 0 \). Moreover, the Lipschitz continuity of \( f' \) is relaxed. It is worth mentioning that the recent approach for analyzing local convergence of Newton’s method and its variants for solving generalized equations, involves the contraction mapping principle for set-valued mappings; see [4, 13, 14] and [16, Theorem 6C.6]. In contrast, our approach is...
based on the Banach Perturbation Lemma obtained by S. M. Robinson in [53, Theorem 2.4]. In this sense, our approach is related to the techniques used in [7,12,36] for studying Newton’s method for generalized equations. The results of this chapter are from the preprint [25].

Chapter 3: This is devoted to the study of Kantorovich’s theorem on Newton’s method for solving the generalized equation (2). In Section 3.1, the main result is stated and some examples of functions satisfying the majorant condition are presented. Section 3.1.1 clarifies the relationship between the majorant function and the function defining the generalized equation. In Section 3.1.2, the main result is proved. In Section 3.2, the analysis of this method under a Lipschitz condition, Smale’s condition, and Nesterov–Nemirovskii’s self-concordant conditions is provided as a special case. The results of this chapter are from the paper [26].

Chapter 4: Therein, an inexact Newton’s method for solving the nonlinear inclusion (3) is stated and analyzed. In some sense, our method is a particular instance of [18]. However, the analysis presented in [18] is local, i.e., a solution is assumed to exist, whereas in our analysis, we do not assume existence of solution. In fact, our aim is to prove a robust Kantorovich’s theorem for solving (3), under assumption $H1$ and an affine invariant majorant condition generalizing $H2$. In particular, we prove the existence of solution for (3). Moreover, the analysis presented shows that the robust analysis of the inexact Newton’s method for solving nonlinear inclusion problems under affine Lipschitz-like and affine Smale’s conditions, can be obtained as a special case of the general theory. Furthermore, for the degenerate cone, where the nonlinear inclusion becomes a nonlinear equation, our analysis retrieves the classical results on semi-local analysis of inexact Newton’s method; see [28]. The first works on this subject include [44, 56]. To our knowledge, this is the first time that the inexact Newton method for solving cone inclusion problems with a relative error tolerance is analyzed. The results of this chapter are from the preprint [24].

Chapter 5: Therein, final remarks and future work are presented.
Chapter 1

Preliminaries

The following notations and results are used throughout this thesis. Let $\mathbb{X}, \mathbb{Y}$ be Banach spaces, the open and closed balls at $x$ with radius $\delta \geq 0$ are denoted, respectively, by $B(x, \delta) = \{y \in X : \|x - y\| < \delta\}$ and $B[x, \delta] = \{y \in X : \|x - y\| \leq \delta\}$. We denote by $L(\mathbb{X}, \mathbb{Y})$ the space consisting of all continuous linear mappings $A : \mathbb{X} \rightarrow \mathbb{Y}$ and the norm of $A$ by $\|A\| := \sup \{\|Ax\| : \|x\| \leq 1\}$. Let $\Omega \subseteq \mathbb{X}$ be an open set and $f : \Omega \rightarrow \mathbb{Y}$ be Fréchet differentiable at all $x \in \Omega$. The Fréchet derivative of $f$ at $x$ is the linear mapping $f'(x) : \mathbb{X} \rightarrow \mathbb{Y}$, which is continuous. The graph of the set-valued mapping $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ is the set $gph F := \{(x, y) \in \mathbb{X} \times \mathbb{Y} : y \in F(x)\}$. The domain and the range of $F$ are, respectively, the sets $\text{dom } F = \{x \in \mathbb{X} : F(x) \neq \emptyset\}$ and $\text{rge } F = \{y \in \mathbb{Y} : y \in F(x) \text{ for some } x\}$. The inverse of $F$ is the set-valued mapping $F^{-1} : \mathbb{Y} \rightrightarrows \mathbb{X}$ defined by $F^{-1}(y) = \{x \in \mathbb{X} : y \in F(x)\}$. The partial linearization of $f + F$ at $x \in \Omega$ is the set-valued mapping $L_f(x, \cdot) : \Omega \rightrightarrows \mathbb{Y}$ defined by

$$L_f(x, y) := f(x) + f'(x)(y - x) + F(y).$$

(1.1)

1.1 Strong Regularity

An important element in the analysis of Newton’s method for solving the equation $f(x) = 0$, is the behavior of the inverse $f'(x)^{-1}$ for $x$ in a neighborhood of a solution $\bar{x}$. The analogous element for the generalized equation (2) is the behavior of the inverse mapping $L_f(x, \cdot)^{-1}$ for $x$ in a neighborhood of a solution $\bar{x}$. It is worth point out that N. H. Josephy in [36] was the first to consider Newton’s method for solving the generalized equation $f(x) + N_C(x) \ni 0$, where $N_C$ is the normal cone of a convex set $C \subset \mathbb{R}^n$, by defining the Newton iteration as $L_f(x_k, x_{k+1})^{-1} \ni 0$ for $k = 0, 1, \ldots$, which is equivalent to (9), to the particular case $F = N_C$. N. H. Josephy in [36], for analyzing Newton’s method, employed the important concept of strong regularity defined by S. M. Robinson [53], which assure “good behavior”
of $L_f(x,\cdot)^{-1}$ for $x$ in a neighborhood of a solution $\bar{x}$. Here we adopt the following definition due to S. M. Robinson; see [53].

**Definition 1.1.1** Let $\Omega \subset \mathbb{X}$ be open and nonempty. The mapping $T : \Omega \rightrightarrows \mathbb{Y}$ is called strongly regular at $x$ for $y$ with modulus $\lambda > 0$, when $y \in T(x)$ and there exist $r_x > 0$, $r_y > 0$ such that $B(x, r_x) \subset \Omega$, the mapping $z \mapsto T^{-1}(z) \cap B(x, r_x)$ is single-valued from $B(y, r_y)$ to $B(x, r_x)$ and Lipschitzian on $B(y, r_y)$ with modulus $\lambda$, i.e., $\|T^{-1}(u) \cap B(x, r_x) - T^{-1}(v) \cap B(x, r_x)\| \leq \lambda\|u - v\|$, for all $u, v \in B(y, r_y)$.

Since $z \mapsto T^{-1}(z) \cap B(x, r_x)$ in Definition 1.1.1 is single-valued, for the sake of simplicity, we have used the notation $w = T^{-1}(z) \cap B(x, r_x)$ instead of $\{w\} := T^{-1}(z) \cap B(x, r_x)$. Hereafter, we use this simplified notation. For a detailed discussion on Definition 1.1.1; see [16,17,53]. The next result is a type of implicit function theorem for generalized equations satisfying the strongly regular condition, its proof is an immediate consequence of [16, Theorem 5F.4] on page 294; see also [53, Theorem 2.1].

**Theorem 1.1.2** Let $\mathbb{X}$, $\mathbb{Y}$ and $\mathbb{Z}$ be Banach spaces, $G : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a set-valued mapping and $g : \mathbb{Z} \times \mathbb{X} \to \mathbb{Y}$ be a continuous function, having partial Fréchet derivative with respect to the second variable $D_x g$ on $\mathbb{Z} \times \mathbb{X}$, which is also continuous. Let $\bar{p} \in \mathbb{Z}$ and suppose that $\bar{x}$ solves the generalized equation

$$g(\bar{p}, x) + G(x) \ni 0.$$ 

Assume that the mapping $g(\bar{p}, \cdot) + G$ is strongly regular at $\bar{x}$ for $0$, with associated Lipschitz modulus $\lambda$. Then, for any $\epsilon > 0$ there exist neighborhoods $U_{\epsilon}$ of $\bar{x}$ and $V_{\epsilon}$ of $\bar{p}$ and a single-valued mapping $s : V_{\epsilon} \to U_{\epsilon}$ such that for any $p \in V_\epsilon$, $s(p)$ is the unique solution in $U_\epsilon$ of the inclusion $g(p, x) + G(x) \ni 0$, and $s(\bar{p}) = \bar{x}$. Moreover, there holds

$$\|s(p') - s(p)\| \leq (\lambda + \epsilon)\|g(p', s(p)) - g(p, s(p))\|, \quad \forall p, p' \in V_{\epsilon}.$$

**Proof.** Since $\mathbb{X}$, $\mathbb{Y}$ are Banach spaces, then they have a shift-invariant metric. Denoting $h : \mathbb{X} \to \mathbb{Y}$ by $h(x) = g(\bar{p}, x) + D_x g(\bar{p}, \bar{x})(x - \bar{x})$, thus $h$ is a strict estimator of $g$ with respect to $x$ uniformly in $p$ at $(\bar{p}, \bar{x})$ with constant $\mu = 0$, see page 37 of [16]. Furthermore, since the mapping $g(\bar{p}, \cdot) + G$ is strongly regular at $\bar{x}$ for $0$, with associated Lipschitz constant $\lambda$, thus $h + G$ is strongly metrically regular at $\bar{x}$ for $0$, with associated Lipschitz constant $\lambda$ such that $\mu \lambda = 0 < 1$; see definition on page 179 of [16]. Therefore, the proof is an immediate consequence of [16, Theorem 5F.4] on page 294.

Indeed, the first version of the Theorem 1.1.2 was proven by S. M. Robinson; see [53, Theorem 2.1], to the particular case $F = N_C$, where $N_C$ is the normal cone of a convex set $C \subset \mathbb{X}$. As an application, a version of the Banach Perturbation Lemma involving the
normal cone was obtained; see [53, Theorem 2.4]. N. H. Josephy in [36] used this version of Banach Perturbation Lemma, see [36, Corollary 1], for proving that the Newton iteration

\[ f(x_k) + f'(x_k)(x_{k+1} - x_k) + N_C(x_{k+1}) \geq 0, \quad k = 0, 1, \ldots, \]

where \( N_C \) is the normal cone of a convex set \( C \subset \mathbb{R}^n \), is quadratically convergent for a solution of the following particular generalized equation \( f(x) + N_C(x) \geq 0 \). In the next lemma we apply Theorem 1.1.2 to obtain a version of the Banach Perturbation Lemma involving a general set-valued mapping. The proof of this result is similar to the correspondent one [36, Corollary 1], we include it here just for sake of completeness.

**Lemma 1.1.3** Let \( \mathbb{X}, \mathbb{Y} \) be Banach spaces, \( a_0 \) be a point of \( \mathbb{Y} \), \( F : \mathbb{X} \Rightarrow \mathbb{Y} \) be a set-valued mapping and \( A_0 : \mathbb{X} \rightarrow \mathbb{Y} \) be a bounded linear mapping. Suppose that \( \bar{x} \in \mathbb{X} \) and \( 0 \in A_0\bar{x} + a_0 + F(\bar{x}) \). Assume that \( A_0 + a_0 + F \) is strongly regular at \( \bar{x} \) for 0 with modulus \( \lambda > 0 \). Then, there exist \( r_0 > 0, r_{A_0} > 0, r_{a_0} > 0 \), and \( r_0 > 0 \) such that, for any \( A \in B(A_0, r_{A_0}) \subset L(\mathbb{X}, \mathbb{Y}) \) and \( a \in B(a_0, r_{a_0}) \subset \mathbb{Y} \) letting \( T(A, a, \cdot) : B(\bar{x}, r_0) \Rightarrow \mathbb{Y} \) be defined as \( T(A, a, x) := Ax + a + F(x) \), the mapping \( y \mapsto T(A, a, y)^{-1} \cap B(\bar{x}, r_0) \) is single-valued from \( B(0, r_0) \subset \mathbb{Y} \) to \( B(\bar{x}, r_0) \). Moreover, for each \( A \in B(A_0, r_{A_0}) \) and \( a \in B(a_0, r_{a_0}) \) there holds \( \lambda\|A - A_0\| < 1 \) and the mapping \( y \mapsto T(A, a, y)^{-1} \cap B(\bar{x}, r_0) \) is also Lipschitzian on \( B(0, r_0) \) as follows

\[ \|T(A, a, y_1)^{-1} \cap B(\bar{x}, r_0) - T(A, a, y_2)^{-1} \cap B(\bar{x}, r_0)\| \leq \frac{\lambda}{1 - \lambda\|A - A_0\|} \|y_1 - y_2\|, \]

for each \( y_1, y_2 \in B(0, r_0) \).

**Proof.** Let \( Z = L(\mathbb{X}, \mathbb{Y}) \times \mathbb{Y} \) and \( g : Z \times \mathbb{X} \rightarrow \mathbb{Y} \) be an operator defined by \( g(A, a, x) = Ax + a \). The operator \( g \) is continuous on \( Z \times \mathbb{X} \) and has partial Fréchet derivative with respect to the variable \( x \) given by \( D_xg(A, a, x) = A \). Note that

\[ A_0x + a_0 + G(x) = g(A_0, a_0, \bar{x}) + D_xg(A_0, a_0, \bar{x})(x - \bar{x}) + G(x), \quad \forall x \in \mathbb{X}, \]

and, by assumption, the mapping \( A_0 + a_0 + G \) is strongly regular at \( \bar{x} \) for 0 with Lipschitz constant \( \lambda \). Then, we may apply Theorem 1.1.2 with \( Z = L(\mathbb{X}, \mathbb{Y}) \times \mathbb{Y}, \bar{p} = (A_0, a_0), p = (A, a) \) and \( g(p, x) = Ax + a \), for concluding that, for any \( \epsilon > 0 \), there exist neighborhoods \( U \) of \( \bar{x} \) and \( V \) of \( (A_0, a_0) \) and a single-valued mapping \( s : V \rightarrow U \) such that for any \( (A, a) \in V \), \( s(A, a) \) is the unique solution in \( U \) of the inclusion

\[ T(A, a, x) = Ax + a + G(x) \geq 0, \]

and \( s(A_0, a_0) = \bar{x} \). Moreover, the following inequality holds

\[ \|s(A, a) - \bar{x}\| \leq (\lambda + \epsilon)\|(A - A_0)\bar{x} + (a - a_0)\|, \quad \forall (A, a) \in V. \quad (1.2) \]
Now, choose neighborhoods $M$ of $A_0$, $N$ de $a_0$ and $W$ of the origin $0 \in \mathbb{Y}$ such that $M \times (N - W) \subset V$. Let $A \in M$, $a \in N$ and $y_1, y_2 \in W$, and let $s(A, a - y_1)$ and $s(A, a - y_2)$ be the solutions associated with $y_1$ and $y_2$, respectively. Since $T(A, a, s(A, a - y_i)) \ni y_i$, i.e., $s(A, a - y_i) = T(A, a, y_i)^{-1} \cap U$, for $i = 1, 2$, after some manipulation, we obtain that

$$y_i + (A_0 - A)s(A, a - y_i) + (a_0 - a) \in A_0s(A, a - y_i) + a_0 + G(s(A, a - y_i)), \quad i = 1, 2. \quad (1.3)$$

Since (1.2) implies that $s : V \to U$ is bounded, we can reduce, if necessary, the neighborhoods $M$, $N$ and $W$ such that

$$y_i + (A_0 - A)s(A, a - y_i) + (a_0 - a) \in W, \quad i = 1, 2.$$

Therefore, taking into account that $A_0 + a_0 + G$ is strongly regular at $\bar{x}$ for 0 with associated Lipschitz constant $\lambda$, last inclusion and (1.3) imply

$$\|s(A, a - y_1) - s(A, a - y_2)\| \leq \lambda \|y_1 + (A_0 - A)s(A, a - y_1) + (a_0 - a) - [y_2 + (A_0 - A)s(A, a - y_2) + (a_0 - a)]\|.$$

Using properties of the norm, last inequality becomes to

$$\|s(A, a - y_1) - s(A, a - y_2)\| \leq \lambda\|y_1 - y_2\| + \lambda\|A_0 - A\|\|s(A, a - y_1) - s(A, a - y_2)\|.$$

Now, if $\tilde{M} \subset M$ and $\lambda\|A - A_0\| < 1$ for each $A \in \tilde{M}$, then last inequality implies that

$$\|s(A, a - y_1) - s(A, a - y_2)\| \leq \frac{\lambda}{1 - \lambda\|A - A_0\|}\|y_1 - y_2\|,$$

and the result follows by noting that $s(A, a - y) = T(A, a, y)^{-1} \cap U$ and $y_1, y_2 \in W$ are arbitrary.

Next, we establish a corollary to Lemma 1.1.3, which plays an important role in the sequel.

**Corollary 1.1.4** Let $X$, $Y$ be Banach spaces, $\Omega \subset X$ be open and nonempty, $f : \Omega \to Y$ be continuous with the Fréchet derivative $f'$ continuous, and $F : X \Rightarrow Y$ be a set-valued mapping. Suppose that $x_0 \in \Omega$ and $L_f(x_0, \cdot) : \Omega \Rightarrow Y$ is strongly regular at $x_1 \in \Omega$ for 0 with modulus $\lambda > 0$. Then, there exist $r_{x_1} > 0$, $r_0 > 0$, and $r_{x_0} > 0$ such that, for each $x \in B(x_0, r_{x_0})$, there holds $\lambda\|f'(x) - f'(x_0)\| < 1$, the mapping $z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1})$ is single-valued from $B(0, r_0)$ to $B(x_1, r_{x_1})$ and Lipschitzian as follows

$$\|L_f(x, u)^{-1} \cap B(x_1, r_{x_1}) - L_f(x, v)^{-1} \cap B(x_1, r_{x_1})\| \leq \frac{\lambda}{1 - \lambda\|f'(x) - f'(x_0)\|}\|u - v\|,$$

for each $u, v \in B(0, r_0)$. 

12
Proof. Since $L_f(x_0, \cdot) : \mathbb{X} \to \mathbb{Y}$ is strongly regular at $x_1$ for $0$ with modulus $\lambda > 0$, applying the first part of Lemma 1.1.3 with $\bar{x} = x_1$, $A_0 = f'(x_0)$ and $a_0 = f(x_0) - f'(x_0)x_0$, we conclude that there exist $r_{x_1} > 0$, $\hat{r} > 0$, $\check{r} > 0$, and $r_0 > 0$ such that, for any $A \in B(f'(x_0), \hat{r}) \subset L(\mathbb{X}, \mathbb{Y})$ and $a \in B(f(x_0) - f'(x_0)x_0, \check{r}) \subset \mathbb{Y}$, letting $T(A, a, \cdot) : B(x_1, r_{x_1}) \to \mathbb{Y}$ be defined as $T(A, a, y) := Ay + a + F(y)$, the mapping $z \mapsto T(A, a, z)^{-1} \cap B(x_1, r_{x_1})$ is a single-valued mapping from $B(0, r_0)$ to $B(x_1, r_{x_1})$. Due to $f$ being continuous with $f'$ continuous, there exists $r_{x_0} > 0$ such that $\lambda\|f'(x) - f'(x_0)\| < 1$,

$$f'(x) \in B(f'(x_0), \hat{r}), \quad f(x) - f'(x)x \in B(f(x_0) - f'(x_0)x_0, \check{r}), \quad \forall x \in B(x_0, r_{x_0}).$$

Hence, we conclude that for each $x \in B(x_0, r_{x_0})$, the mapping $z \mapsto T(f'(x), f(x) - f'(x)x, z)^{-1} \cap B(x_1, r_{x_1})$ is single-valued from $B(0, r_0)$ to $B(x_1, r_{x_1})$, where

$$T(f'(x), f(x) - f'(x)x, y) := f'(x)y + f(x) - f'(x)x + F(y) = f(x) + f'(x)(y - x) + F(y). \quad (1.4)$$

Since (1.1) and (1.4) imply that $L_f(x, y) = T(f'(x), f(x) - f'(x)x, y)$, for all $x \in B(x_0, r_{x_0})$ and $y \in B(x_1, r_{x_1})$, after some manipulations, we have, for each $z \in B(0, r_0)$,

$$L_f(x, z)^{-1} \cap B(x_1, r_{x_1}) = T(f'(x), f(x) - f'(x)x, z)^{-1} \cap B(x_1, r_{x_1}), \quad \forall x \in B(x_0, r_{x_0}). \quad (1.5)$$

Therefore, for $x \in B(x_0, r_{x_0})$, (1.5) and (1.4) imply that $z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1})$ is single-valued from $B(0, r_0)$ to $B(x_1, r_{x_1})$, which proves the first part of the corollary. Finally, considering (1.5) and the second part of Lemma 1.1.3, we also conclude that the mapping $z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1})$ is Lipschitzian from $B(0, r_0)$ to $B(x_1, r_{x_1})$ with the Lipschitz constant $\lambda/[1 - \lambda\|f'(x) - f'(x_0)\|]$, which concludes the proof.

\(\blacksquare\)

### 1.2 Majorant condition

In this section, we define the majorant condition and some classes of functions which satisfies this concept are presented. Firstly, we define the majorant condition for the local case.

**Definition 1.2.1** Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces, $\Omega \subset \mathbb{X}$ be open, $f : \Omega \to \mathbb{Y}$ be continuous with Fréchet derivative $f'$ continuous in $\Omega$. Let $\bar{x} \in \Omega$, $R > 0$, and $\kappa := \sup\{t \in [0, R) : B(\bar{x}, t) \subset \Omega\}$. A twice continuously differentiable function $\psi : [0, R) \to \mathbb{R}$ is a majorant function for $f$ on $B(\bar{x}, \kappa)$ with modulus $\lambda > 0$, if it satisfies the following inequality

$$\lambda\|f'(x) - f'(\bar{x} + \tau(x - \bar{x}))\| \leq \psi' (\|x - \bar{x}\|) - \psi' (\|x - \bar{x}\|), \quad (1.6)$$

for all $\tau \in [0, 1]$, $x \in B(\bar{x}, \kappa)$ and, moreover, there hold:

- **h1)** $\psi(0) = 0$ and $\psi'(0) = -1$;
Let $\Omega$ be differentiable in $\Omega$, is convex, and satisfies

$$\forall x, y \in \Omega$$

Then, for all $x, y \in B(\bar{x}, 1/K) \subset \Omega$. Consider the following class of functions

$$L_\lambda = \{ f \in C^1(\Omega; \mathbb{Y}) : \lambda \| f'(y) - f'(x) \| \leq K \| y - x \|, \ x, y \in B(\bar{x}, 1/K) \}.$$ 

Let $R > 0$ and $\psi : [0, R) \to \mathbb{R}$ be given by $\psi(t) = (K/2)t^2 - t$. Since $\psi'(t) = K t - 1$, it is easy to see that $\psi'(|x - \bar{x}|) - \psi'(0) = K ||x - \bar{x}||$, for all $x \in B(\bar{x}, \kappa)$. Thus, we conclude that all $f \in L_\lambda$ and $\psi$ satisfy (1.6), and the conditions h1 and h2 hold.

The next result gives us an easier condition to check than (1.6), whenever $f$ and $\psi$ are twice continuously differentiable. The proof of this result is similar to Lemma 22 of [23] and is omitted here.

**Lemma 1.2.3** Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces, $\Omega \subseteq \mathbb{X}$ be an open set, $f : \Omega \to \mathbb{Y}$ be twice continuously differentiable. Let $\bar{x} \in \Omega$, $R > 0$ and $\kappa = \sup \{ t \in [0, R) : B(\bar{x}, t) \subset \Omega \}$. Let $\lambda > 0$ and $\psi : [0, R) \to \mathbb{R}$ be twice continuously differentiable such that $\lambda \| f''(x) \| \leq \psi''(|x - \bar{x}|)$, for all $x \in B(\bar{x}, \kappa)$, then $f$ and $\psi$ satisfy (1.8).

**Example 1.2.4** Let $\Omega \subset \mathbb{R}^n$ be an open set, $\bar{x} \in \mathbb{R}^n$, and $\lambda > 0$. Consider a class of analytic functions $f : \Omega \to \mathbb{R}^n$ satisfying Smale’s condition at $\bar{x}$, that is,

$$S_\lambda = \left\{ f : \Omega \to \mathbb{R}^n : f \text{ is analytic and } \gamma := \sup_{n>1} \left\| (\lambda^f(n)(\bar{x})) / n! \right\|^{1/(n-1)} < +\infty \right\},$$

see [6, 59]. Assume that $B(\bar{x}, 1/\gamma) \subset \Omega$. Then, based on the ideas of [2, Lemma 5.3] we can prove that, for all $x \in B(\bar{x}, 1/\gamma)$ and $f \in S_\lambda$, it holds that $\lambda \| f''(x) \| \leq 2\gamma/(1 - \gamma \| x - \bar{x} \|)^3$, see also, for example, [22, Lemma 21]. Since all polynomial functions are in $S_\lambda$, we conclude that $S_\lambda$ is nonempty. Let $\psi : [0, 1/\gamma) \to \mathbb{R}$ be defined by $\psi(t) = t/[1 - \gamma t] - 2t$. We can show that $\psi$ satisfies h1 and h2. Therefore, since $\psi''(t) = 2\gamma/(1 - \gamma t)^3$ and $\lambda \| f''(x) \| \leq 2\gamma/(1 - \gamma \| x - \bar{x} \|)^3$, for all $x \in B(\bar{x}, 1/\gamma)$, thus using Lemma 1.2.9, we conclude that $f \in S_\lambda$ and $\psi$ satisfy (1.6), for all $x, y \in B(\bar{x}, 1/\gamma)$.

**Example 1.2.5** Let $\Omega \subset \mathbb{R}^n$ be a convex set. A function $g : \Omega \to \mathbb{R}$ is called $\alpha$-self-concordant with parameter $\alpha > 0$ if $g \in C^3(\Omega; \mathbb{R})$, i.e., $g$ is three times continuously differentiable in $\Omega$, is convex, and satisfies

$$|g''(x)[h, h, h]| \leq 2a^{-1/2}(g''(x)[h, h])^{3/2}, \quad x \in \Omega, \ h \in \mathbb{R}^n. \quad (1.7)$$
Let \( \bar{x} \in \Omega \) such that \( g''(\bar{x}) \) is invertible. Define the space \( \mathbb{X} := (\mathbb{R}^n, \langle \cdot, \cdot \rangle_\mathbb{X}) \) as the Euclidean space \( \mathbb{R}^n \) with a new inner product and the associated norm, defined by \( \langle u, v \rangle_\mathbb{X} := a^{-1}\langle g''(\bar{x})u, v \rangle \) and \( \|u\|_\mathbb{X} := \sqrt{\langle u, u \rangle_\mathbb{X}} \), for all \( u \in \mathbb{R}^n \), respectively. Thus, the open ball of radius \( r > 0 \) centered at \( x_0 \) (Dikin’s ellipsoid of radius \( r > 0 \) centered at \( \bar{x} \)) in \( \mathbb{X} \) is defined by \( W'_r(\bar{x}) := \{ x \in \mathbb{R}^n : \| x - \bar{x} \|_\mathbb{X} < r \} \). We can prove that \( W'_1(\bar{x}) \subset \Omega \). For more details about self-concordant functions, see \([45]\). Let \( \lambda > 0 \) and consider the class of functions \( f : \Omega \to \mathbb{R}^n \) such that \( \lambda \| g''(\bar{x})\|_x f = g' \) and \( g \) is \( a \)-self-concordant, i.e.,

\[ \mathcal{A}_\lambda = \{ f : \Omega \to \mathbb{R}^n : f = [\lambda \| g''(\bar{x})\|_x]^{-1} g' \text{ and } g \text{ is } a \text{-self-concordant} \} . \]

Using \([2, \text{Lemma } 5.1]\) we can prove that, for all \( x \in W'_1(\bar{x}) \) and \( f \in \mathcal{A}_\lambda \), it holds that \( \lambda \| f''(x) \| \leq 2/(1 - \| x - \bar{x} \|)^3 \). Let \( \psi : [0, 1) \to \mathbb{R} \) be defined by \( \psi(t) = t/[1 - t] - 2t \). Note that \( \psi \) satisfies \( h_1 \) and \( h_2 \). Therefore, since \( \psi''(t) = 2/3(1 - t)^{-3} \), we have \( \lambda \| f''(x) \| \leq \psi''(\| x - \bar{x} \|) \), for all \( x \in W'_1(\bar{x}) \). Thus, using Lemma 1.2.9, we conclude that \( f \in \mathcal{A}_\lambda \) and \( \psi \) satisfy (1.6), for all \( x, y \in W'_1(\bar{x}) \).

**Example 1.2.6** Let \( C^2(\mathbb{R}^n; \mathbb{R}^n) \) be the space of function \( f : \mathbb{R}^n \to \mathbb{R}^n \) twice continuously differentiable, \( \lambda > 0 \), and \( \mu > 1 \). Consider the following class of functions

\[ \mathcal{F}_\lambda = \left\{ f \in C^2(\mathbb{R}^n; \mathbb{R}^n) : \lambda \| f''(x) \| \leq \| x \|^{\mu - 1}, \ x \in \mathbb{R}^n \right\} . \]

The class \( \mathcal{F}_\lambda \) is nonempty. Indeed, define \( f : \mathbb{R}^n \to \mathbb{R}^n \) by \( f(x) = \| x \|^{\mu - 1} x / \| x \|^{\mu} \), where \( \bar{x} \in \mathbb{R}^n \). Some calculations show that

\[ f''(x)(v, v) = \frac{1}{\lambda \mu(1 + \mu)} \left[ \mu(\mu - 2)\| x \|^{\mu - 4} \langle x, v \rangle^2 x + 2\mu\| x \|^{\mu - 2} \langle x, v \rangle v + \mu\| x \|^{\mu - 2}\| v \|^2 x \right] , \]

for all \( x, v \in \mathbb{R}^n \), \( x \neq 0 \), and \( f''(0) = 0 \). Hence, we have \( \lambda \| f''(x) \| \leq \| x \|^{\mu - 1} \), for all \( x \in \mathbb{R}^n \) and thus \( f \in \mathcal{F}_\lambda \). On the other hand, defining \( \psi : [0, +\infty) \to \mathbb{R} \) by \( \psi(t) = t^{1+\mu}/[\mu(1+\mu)] - t \), we conclude from Lemma 1.2.9 that \( \psi \) and any \( f \in \mathcal{F}_\lambda \) satisfy (1.6). We can also show that \( \psi \) satisfies \( h_1 \) and \( h_2 \).

Next, we establish the definition of majorant condition for the semi-local case.

**Definition 1.2.7** Let \( \mathbb{X}, \ Y \) be Banach spaces, \( \Omega \subset \mathbb{X} \) be open, \( f : \Omega \to \mathbb{Y} \) be continuous with Fréchet derivative \( f' \) continuous in \( \Omega \). Let \( x_0 \in \Omega, R > 0 \), and \( \kappa := \sup\{ t \in [0, R) : B(x_0, t) \subset \Omega \} \). A twice continuously differentiable function \( \psi : [0, R) \to \mathbb{R} \) is a majorant function for \( f \) on \( B(x_0, \kappa) \) with modulus \( \lambda > 0 \), if it satisfies the following inequality

\[ \lambda \| f'(y) - f'(x) \| \leq \psi'(\| y - x \| + \| x - x_0 \|) - \psi'(\| x - x_0 \|), \]

for all \( x, y \in B(x_0, \kappa) \) and \( \| y - x \| + \| x - x_0 \| < R \). Moreover, suppose that the following conditions hold:
Lemma 1.2.9

\( \psi(0) > 0, \psi'(0) = -1; \)

\( \psi' \) is convex and strictly increasing;

\( \psi(t) = 0 \) for some \( t \in (0, R) \) and let \( t_* := \min\{t \in [0, R) : \psi(t) = 0\}. \)

Now, we present some examples of functions \( f \) and \( \psi \) satisfying the condition the condition (1.8). The first one is a class of functions satisfying a Lipschitz-type condition.

Example 1.2.8 Let \( X, Y \) be Banach spaces, \( \Omega \subseteq X \) be an open set and \( f \in C^1(\Omega; Y) \), i.e., \( f : \Omega \to Y \) be continuous with Fréchet derivative \( f' \) continuous. Let \( x_0 \in \Omega \) and \( K > 0 \) such that \( B(x_0, 1/K) \subset \Omega \). Consider the following class of functions

\[ \mathcal{L}_\lambda = \{ f \in C^1(\Omega; Y) : \lambda \|f'(y) - f'(x)\| \leq K\|y - x\|, \ x, y \in B(x_0, 1/K) \}. \]

Let \( R > 0, b > 0 \) and \( \psi : [0, R) \to \mathbb{R} \) be given by \( \psi(t) = (K/2)t^2 - t + b \). It can be seen that \( \psi'(\|y - x\| + \|x - x_0\|) - \psi'(\|x - x_0\|) = K\|y - x\| \), for all \( x, y \in X \) such that \( \|y - x\| \leq R \). Thus, we conclude that all \( f \in \mathcal{L}_\lambda \) and \( \psi \) satisfy (1.8), for all \( x, y \in X \) such that \( \|y - x\| \leq R \). Moreover, if \( bK \leq 1/2 \) then \( \psi \) satisfies \( a_1, a_2, \) and \( a_3 \). Additionally, if \( bK < 1/2 \) then \( \psi \) satisfies \( a_4 \).

The next result gives us an easier condition to check than (1.8), whenever \( f \) and \( \psi \) are twice continuously differentiable. The proof of this result is similar to Lemma 22 of [23] and is omitted here.

Lemma 1.2.9 Let \( X, Y \) be Banach spaces, \( \Omega \subseteq X \) be an open set, \( f : \Omega \to Y \) be twice continuously differentiable. Let \( x_0 \in \Omega, \ R > 0 \) and \( \kappa = \sup\{t \in [0, R) : B(x_0, t) \subset \Omega\} \). Let \( \lambda > 0 \) and \( \psi : [0, R) \to \mathbb{R} \) be twice continuously differentiable such that \( \lambda \|f''(x)\| \leq \psi''(\|x - x_0\|), \) for all \( x \in B(x_0, \kappa) \), then \( f \) and \( \psi \) satisfy (1.8).

Example 1.2.10 Let \( \Omega \subseteq \mathbb{R}^n \) be an open set, \( x_0 \in \mathbb{R}^n \), and \( \lambda > 0 \). Consider a class of analytic functions \( f : \Omega \to \mathbb{R}^n \) satisfying Smale’s condition at \( x_0 \), that is,

\[ \mathcal{S}_\lambda = \left\{ f : \Omega \to \mathbb{R}^n : f \text{ is analytic and } \gamma := \sup_{n \geq 1} \left\| [\lambda f^{(n)}(x_0)]/n! \right\|^{1/(n-1)} < +\infty \right\}, \]

see [6,59]. Assume that \( B(x_0, 1/\gamma) \subset \Omega \). Then, based on the ideas of [2, Lemma 5.3] we can prove that, for all \( x \in B(x_0, 1/\gamma) \) and \( f \in \mathcal{S}_\lambda \), it holds that \( \lambda \|f''(x)\| \leq 2\gamma/(1 - \gamma \|x - x_0\|)^3 \), see also, for example, [22, Lemma 21]. Since all polynomial functions are in \( \mathcal{S}_\lambda \), we conclude that \( \mathcal{S}_\lambda \) is nonempty. Let \( b > 0 \) and let \( \psi : [0, 1/\gamma) \to \mathbb{R} \) be defined by \( \psi(t) = t/[1 - \gamma t] - 2t + b \). We can show that \( \psi \) satisfies \( a_1, a_2, a_3, \) and \( a_4 \). Therefore, since \( \psi''(t) = 2\gamma/(1 - \gamma t)^3 \) and \( \lambda \|f''(x)\| \leq 2\gamma/(1 - \gamma \|x - x_0\|)^3 \), for all \( x \in B(x_0, 1/\gamma) \), thus using Lemma 1.2.9, we conclude that \( f \in \mathcal{S}_\lambda \) and \( \psi \) satisfy (1.8), for all \( x, y \in B(x_0, 1/\gamma) \) and \( \|y - x\| + \|x - x_0\| < 1/\gamma \).
Example 1.2.11 Let $\Omega \subset \mathbb{R}^n$ be a convex set and $g : \Omega \to \mathbb{R}$ be an $\alpha$-self-concordant function, as defined in Example 1.2.5. Let $x_0 \in \Omega$ such that $g''(x_0)$ is invertible. Define the space $X := (\mathbb{R}^n, \langle \cdot, \cdot \rangle_{x_0})$ as the Euclidean space $\mathbb{R}^n$ with a new inner product and the associated norm, defined by $\langle u, v \rangle_{x_0} := a^{-1} \langle g''(x_0)u, v \rangle$ and $\|u\|_{x_0} := \sqrt{\langle u, u \rangle_{x_0}}$, for all $u \in \mathbb{R}^n$, respectively. Thus, the open ball of radius $r > 0$ centered at $x_0$ (Dikin’s ellipsoid of radius $r > 0$ centered at $x_0$) in $X$ is defined by $W_r(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} < r\}$. We can prove that $W_1(x_0) \subset \Omega$. Let $\lambda > 0$ and consider the class of functions $f : \Omega \to \mathbb{R}^n$ such that $\lambda\|g''(x_0)\|_{x_0} f = g'$ and $g$ is $\alpha$-self-concordant, i.e.,

$$A_\lambda = \{ f : \Omega \to \mathbb{R}^n : f = [\lambda\|g''(x_0)\|_{x_0}]^{-1}g' \text{ and } g \text{ is } \alpha\text{-self-concordant} \}. $$

Using [2, Lemma 5.1] we can prove that, for all $x \in W_1(x_0)$ and $f \in A_\lambda$, it holds that $\lambda\|f''(x)\| \leq 2/(1 - \|x - x_0\|^3)$. Let $b > 0$ and let $\psi : [0,1) \to \mathbb{R}$ be defined by $\psi(t) = t/[1-t] - 2t + b$. Note that $\psi$ satisfies a1, a2, a3, and a4. Therefore, since $\psi''(t) = 2/(1 - t)^3$, we have $\lambda\|f''(x)\| \leq \psi''(\|x - x_0\|)$, for all $x \in W_1(x_0)$. Thus, using Lemma 1.2.9, we conclude that $f \in A_\lambda$ and $\psi$ satisfy (1.8), for all $x, y \in W_1(x_0)$ and $\|y - x\| + \|x - x_0\| < 1$.

Example 1.2.12 Let $C^2(\mathbb{R}^n; \mathbb{R}^n)$ be the space of functions $f : \mathbb{R}^n \to \mathbb{R}^n$ twice continuously differentiable, $\lambda > 0$, and $\mu > 1$. Consider the following class of functions

$$F_\lambda = \{ f \in C^2(\mathbb{R}^n, \mathbb{R}^n) : \lambda\|f''(x)\| \leq \|x\|^\mu, \ x \in \mathbb{R}^n \}. $$

The class $F_\lambda$ is nonempty. Indeed, define $f : \mathbb{R}^n \to \mathbb{R}^n$ by $f(x) = [(\|x\|^\mu x - \bar{x})/\lambda\mu(1 + \mu)]$, where $\bar{x} \in \mathbb{R}^n$. Some calculations show that

$$f''(x,v,v) = \frac{1}{\lambda\mu(1 + \mu)} \left[ \mu(\mu - 2)\|x\|^{\mu-4}\langle x, v \rangle^2 x + 2\mu\|x\|^{\mu-2}\langle x, v \rangle v + \mu\|x\|^{\mu-2}\|v\|^2 x \right],$$

for all $x, v \in \mathbb{R}^n$, $x \neq 0$, and $f''(0) = 0$. Hence, we have $\lambda\|f''(x)\| \leq \|x\|^{\mu-1}$, for all $x \in \mathbb{R}^n$ and thus $f \in F_\lambda$. On the other hand, letting $b > 0$ and defining $\psi : [0, +\infty) \to \mathbb{R}$ by $\psi(t) = t^{1+\mu}/[\mu(1 + \mu)] - t + b$, we conclude from Lemma 1.2.9 that $\psi$ and any $f \in F_\lambda$ satisfy (1.8) with $x_0 = 0$. We can also show that $\psi$ satisfies a1, a2, a3, and a4.
Chapter 2

Local convergence analysis of Newton’s method for solving strongly regular generalized equations

In this chapter, we present an analysis of the behavior of the sequence generated by Newton’s method for solving the generalized equation (2). For this purpose, we assume that $L_f(\bar{x},.)$ is strongly regular at $\bar{x}$ for 0 with modulus $\lambda > 0$. Further, we assume that Lipschitz continuity of $f'$ is relaxed, i.e., we assume that $f'$ satisfies the majorant condition.

2.1 Local analysis of Newton’s method

The statement of our main result in this chapter is:

**Theorem 2.1.1** Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces, $\Omega \subset \mathbb{X}$ be open, $f : \Omega \to \mathbb{Y}$ be continuous with Fréchet derivative $f'$ continuous in $\Omega$, $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a set-valued mapping with closed graph and $\bar{x} \in \Omega$. Suppose that $L_f(\bar{x},.) : \Omega \rightrightarrows \mathbb{Y}$ is strongly regular at $\bar{x}$ for 0 with modulus $\lambda > 0$. Let $R > 0$, $\kappa := \sup\{t \in [0,R) : B(\bar{x},t) \subset \Omega\}$ and assume that $\psi : [0, R) \to \mathbb{R}$ is a majorant function for $f$ on $B(\bar{x},\kappa)$ with modulus $\lambda > 0$. Let $\nu := \sup\{t \in [0, R) : \psi'(t) < 0\}$, $\rho := \sup\{t \in (0, \nu) : \psi(t)/(t\psi'(t)) - 1 < 1\}$ and $r := \min\{\kappa, \rho\}$. Then, there exists $r_{\bar{x}} > 0$ with $r_{\bar{x}} \leq r$ such that the sequences with initial point $x_0 \in B(\bar{x},r_{\bar{x}})/\{\bar{x}\}$ and $t_0 = \|\bar{x} - x_0\|$, respectively,

$$0 \in f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1}), \quad t_{k+1} = |t_k - \psi(t_k)/\psi'(t_k)|, \quad k = 0, 1, \ldots,$$

(2.1)

are well defined, $\{t_k\}$ is strictly decreasing, is contained in $(0,r)$ and converges to 0, $\{x_k\}$ is contained in $B(\bar{x},r_{\bar{x}})$ and converges to the point $\bar{x}$, which is the unique solution of $f(x) + $
\( F(x) \ni 0 \) in \( B(\bar{x}, \bar{\sigma}) \), where \( 0 < \bar{\sigma} \leq \min \{ r_\bar{x}, \sigma \} \) and \( \sigma := \sup \{ 0 < t < \kappa : \psi(t) < 0 \} \) and there hold
\[
\lim_{k \to \infty} \frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} = 0, \quad \lim_{k \to \infty} \frac{t_{k+1}}{t_k} = 0. \quad (2.2)
\]
Moreover, given \( 0 \leq p \leq 1 \) and assuming that
\begin{itemize}
  \item [h3)] the function \((0, \nu) \ni t \mapsto (\psi(t)/\psi'(t) - t)/t^{p+1}\) is strictly increasing,
\end{itemize}
then the sequence \( \{t_{k+1}/t_k^{p+1}\} \) is strictly decreasing and satisfies the following condition
\[
\|x_{k+1} - \bar{x}\| \leq \frac{t_{k+1}}{t_k^{p+1}} \|x_k - \bar{x}\|^{p+1}, \quad k = 0, 1, \ldots. \quad (2.3)
\]
If, additionally, \( \psi(\rho)/(|\rho\psi'(\rho)) - 1 = 1 \) and \( \rho < \kappa \), then \( r_\bar{x} = \rho \) is the largest radius of convergence.

Remark 2.1.2 The first equality in (2.2) means that \( \{x^k\} \) converges superlinearly to \( \bar{x} \). For \( \psi \) such that the derivative \( \psi' \) is convex, the condition h3 holds with \( p = 1 \). In this case, the following condition is satisfied
\[
\frac{t_{k+1}}{t_k^2} \leq \frac{\psi''(t_0)}{2|\psi'(t_0)|}, \quad k = 0, 1, \ldots,
\]
and \( \{x^k\} \) converges quadratically. Indeed, convexity of \( \psi' \) is necessary to obtain quadratic convergence; see Example 2 in [21]. Moreover, as \( \{t_{k+1}/t_k^{p+1}\} \) is strictly decreasing, we have \( t_{k+1}/t_k^{p+1} \leq t_1/t_0^{p+1} \), for \( k = 0, 1, \ldots \). Thus, (2.3) implies \( \|x_{k+1} - \bar{x}\| \leq (t_1/t_0^{p+1}) \|x_k - \bar{x}\|^{p+1} \), for \( k = 0, 1, \ldots \). Consequently, if \( p = 0 \), then \( \|x_k - \bar{x}\| \leq t_0[t_1/t_0]^k \), for \( k = 0, 1, \ldots \) and if \( 0 < p \leq 1 \), then there holds
\[
\|x_k - \bar{x}\| \leq t_0(t_1/t_0)^{(p+1)k-1}/p, \quad k = 0, 1, \ldots.
\]

Remark 2.1.3 Throughout the proof of the above theorem, we assume that if \( F \equiv \{0\} \), then the constant \( r_\bar{x} = \nu \). In this case, Theorem 2.1.1 merges into Theorem 2 of [21].

Hereafter, we assume that the hypotheses of Theorem 2.1.1 hold, with the exception of h3, which will be considered to hold only when explicitly stated.

### 2.2 Preliminary results

In this section, our first goal is to prove all statements in Theorem 2.1.1 concerning the sequence \( \{t_k\} \) associated with the majorant function \( \psi \) defined in (2.1). Moreover, we obtain
a few relationships between the majorant function $\psi$ and the set-valued mapping $f + F$, which will play an important role throughout the thesis. Furthermore, the results in Theorem 2.1.1 related to the uniqueness of the solution and the optimal convergence radius will be proved. We begin with some observations on the majorant function.

As proved in Proposition 2.5 of [20], the constants $\kappa$, $\nu$ and $\sigma$ in Definition 1.2.7 and Theorem 2.1.1 are all positive and $t - \psi(t)/\psi'(t) < 0$, for all $t \in (0, \nu)$. According to $h_2$ and the definition of $\nu$, we have $\psi'(t) < 0$, for all $t \in [0, \nu)$. Therefore, the Newton iteration map $n_\psi : [0, \nu) \to (-\infty, 0]$ given by

$$n_\psi(t) = t - \frac{\psi(t)}{\psi'(t)}, \quad t \in [0, \nu),$$

is well defined. Proposition 2.2.1, which follows below, was proved in [21, Proposition 4].

**Proposition 2.2.1** Assume that the hypotheses in Theorem 2.1.1 about $\psi$ hold. Then $\lim_{t \to 0} |n_\psi(t)|/t = 0$ and the constant $\rho$ is positive. As a consequence, $|n_\psi(t)| < t$ for all $t \in (0, \rho)$.

Using (2.4), it is easy to see that the sequence $\{t_k\}$ defined in (2.1) is rewritten as

$$t_0 = \|\bar{x} - x_0\|, \quad t_{k+1} = |n_\psi(t_k)|, \quad k = 0, 1, \ldots$$

(2.5)

Next result, which is a consequence of the above proposition, contains the main convergence properties of the sequence $\{t_k\}$ and its proof can be found in Corollary 5 of [21].

**Corollary 2.2.2** The sequence $\{t_k\}$ is well defined, is strictly decreasing and is contained in $(0, \rho)$. Moreover, $\{t_k\}$ converges to 0 with superlinear rate, i.e., $\lim_{k \to \infty} t_{k+1}/t_k = 0$. If additionally $h_3$ holds, then the sequence $t_{k+1}/t_k^{p+1}$ is strictly decreasing.

In the sequel, we study the linearization error of the function $f$ in $\Omega$ defined by

$$E_f(x, y) := f(y) - \left[ f(x) + f'(x)(y - x) \right], \quad x, y \in \Omega.$$  \hspace{1cm} (2.6)

We show that the above error is bounded by the linearization error of the function $\psi$ defined by

$$e_\psi(t, u) := \psi(u) - \left[ \psi(t) + \psi'(t)(u - t) \right], \quad t, u \in [0, R),$$

which can be stated as:

**Lemma 2.2.3** There holds $\lambda \|E_f(x, \bar{x})\| \leq e_\psi(\|x - \bar{x}\|, 0)$, for all $x \in B(\bar{x}, \kappa)$.  \hspace{1cm} (2.7)
Proof. Since $\bar{x}+(1-u)(x-\bar{x}) \in B(\bar{x}, \kappa)$, for all $0 \leq u \leq 1$ and $f$ is continuously differentiable in $\Omega$, with the definition of $E_f$, and after some simple manipulation, we obtain

$$\lambda\|E_f(x, \bar{x})\| \leq \int_0^1 \lambda\|f'(x) - f'(\bar{x} + (1-u)(x-\bar{x}))\|\|x-\bar{x}\| \, du.$$  

Using (1.6) in the last inequality, carrying out the integration and using the definition of $e_\psi$, leads to the desired result.  

Next lemma, states that, if $L_f(\bar{x},.)$ is strongly regular at $\bar{x}$ for $0$ with modulus $\lambda > 0$ and (1.6) holds, then there exists a neighborhood of $\bar{x}$ such that for all $x$ in this neighborhood, $f+F$ is also strongly regular at $x$ for $0$ with modulus $\lambda/(\|\psi(\|x-\bar{x}\|)\|)$. The result is a consequence of Corollary 1.1.4:

**Lemma 2.2.4** There exists a constant $r_x \leq r$ such that, the mapping $x \mapsto L_f(x,0)^{-1} \cap B(\bar{x}, r_x)$ is single-valued in $B(\bar{x}, r_x)$ and the following condition is satisfied

$$\|\bar{x} - L_f(x,0)^{-1} \cap B(\bar{x}, r_x)\| \leq \frac{\lambda}{\psi(\|x-\bar{x}\|)}\|E_f(x, \bar{x})\|, \quad \forall \ x \in B(\bar{x}, r_x).$$

**Proof.** Let $x \in B(\bar{x}, r)$. Since $r < \nu$, we have $\|x-\bar{x}\| < \nu$. Thus, $\psi(\|x-\bar{x}\|) < 0$ which, together with (1.6) and h1, imply that

$$\lambda\|f'(x) - f'(\bar{x})\| \leq \psi(\|x-\bar{x}\|) - \psi(0) < -\psi(0) = 1, \quad \forall \ x \in B(\bar{x}, r).$$

(2.7)

Since $f+F$ is strongly regular at $\bar{x}$ for $0$ with modulus $\lambda > 0$, we can apply Corollary 1.1.4 to obtain $r_x > 0$ and $r_0 > 0$ such that, for any $x \in B(\bar{x}, r_x)$, the mapping $z \mapsto L_f(x,z)^{-1} \cap B(\bar{x}, r_x)$ is single-valued from $B(0, r_0)$ to $B(\bar{x}, r_x)$. In particular, we conclude that the mapping $x \mapsto L_f(x,0)^{-1} \cap B(\bar{x}, r_x)$ is single-valued in $B(\bar{x}, r_x)$. Moreover, Corollary 1.1.4 implies that

$$\|L_f(x,u)^{-1} \cap B(\bar{x}, r_x) - L_f(x,v)^{-1} \cap B(\bar{x}, r_x)\| \leq \frac{\lambda\|u-v\|}{1 - \lambda\|f'(x) - f'(\bar{x})\|}, \quad \forall \ u,v \in B(0, r_0).$$

If necessary, we shrink $r_x$ such that $r_x \leq r$, in order to combine the last inequality with the first inequality in (2.7) and h1, to conclude that, for all $x \in B(\bar{x}, r_x)$ the following condition is satisfied

$$\|L_f(x,u)^{-1} \cap B(\bar{x}, r_x) - L_f(x,v)^{-1} \cap B(\bar{x}, r_x)\| \leq \frac{\lambda\|u-v\|}{\psi(\|x-\bar{x}\|)}, \quad \forall \ u,v \in B(0, r_0).$$

(2.8)

On the other hand, since $f$ and $f'$ are continuous in $\Omega$, we have $\lim_{x \to \bar{x}} E_f(x, \bar{x}) = 0$. Thus, we can shrink $r_x$, if necessary, such that

$$E_f(x, \bar{x}) \in B(0, r_0), \quad \forall \ x \in B(\bar{x}, r_x).$$

(2.9)
Let \( x \in B(\bar{x}, r_x) \). After some simple algebraic manipulation, and by using the linearization error (2.6), we obtain
\[
0 \in f(\bar{x}) + F(\bar{x}) = f(x) + f'(x)(\bar{x} - x) - f(x) - f'(x)(\bar{x} - x) + f(\bar{x}) + F(\bar{x}) = f(x) + f'(x)(\bar{x} - x) + E_f(x, \bar{x}) + F(\bar{x}).
\]
Hence, (1.1) implies \(-E_f(x, \bar{x}) \in L_f(x, \bar{x}) = f(x) + f'(x)(\bar{x} - x) + F(\bar{x})\). Thus, from (2.9) and due to the mapping \( z \mapsto L_f(x, z)^{-1} \cap B(\bar{x}, r_x) \) being single valued from \( B(0, r_0) \) to \( B(\bar{x}, r_x) \), we conclude that
\[
\bar{x} = L_f(x, -E_f(x, \bar{x}))^{-1} \cap B(\bar{x}, r_x).
\]
Therefore, substituting \( u = -E_f(x, \bar{x}) \) and \( v = 0 \) into (2.8) the desired inequality follows.

Lemma 2.2.4 guarantees, in particular, that the mapping \( x \mapsto L_f(x, 0)^{-1} \cap B(\bar{x}, r_x) \) is single valued in \( B(\bar{x}, r_x) \) and consequently, the Newton iteration mapping is well defined. Denoting the Newton iteration mapping for \( f + F \) in that region by \( N_{f+F} \), namely, \( N_{f+F} : B(\bar{x}, r_x) \to X \) is defined by
\[
N_{f+F}(x) := L_f(x, 0)^{-1} \cap B(\bar{x}, r_x), \quad \forall \ x \in B(\bar{x}, r_x).
\] (2.10)
Using (1.1) and definition of \( L_f(x, \cdot)^{-1} \), the Newton iteration mapping in (2.10) is equivalent to
\[
0 \in f(x) + f'(x)(N_{f+F}(x) - x) + F(N_{f+F}(x)), \quad N_{f+F}(x) \in B(\bar{x}, r_x), \quad \forall \ x \in B(\bar{x}, r_x).
\] (2.11)
for each \( x \in B(\bar{x}, r_x) \). Therefore, since Lemma 2.2.4 guarantees that \( N_{f+F}(x) \) is single valued in \( B(\bar{x}, r_x) \), see (2.10), we can apply a single Newton iteration for any \( x \in B(\bar{x}, r_x) \) to obtain \( N_{f+F}(x) \) which may not belong to \( B(\bar{x}, r_x) \), or even may not belong to the domain of \( f \). Thus, this allow us to guarantee the well-definedness of only one iteration in Newton’s method. In particular, the next result shows that for any \( x \in B(\bar{x}, r_x) \), Newton iterations, see (2.11), may be repeated indefinitely.

**Lemma 2.2.5** If \( \|x - \bar{x}\| \leq t < r_x \), then \( \|N_{f+F}(x) - \bar{x}\| \leq |n_\psi(\|x - \bar{x}\|)|. \) As a consequence, \( N_{f+F}(B(\bar{x}, r_x)) \subset B(\bar{x}, r_x) \). Moreover, if h3 holds then
\[
\|N_{f+F}(x) - \bar{x}\| \leq \frac{|n_\psi(t)|}{tp+1} \|x - \bar{x}\|^{p+1}.
\]

**Proof.** Since \( 0 \in f(\bar{x}) + F(\bar{x}) \), we have \( \bar{x} = N_{f+F}(\bar{x}) \). Thus, the inequalities of the lemma are trivial for \( x = \bar{x} \). Now, we assume that \( 0 < \|x - \bar{x}\| \leq t \). Hence, Lemma 2.2.4 implies that the mapping \( x \mapsto L_f(x, 0)^{-1} \cap B(\bar{x}, r_x) \) is single-valued in \( B(\bar{x}, r_x) \) and Lipschitz continuous with modulus \( \lambda/|\psi'(\|x - \bar{x}\|)| \). Using (2.10) and Lemma 2.2.4, it is easy to conclude that
\[
\|\bar{x} - N_{f+F}(x)\| \leq \frac{\lambda}{|\psi'(\|x - \bar{x}\|)|} \|E_f(x, \bar{x})\|.
\]
Thus, application of the last inequality and Lemma 2.2.3 leads to
\[
\| \bar{x} - N_{f+F}(x) \| \leq \frac{e_{\psi}(\| x - \bar{x} \|, 0)}{|\psi'(\| x - \bar{x} \|)|}.
\]
On the other hand, taking into account that \( \psi(0) = 0 \), the definitions of \( e_{\psi} \) and \( n_{\psi} \) imply that
\[
\frac{e_{\psi}(\| x - \bar{x} \|, 0)}{|\psi'(\| x - \bar{x} \|)|} = \psi(\| x - \bar{x} \|) - \| x - \bar{x} \| = |n_{\psi}(\| x - \bar{x} \|)|.
\]
Hence, the first part of the lemma follows by combining the two previous expressions. For proving the inclusion of the lemma, let \( x \in B(\bar{x}, r_{x}) \). Since \( \| x - \bar{x} \| < r_{x}, r_{x} \leq \rho \) and \( \| N_{f+F}(x) - \bar{x} \| \leq |n_{\psi}(\| x - \bar{x} \|)| \), by using the second part of Proposition 2.2.1 we conclude that \( \| N_{f+F}(x) - \bar{x} \| < \| x - \bar{x} \| \) which proves the inclusion. Next, we prove the last inequality of the lemma. If \( 0 < \| x - \bar{x} \| \leq t \) then assumption \( h3 \) and (2.4) yields
\[
\frac{|n_{\psi}(\| x - \bar{x} \|)|}{\| x - \bar{x} \|^{p+1}} < \frac{|n_{\psi}(t)|}{\rho^{p+1}}.
\]
Therefore, using the first part of Lemma 2.2.5 the desired inequality follows.

In the next result we obtain uniqueness of the solution for (2) in the neighborhood \( B[\bar{x}, \sigma] \).

**Lemma 2.2.6** There exists \( \bar{\sigma} \leq \min\{r_{x}, \sigma\} \) such that \( \bar{x} \) is the unique solution of (2) in \( B[\bar{x}, \bar{\sigma}] \).

**Proof.** Let \( r_{x} > 0 \) be the constant given by Lemma 2.2.4. Thus, Corollary 1.1.4 implies that there exists \( r_{0} > 0 \), such that for any \( x \in B(\bar{x}, r_{x}) \), the mapping \( z \mapsto L_{f}(x, z)^{-1} \cap B(\bar{x}, r_{x}) \) is single-valued from \( B(0, r_{0}) \) to \( B(\bar{x}, r_{x}) \) and the following inequality is satisfied
\[
\| L_{f}(x, u)^{-1} \cap B(\bar{x}, r_{x}) - L_{f}(x, v)^{-1} \cap B(\bar{x}, r_{x}) \| \leq \frac{\lambda\| u - v \|}{1 - \lambda\| f'(x) - f'(\bar{x}) \|},
\]
for each \( u, v \in B(0, r_{0}) \). Now, due to \( f \) being continuous, we have \( \lim_{x \to \bar{x}} E_{f}(\bar{x}, x) = 0 \). Thus, we can take \( \bar{\sigma} \leq \min\{r_{x}, \sigma\} \), such that
\[
E_{f}(\bar{x}, x) \in B(0, r_{0}), \quad \forall x \in B(\bar{x}, \bar{\sigma}). \tag{2.12}
\]
Let \( y \in B(\bar{x}, \bar{\sigma}) \) and assume that \( 0 \in f(y) + F(y) \). Then, after some manipulations, we obtain
\[
0 \in f(y) + F(y) = f(y) - f(\bar{x}) - f'(\bar{x})(y - \bar{x}) + f(\bar{x}) + f'(\bar{x})(y - \bar{x}) + F(y) = E_{f}(\bar{x}, y) + L_{f}(\bar{x}, y),
\]
which implies that \( -E_{f}(\bar{x}, y) \in L_{f}(\bar{x}, y) \). Since \( y \in B(\bar{x}, \bar{\sigma}) \), from (2.12), and due to the mapping \( z \mapsto L_{f}(\bar{x}, z)^{-1} \cap B(\bar{x}, r_{x}) \) being single-valued from \( B(0, r_{0}) \) to \( B(\bar{x}, r_{x}) \), we have,
\[
y = L_{f}(\bar{x}, -E_{f}(\bar{x}, y))^{-1} \cap B(\bar{x}, r_{x}), \quad \bar{x} = L_{f}(\bar{x}, 0)^{-1} \cap B(\bar{x}, r_{x}).
\]
Hence, substituting $x = \bar{x}$, $u = 0$ and $v = -E_f(\bar{x}, y)$, in the above inequality, we conclude that
\[
\|\bar{x} - y\| = \|L_f(\bar{x}, 0)^{-1} \cap B(\bar{x}, r_\bar{x}) - L_f(\bar{x}, -E_f(\bar{x}, y))^{-1} \cap B(\bar{x}, r_\bar{x})\| \leq \lambda\|E_f(\bar{x}, y)\|.
\]

Using (2.6), and last inequality we obtain that
\[
\|\bar{x} - y\| \leq \lambda\|f(y) - f(\bar{x}) - f'(\bar{x})(y - \bar{x})\| \leq \int_0^1 \lambda\|f'(\bar{x} + u(y - \bar{x})) - f'(\bar{x})\|\|y - \bar{x}\|du.
\]

Letting $x = \bar{x} + u(y - \bar{x})$ and $\tau = 0$ in (1.6), the latter inequality yields
\[
\|\bar{x} - y\| \leq \int_0^1 [\psi'(u\|y - \bar{x}\|)) - \psi'(0)]\|y - \bar{x}\|du.
\]

Carrying out the integral in the above inequality, we have $0 \leq \psi(\|y - \bar{x}\|)$, and since $\psi(t) < 0$ for $t \in (0, \sigma)$ and $\|y - \bar{x}\| \leq \sigma$, we obtain $\psi(\|y - \bar{x}\|) = 0$. Since, $0 \leq \|y - \bar{x}\| \leq \sigma$ and $0$ is the unique zero of $\psi$ in $[0, \sigma]$, we conclude that $\|y - \bar{x}\| = 0$ and $\bar{x}$ is the unique solution of (2) in $B[\bar{x}, \sigma]$.

The next result leads to the largest radius of convergence; its proof is similar to the proof of Lemma 2.15 of [20].

**Lemma 2.2.7** If $\psi(\rho)/(\rho\psi'(\rho)) - 1 = 1$ and $\rho < \kappa$, then $r_\bar{x} = \rho$ is the largest radius of convergence.

### 2.2.1 Proof of Theorem 2.1.1

In this section, we will prove the statements in Theorem 2.1.1 involving the Newton sequence $\{x_k\}$.

First, note that the inclusion in (2.1) together with (2.10) and (2.11) imply that $\{x_k\}$ satisfies
\[
x_{k+1} = N_{f+F}(x_k), \quad k = 0, 1, \ldots, \tag{2.13}
\]
which is indeed an equivalent definition of this sequence.

**Proof.** All statements involving $\{t_k\}$ were proved in Corollary 2.2.2. Since Lemma 2.2.4 and (2.10) imply that there exist constants $r_\bar{x} > 0$ and $r_0 > 0$ such that $r_\bar{x} \leq r$, and for any $x \in B(\bar{x}, r_\bar{x})$, the mapping $N_{f+F}$ is single valued in $B(\bar{x}, r_\bar{x})$. Thus, considering that Lemma 2.2.5 implies $N_{f+F}(B(\bar{x}, r_\bar{x})) \subset B(\bar{x}, r_\bar{x})$, we can conclude from $x_0 \in B(\bar{x}, r_\bar{x})$ and (2.13) that $\{x_k\}$ is well defined and remains in $B(\bar{x}, r_\bar{x})$. Now, we will prove that $\{x_k\}$ converges towards $\bar{x}$. Without loss of generality, we assume that $\{x_k\}$ is an infinity sequence.
Since \( 0 < \|x_k - \bar{x}\| < r_\bar{x} \leq \rho \), for \( k = 0, 1, \ldots \), we obtain from (2.13), Lemma 2.2.5 and second part of Proposition 2.2.1 that

\[
\|x_{k+1} - \bar{x}\| \leq |n_\psi(\|x_k - \bar{x}\|)| < \|x_k - \bar{x}\|, \quad k = 0, 1, \ldots \tag{2.14}
\]

Thus, \( \{\|x_k - \bar{x}\|\} \) is strictly decreasing and convergent. Let \( \bar{\alpha} = \lim_{k \to \infty} \|x_k - \bar{x}\| \). Because \( \{\|x_k - \bar{x}\|\} \) is contained in \((0, \rho)\) and is strictly decreasing, we have \( 0 \leq \bar{\alpha} < \rho \). Then, by continuity of \( n_\psi \) and (2.14), we obtain \( 0 \leq \bar{\alpha} = |n_\psi(\bar{\alpha})| \), and from second part of Proposition 2.2.1, we have \( \bar{\alpha} = 0 \). Therefore, the convergence of \( \{x_k\} \) to \( \bar{x} \) is proved. Now, we will show that \( \bar{x} \) is a solution of the generalized equation \( f(x) + F(x) \ni 0 \). From (2.1) we conclude that

\[
(x_{k+1}, -f(x_k) - f'(x_k)(x_{k+1} - x_k)) \in \text{gph } F, \quad k = 0, 1, \ldots
\]

Based on the assumption that the set-valued mapping \( F \) has closed graph, and \( f \) and \( f' \) are continuous, the last inclusion leads to

\[
\lim_{k \to \infty} (x_{k+1}, -f(x_k) - f'(x_k)(x_{k+1} - x_k)) = (\bar{x}, -f(\bar{x})) \in \text{gph } F,
\]

which implies that \( f(\bar{x}) + F(\bar{x}) \ni 0 \). Now, we will show the first equality in (2.2). Note that (2.14) implies that

\[
\frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} \leq \frac{|n_\psi(\|x_k - \bar{x}\|)|}{\|x_k - \bar{x}\|}, \quad k = 0, 1, \ldots
\]

Since \( \lim_{k \to \infty} \|x_k - \bar{x}\| = 0 \), the desired equality follows from the first statement in Proposition 2.2.1. To prove (2.3), firstly we will show by induction that \( \{x_k\} \) and \( \{t_k\} \) defined in (2.1) satisfy

\[
\|x_k - \bar{x}\| \leq t_k, \quad k = 0, 1, \ldots \tag{2.15}
\]

Since \( t_0 = \|x_0 - \bar{x}\| \), the above inequality holds for \( k = 0 \). Now, we assume that \( \|x_k - \bar{x}\| \leq t_k \). Using (2.13), second part of Lemma 2.2.5, the induction assumption and (2.5), we have

\[
\|x_{k+1} - \bar{x}\| = \|N_{f+F}(x_k) - \bar{x}\| \leq \frac{n_\psi(t_k)}{t_{k+1}} \|x_k - \bar{x}\|^{p+1} = \frac{t_{k+1}}{t_k} \|x_k - \bar{x}\|^{p+1} \leq t_{k+1},
\]

and the proof by induction is complete. Thus, the inequality (2.3) follows from the combination of (2.15) and the second part of Lemma 2.2.5. Finally, the uniqueness follows from Lemma 2.2.6, and the last statement in the theorem follows from Lemma 2.2.7. \( \blacksquare \)
2.3 Particular cases

In this section, some special cases of Theorem 2.1.1 will be considered. We begin by remarking that Theorem 2.1.1 for $F = N_C$, the normal cone of a convex set $C$ and $\psi(t) = \lambda L t^2/(2) - t$, subject to some minor modifications is equivalent to Theorem 1 of [36], where $\lambda$ is the modulus of strong regularity of $f + N_C$ and $L$ is the Lipschitz constant of $f'$. It has been shown in [15, Theorem 1] that, if $C \subset \mathbb{R}^n$ is a polyhedral convex set, then the Aubin continuity of the inverse of $f(x) + f'(x)(\cdot - x) + N_C$ is equivalent to strong regularity of $f + N_C$. In particular, we can conclude that [13, Theorem 2] follows from Theorem 2.1.1. In this case, $\psi : [0, 1/L) \to \mathbb{R}$, defined by $\psi(t) := (\lambda L/2)t^2 - t$, is the function associated with $f$ satisfying (1.6). In the following sections, we shall discuss in more detail some other particular instances of Theorem 2.1.1.

2.3.1 Under Hölder-type condition

The next result, which is a consequence of our main result Theorem 2.1.1, is a version of a classical convergence theorem for Newton’s method under Hölder-type condition for solving generalized equations of type (2). Some classical versions for $F \equiv \{0\}$ have appeared in [34, 43, 50, 60].

**Theorem 2.3.1** Let $X$, $Y$ be Banach spaces, $\Omega \subseteq X$ an open set and $f : \Omega \to Y$ be continuous with Fréchet derivative $f'$ being continuous in $\Omega$, $F : X \rightrightarrows Y$ be a set-valued mapping with closed graph, and $\bar{x} \in \Omega$. Suppose that $L_f(\bar{x}, \cdot) : \Omega \rightrightarrows Y$ is strongly regular at $\bar{x}$ for $0$ with modulus $\lambda > 0$ and there exist constants $K > 0$ and $0 < p \leq 1$ such that

$$\lambda \| f'(x) - f'(\bar{x} + \tau (x - \bar{x})) \| \leq (K - \tau^p) \| x - \bar{x} \|^p, \quad \forall x \in B(\bar{x}, K), \quad \forall \tau \in [0, 1]. \quad (2.16)$$

Let $r := \min \{\kappa, \left[\frac{(p+1)}{(2p+1)K}\right]^{1/p}\}$, where $\kappa := \sup \{t > 0 : B(\bar{x}, t) \subset \Omega\}$. Then, there exists a radius of convergence $r_\bar{x} > 0$, with $r_\bar{x} \leq r$ such that the sequences with the initial value $x_0 \in B(\bar{x}, r_\bar{x})/\{\bar{x}\}$ and $t_0 = \|\bar{x} - x_0\|$, respectively,

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad t_{k+1} = \frac{K p t_k^{p+1}}{(p+1)(1 - K t_k^p)}, \quad k = 0, 1, \ldots, \quad (2.17)$$

are well defined, $\{t_k\}$ is strictly decreasing, is contained in $(0, r)$ and converges to $0$, $\{x_k\}$ is contained in $B(\bar{x}, r_\bar{x})$ and converges to the point $\bar{x}$ which is a unique solution of $f(x) + F(x) \ni 0$ in $B(\bar{x}, \bar{\sigma})$, where $\bar{\sigma} \leq \min \{r_\bar{x}, \left[\frac{(p+1)}{K}\right]^{1/p}\}$. Moreover, $\{t_{k+1}/t_k^{1+p}\}$ is strictly decreasing, $t_{k+1}/t_k^{1+p} < (Kp)/[(1 + p)(1 - K\|\bar{x} - x_0\|^p)]$ and

$$\|\bar{x} - x_{k+1}\| \leq \frac{K p}{(p+1)(1 - K t_k^p)} \|\bar{x} - x_k\|^{p+1} \leq \frac{K p \|\bar{x} - x_k\|^{p+1}}{(p+1)(1 - K\|x_0 - \bar{x}\|^p)}, \quad k = 0, 1, \ldots.$$

26
If, additionally, \([p+1]/((2p+1)K)]^{1/p} < \kappa\), then \(r_\bar{x} = [(p+1)/((2p+1)K)]^{1/p}\) is the largest radius of convergence.

**Proof.** Using (2.16), we can immediately prove that \(f, \bar{x}\) and \(\psi : [0, \kappa) \to \mathbb{R},\) defined by \(\psi(t) = Kt^{p+1}/(p+1) - t\), satisfy the inequality (1.6) and the conditions \(h_1, h_2\) and \(h_3\) in Theorem 2.1.1. In this case, it is easy to see that \(\rho\) and \(\nu\), as defined in Theorem 2.1.1, satisfy \(\rho = [(p+1)/((2p+1)K)]^{1/p} \leq \nu = [1/K]^{1/p}\) and, as a consequence, \(r := \min\{\kappa, [(p+1)/((2p+1)K)]^{1/p}\}\). Moreover, \(\psi(\rho)/(\psi'(\rho)) - 1 = 1\), \(\psi(0) = \psi([(p+1)/K]^{1/p}) = 0\) and \(\psi(t) < 0\) for all \(t \in (0, [(p+1)/K]^{1/p})\). Also, the sequence \(\{t_k\}\) in Theorem 2.1.1 is given by (2.17) and satisfies

\[
\frac{t_{k+1}}{t_k} = \frac{Kp}{(p+1)(1-Kt_k^p)} < \frac{Kp}{(p+1)[1-K\|x_0 - \bar{x}\|^p]}, \quad k = 0, 1, \ldots.
\]

Therefore, the result follows by invoking Theorem 2.1.1.

**Remark 2.3.2** Theorem 2.3.1 contains, as particular cases, some results on Newton’s method as we can see in Rall [50], and Traub and Wozniakowski [60].

### 2.3.2 Under Smale’s-type condition

In this section, we assume that \(f\) is an analytic function and using the ideas of [2], we present a version of the classical convergence theorem for Newton’s method for solving the generalized equation (2). The classical version appeared in corollary of Proposition 3 pp. 195 of Smale [59], see Proposition 1 pp. 157 and Remark 1 pp. 158 of Blum, Cucker, Shub, and Smale [6] and [20]. In [1, Theorem 4.2] appears a version of this theorem for metrically regular mapping, which in some sense will be improved in our result below. For stating the result, we need the following definition. Let \(\Omega \subseteq \mathbb{X}\) and \(f : \Omega \to \mathbb{Y}\) be an analytic function. The \(n\)-th derivative of \(f\) at \(x\) is an \(n\)-th multilinear map \(f^n(x) : \mathbb{X} \times \ldots \times \mathbb{X} \to \mathbb{X}\), and its norm is defined by

\[
\|f^n(x)\| = \sup \{\|f^n(x)(v_1, \ldots, v_n)\| : v_1, \ldots, v_n \in \mathbb{X}, \ |v_i| \leq 1, i = 1, \ldots, n\}.
\]

**Theorem 2.3.3** Let \(\mathbb{X}, \mathbb{Y}\) be Banach spaces, \(\Omega \subseteq \mathbb{X}\) an open set and \(f : \Omega \to \mathbb{Y}\) be an analytic function, \(F : \mathbb{X} \rightrightarrows \mathbb{Y}\) be a set-valued mapping with closed graph and \(\bar{x} \in \Omega\). Suppose that \(L_f(\bar{x}, \cdot) : \Omega \rightrightarrows \mathbb{Y}\) is strongly regular at \(\bar{x}\) for 0 with modulus \(\lambda > 0\). Assume that

\[
\gamma := \sup_{n > 1} \left\| \frac{\lambda f^n(\bar{x})}{n!} \right\|^{1/(n-1)} < +\infty.
\]  

(2.18)

Let \(r := \min\{\kappa, (5 - \sqrt{17})/(4\gamma)\}\), where \(\kappa := \sup\{t > 0 : B(\bar{x}, t) \subseteq \Omega\}\). Then, there exists \(r_{\bar{x}} > 0\) with \(r_{\bar{x}} \leq r\) such that the sequences with initial value \(x_0 \in B(\bar{x}, r_{\bar{x}})/\{\bar{x}\}\) and
If twice continuously differentiable, and its proof is similar to Lemma 5.3 of [2]. The second is easier to check than the condition (1.6), whenever the functions under consideration are decreasing, as defined in Theorem 2.1.1, satisfy

\[ \|x_{k+1} - \bar{x}\| \leq \frac{\gamma}{2(1 - \gamma t_k^2) - 1} \|x_k - \bar{x}\|^2 \leq \frac{\gamma}{2(1 - \gamma \|x_0 - \bar{x}\|)^2 - 1} \|x_k - \bar{x}\|^2, \quad k = 0, 1, \ldots \]

If, additionally, \((5 - \sqrt{17})/(4\gamma) < \kappa\), then \(r_x = (5 - \sqrt{17})/(4\gamma)\) is the largest radius of convergence.

To prove Theorem 2.3.3, we need the following results. The first gives us a condition that is easier to check than the condition (1.6), whenever the functions under consideration are twice continuously differentiable, and its proof is similar to Lemma 5.3 of [2]. The second gives a relationship between the second derivatives \(f''\) and \(\psi''\), which allows us to show that \(f\) and \(\psi\) satisfy (1.6), and its proof is along the lines of Lemma 22 of [22].

**Lemma 2.3.4** If \(f : \Omega \subset \mathbb{X} \rightarrow \mathbb{Y}\) is an analytic function, \(\bar{x} \in \Omega\) and \(B(\bar{x}, 1/\gamma) \subset \Omega\), where \(\gamma\) is defined in (2.18), then for all \(x \in B(\bar{x}, 1/\gamma)\), \(\|f''(x)\| \leq 2\gamma/(1 - \gamma \|x - \bar{x}\|)^3\) is satisfied.

**Lemma 2.3.5** Let \(\mathbb{X}, \mathbb{Y}\) be Banach spaces, \(\Omega \subset \mathbb{X}\) be an open set, and \(f : \Omega \rightarrow \mathbb{Y}\) be twice continuously differentiable. Let \(\bar{x} \in \Omega\), \(R > 0\) and \(\kappa = \sup\{t \in [0, R) : B(\bar{x}, t) \subset \Omega\}\). Let \(\lambda > 0\) and \(\psi : [0, R) \rightarrow \mathbb{R}\) be twice continuously differentiable such that \(\lambda \|f''(x)\| \leq \psi''(\|x - \bar{x}\|)\), for all \(x \in B(\bar{x}, \kappa)\), then \(f\) and \(\psi\) satisfy (1.6).

**[Proof of Theorem 2.3.3]**. Let \(\psi : [0, 1/\gamma) \rightarrow \mathbb{R}\) be defined by \(\psi(t) = t/(1 - \gamma t) - 2t\). It is easy to see that \(\psi\) is analytic and \(\psi(0) = 0\), \(\psi'(t) = 1/(1 - \gamma t)^2 - 2\), \(\psi''(0) = -1\), \(\psi''(t) = 2\gamma/(1 - \gamma t)^3\). Moreover, \(\psi'\) is convex. Hence, \(\psi\) satisfies h1, h2 and h3. Now, we combine Lemma 2.3.5 with Lemma 2.3.4, to conclude that \(f\) and \(\psi\) satisfy (1.6). The constants, \(\nu\), \(\rho\) and \(r\), as defined in Theorem 2.1.1, satisfy

\[ \rho = \frac{5 - \sqrt{17}}{4\gamma} < \nu = \frac{\sqrt{2} - 1}{\sqrt{2}\gamma} < \frac{1}{\gamma}, \quad r = \min \left\{ \kappa, \frac{5 - \sqrt{17}}{4\gamma} \right\} \]

Moreover, \(\psi(\rho)/(\rho \psi'(\rho)) - 1 = 1\), \(\psi(0) = \psi(1/(2\gamma)) = 0\) and \(\psi(t) < 0\) for \(t \in (0, 1/(2\gamma))\). Also, \(\{t_k\}\) satisfy

\[ t_{k+1}/t_k^2 = \frac{\gamma}{2(1 - \gamma t_k^2) - 1} < \frac{\gamma}{2(1 - \gamma \|x_0 - \bar{x}\|)^2 - 1}, \quad k = 0, 1, \ldots \]

Therefore, the result follows by considering Remark 2.1.2 and by applying Theorem 2.1.1.
2.3.3 Under Nesterov-Nemirovskii’s condition

In this section, we show a corresponding theorem to Theorem 2.1.1 under the Nesterov-Nemirovskii condition. See for instance [45].

**Theorem 2.3.6** Let \( \Omega \subset \mathbb{R}^n \) be a convex set, \( \bar{x} \in \Omega \), and \( f \in A_\lambda \), where \( A_\lambda \) is defined in Example 1.2.5. Let \( F : \mathbb{R}^n \rightharpoonup \mathbb{R}^n \) be a set-valued mapping with closed graph. Suppose that \( L_f(\bar{x}, \cdot) : \Omega \rightharpoonup Y \) is strongly regular at \( \bar{x} \) for \( 0 \) with modulus \( \lambda > 0 \), and \( W_1(x_0) := \{ x \in \mathbb{R}^n : \| x - x_0 \|_x < 1 \} \subseteq \Omega \). Let \( r := \min \{ \kappa, (5 - \sqrt{17})/4 \} \), where \( \kappa := \sup \{ t > 0 : B(\bar{x}, t) \subseteq \Omega \} \). Then, there exists \( r_{\bar{x}} > 0 \) with \( r_{\bar{x}} \leq r \) such that the sequences with initial value \( x_0 \in B(\bar{x}, r_{\bar{x}}) \) and \( t_0 = \| \bar{x} - x_0 \| \), respectively,

\[
\begin{align*}
  f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1}) &\ni 0, \quad t_{k+1} = \frac{t_k^2}{1 - 2(1 - t_k)^2}, \quad k = 0, 1, \ldots , \\
\| x_{k+1} - \bar{x} \| &\leq \frac{1}{2(1 - t_k)^2 - 1} \| x_k - \bar{x} \|^2 \leq \frac{1}{2(1 - \| x_0 - \bar{x} \|^2 - 1) - 1} \| x_k - \bar{x} \|^2, \quad k = 0, 1, \ldots .
\end{align*}
\]

are well defined, \( \{ t_k \} \) is strictly decreasing, is contained in \((0, r)\) and converges to \(0\), \( \{ x_k \} \) is contained in \( B(\bar{x}, r_{\bar{x}}) \) and converges to the point \( \bar{x} \) which is a unique solution of \( f(x) + F(x) \ni 0 \) in \( B(\bar{x}, \bar{\sigma}) \), where \( 0 < \bar{\sigma} \leq \min \{ r_{\bar{x}}, 1/2 \} \). Moreover, \( \{ t_{k+1}/t_k^2 \} \) is strictly decreasing, \( t_{k+1}/t_k^2 < 1/[2(1 - \| x_0 - \bar{x} \|^2 - 1)] \), for \( k = 0, 1, \ldots \) and

If, additionally, \( (5 - \sqrt{17})/4 < \kappa \), then \( r_{\bar{x}} = (5 - \sqrt{17})/4 \) is the largest radius of convergence.
Chapter 3

Kantorovich’s theorem on Newton’s method for solving strongly regular generalized equation

In this chapter, our objective is to state and prove Kantorovich’s theorem for Newton’s method for solving (2). To state the theorem, we need to set some important constants. We refer to the real numbers

\[ r_{x_1} > 0, \quad r_0 > 0, \quad r_{x_0} > 0, \]  

(3.1)
as the three constants given by Corollary 1.1.4.

3.1 Kantorovich’s theorem

The statement of the main result in this chapter is:

Theorem 3.1.1 Let \( X, Y \) be Banach spaces, \( \Omega \subset X \) be open, \( f : \Omega \to Y \) be continuous with Fréchet derivative \( f' \) continuous, and \( F : X \rightrightarrows Y \) be a set-valued mapping with a closed graph. Assume that \( L_f(x_0) : \Omega \rightrightarrows Y \) is strongly regular at \( x_1 \in \Omega \) for 0 with modulus \( \lambda > 0 \) and there exists \( \psi : [0, R) \to \mathbb{R} \) a majorant function for \( f \). Moreover, suppose that

\[ \|x_1 - x_0\| \leq \psi(0). \]  

(3.2)

Additionally, for the constants \( r_0 \) and \( r_{x_0} \) fixed in (3.1), suppose that the following inequalities hold:

\[ t_* \leq r_{x_0}, \quad \frac{\psi''(t_*)}{2\lambda} \psi(0)^2 < r_0. \]  

(3.3)
Then, the sequences generated by Newton’s method for solving \(0 \in f(x) + F(x)\) and \(\psi(t) = 0\), with starting point \(x_0\) and \(t_0 = 0\), defined respectively by,

\[
x_{k+1} := L_f(x_k, 0)^{-1} \cap B(x_1, r_{x_1}), \quad t_{k+1} = t_k - \psi(t_k)/\psi'(t_k), \quad k = 0, 1, \ldots ,
\]

are well defined, \(\{t_k\}\) is strictly increasing, \(\{t_k\} \subset (0, t_*)\) and converges to \(t_*\), and \(\{x_k\} \subset B(x_0, t_*)\) and converges to \(x_* \in B[x_0, t_*]\), which is the unique solution of \(0 \in f(x) + F(x)\) in \(B[x_0, t_*] \cap B[x_1, r_{x_1}]\). Moreover, \(\{x_k\}\) and \(\{t_k\}\) satisfy

\[
\|x_* - x_k\| \leq t_* - t_k, \quad \|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\|x_* - x_k\|^2,
\]

for all \(k = 0, 1, \ldots\), and the sequences \(\{x_k\}\) and \(\{t_k\}\) converge \(Q\)-linearly as follows

\[
\|x_* - x_{k+1}\| \leq \frac{1}{2}\|x_* - x_k\|, \quad t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k), \quad k = 0, 1, \ldots .
\]

Additionally, if the following condition holds

\textbf{a4)} \(\psi'(t_*) < 0\),

then the sequences, \(\{x_k\}\) and \(\{t_k\}\) converge \(Q\)-quadratically as follows

\[
\|x_* - x_{k+1}\| \leq \frac{\psi''(t_*)}{-2\psi'(t_*)}\|x_* - x_k\|^2, \quad t_* - t_{k+1} \leq \frac{\psi''(t_*)}{-2\psi'(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \ldots .
\]

In Section 3.2, we present particular instances of Theorem 3.1.1 for the classes of functions in the above examples. Hereafter, we consider that all the assumptions in Theorem 3.1.1 hold.

### 3.1.1 Basic results

In this section, we establish some results about \(\psi : [0, R) \to \mathbb{R}\) and some relationships with \(f + F\). We begin by reminding that Proposition 3 of [27] states that \(\psi\) has the smallest root \(t_* \in (0, R)\), is strictly convex, \(\psi(t) > 0\), and \(\psi'(t) < 0\), for all \(t \in [0, t_*)\). Moreover, \(\psi'(t_*) \leq 0\) and \(\psi'(t_*) < 0\) if, and only if, there exists \(t \in (t_*, R)\) such that \(\psi(t) < 0\). Since \(\psi'(t) < 0\) for all \(t \in [0, t_*)\), the Newton iteration of \(\psi\) is well defined in \([0, t_*)\). Let us call it \(n_\psi : [0, t_*) \to \mathbb{R}\) such that

\[
n_\psi(t) = t - \frac{\psi(t)}{\psi'(t)}.
\]

The next result is used to obtain the convergence rate of the sequence generated by Newton’s method for solving \(\psi(t) = 0\). Its proof can be found in [27, Proposition 4].
Lemma 3.1.2 There following statements hold: \( n_\psi(t) \in [0,t_\ast), t < n_\psi(t), \) and \( t_\ast - n_\psi(t) \leq (t_\ast - t)/2, \) for \( t \in [0,t_\ast). \) Moreover, the mapping \([0,t_\ast) \ni -\psi(t)'(t) \in [0, +\infty) \) decreases. If \( \psi \) also satisfies a4, then \( t_\ast - n_\psi(t) \leq [D^-\psi'(t_\ast)/(-2\psi'(t_\ast))](t_\ast - t)^2, \) for \( t \in [0,t_\ast). \)

Using (3.8), the definition of the sequence \( \{t_k\} \) in (3.4) is equivalent to the following one

\[
t_0 = 0, \quad t_{k+1} = n_\psi(t_k), \quad k = 0, 1 \ldots	ag{3.9}
\]

The next result contains the main convergence properties of the above sequence and its proof, which is a consequence of Lemma 3.1.2, follows the same pattern as the proof of Corollary 2.15 of [23].

Corollary 3.1.3 The sequence \( \{t_k\} \) is well defined, strictly increasing, and \( \{t_k\} \subset [0,t_\ast). \) Moreover, \( \{t_k\} \) converges \( Q \)-linearly to \( t_\ast \) as the second inequality in (3.6). Additionally, if a4 holds, then \( \{t_k\} \) converges \( Q \)-quadratically to \( t_\ast \) as the second inequality in (3.7) and converges \( Q \)-quadratically.

Therefore, we have obtained all the statements about \( \{t_k\} \) in Theorem 3.1.1. Now, we are going to establish some relationships between \( \psi \) and \( f + F. \) The next result is a consequence of Corollary 1.1.4.

Proposition 3.1.4 For any \( x \in B(x_0,t_\ast) \), the mapping \( z \mapsto L_f(x,z)^{-1} \cap B(x_1,r_{x_1}) \) is single-valued from \( B(0,r_0) \) to \( B(x_1,r_{x_1}) \) and there holds

\[
\|L_f(x,u)^{-1} \cap B(x_1,r_{x_1}) - L_f(x,v)^{-1} \cap B(x_1,r_{x_1})\| \leq -\frac{\lambda}{\psi'(|x-x_0|)|u-v|}, \forall u, v \in B(0,r_0).
\]

Proof. Definitions of \( r_{x_1}, r_0, \) and \( r_{x_0} \) in (3.1) together with Corollary 1.1.4 imply that, for any \( x \in B(x_0,r_{x_0}), \) the mapping \( z \mapsto L_f(x,z)^{-1} \cap B(x_1,r_{x_1}) \) is single-valued from \( B(0,r_0) \) to \( B(x_1,r_{x_1}) \) and there holds

\[
\|L_f(x,u)^{-1} \cap B(x_1,r_{x_1}) - L_f(x,v)^{-1} \cap B(x_1,r_{x_1})\| \leq \frac{\lambda(1-\lambda f'(x)-f'(x))}{1-\lambda f'(x)}|u-v|, \tag{3.10}
\]

for all \( u, v \in B(0,r_0). \) Since \( |x-x_0| < t_\ast \) thus \( \psi'(|x-x_0|) < 0. \) Hence, (1.8) together with a1 imply that

\[
\lambda f'(x) - f'(x_0) \leq \psi'(|x-x_0|) - \psi'(0) < 1, \quad \forall x \in B(x_0,t_\ast),
\]

and then, using (3.3), i.e., \( t_\ast \leq r_{x_0}, \) (3.10) and a1, the inequality of the proposition follows.

For stating the next result, we need to define the linearization error of \( f \) at points in \( \Omega, \)

\[
E_f(x,y) := f(y) - [f(x) + f'(x)(y-x)], \quad \forall y, x \in \Omega.	ag{3.11}
\]
In the next result, we bound this error by the linearization error of the function $\psi$, namely, 

$$e_{\psi}(t, u) := \psi(u) - [\psi(t) + \psi'(t)(u - t)], \quad \forall \ t, u \in [0, R].$$  \hfill (3.12)

**Lemma 3.1.5** Consider $x, y \in B(x_0, R)$ and $0 \leq t < v < R$. If $\|x - x_0\| \leq t$ and $\|y - x\| \leq v - t$, then

$$\lambda \|E_f(x, y)\| \leq e_{\psi}(t, v) \|y - x\|^2 \leq \frac{1}{2} \psi''(v)(v - t)^2.$$  \hfill (3.13)

**Proof.** Since $x + \tau(y - x) \in B(x_0, R)$, for all $\tau \in [0, 1]$. The linearization error of $f$ in (3.11) is equivalent to

$$E_f(x, y) = \int_0^1 [f'(x + \tau(y - x)) - f'(x)](y - x)d\tau,$$

which combined with the assumption in (1.8) and after some simple algebraic manipulations yields

$$\lambda \|E_f(x, y)\| \leq \int_0^1 [\psi'(|x - x_0| + \tau\|y - x\|) - \psi'(|x - x_0|)]\|y - x\|d\tau.$$  \hfill (3.14)

Using a2, we know that $\psi'$ is convex. Thus, since $\|x - x_0\| \leq t$, $\|y - x\| < v - t$ and $v < R$, we have

$$\lambda \|E_f(x, y)\| \leq \int_0^1 [\psi'(t + \tau\|y - x\|) - \psi'(t)]\|y - x\|^2 \frac{d\tau}{v - t},$$

which, after performing the integration yields the first inequality in (3.13). Now, we prove the last inequality in (3.13). We know that $\psi'$ is convex and differentiable, thus using (3.12), we conclude that

$$e_{\psi}(t, v) = \int_0^1 [\psi'(t + \tau(t - v)) - \psi'(t)](v - t)d\tau \leq \int_0^1 \psi''(v)\tau(v - t)^2d\tau = \frac{1}{2} \psi''(v)(v - t)^2,$$

which, using the first inequality in (3.13) and considering that $\|y - x\| \leq v - t$, gives the desired inequality. \hfill \blacksquare

Proposition 3.1.4 guarantees that, for $x \in B(x_0, t_*)$, the mapping $z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1})$ is single-valued from $B(0, r_0)$ to $B(x_1, r_{x_1})$. Thus, we define the Newton iteration mapping $N_{f+F} : B(x_0, t_*) \to X$ by

$$N_{f+F}(x) := L_f(x, 0)^{-1} \cap B(x_1, r_{x_1}).$$  \hfill (3.15)

We remark that the definition of $N_{f+F}$ can be equivalently stated as the following inclusions

$$0 \in f(x) + f'(x)(N_{f+F}(x) - x) + F(N_{f+F}(x)), \quad N_{f+F}(x) \in B(x_1, r_{x_1}), \quad x \in B(x_0, t_*).$$  \hfill (3.16)
Therefore, one can apply a single Newton iteration on any \( x \in B(x_0, t_*) \) to obtain \( N_{f+\mathbf{F}}(x) \), which may not belong to \( B(x_0, t_*) \). Thus, this is adequate to ensure the well-definedness of only one Newton iteration. To ensure that Newtonian iterations may be repeated indefinitely or, in particular, invariant on subsets of \( B(x_0, t_*) \), we need some additional results. First, define some subsets of \( B(x_0, t_*) \), in which, as we shall prove, Newton iteration mapping (3.15) are “well behaved”. Define

\[
K(t) := \left\{ x \in \Omega : \|x - x_0\| \leq t, \quad \|L_f(x, 0)^{-1} \cap B(x_1, r_{x_1}) - x\| \leq -\frac{\psi(t)}{\psi'(t)} \right\}, \quad t \in [0, t_*),
\]

and consequently, by using (3.16) and the definition in (1.1), we have

\[
\|N_{f+\mathbf{F}}(x) - x_0\| \leq \|x - x_0\| + \|N_{f+\mathbf{F}}(x) - x\| = \|x - x_0\| + \|L_f(x, 0)^{-1} \cap B(x_1, r_{x_1}) - x\|,
\]

and consequently, by using (3.8) and (3.19), the last inequality becomes

\[
\|N_{f+\mathbf{F}}(x) - x_0\| \leq t - \frac{\psi(t)}{\psi'(t)} = n_\psi(t) < t_*.
\]

For the sake of simplicity with regard to the notations, let \( x_+ = N_{f+\mathbf{F}}(x) \in B(x_1, r_{x_1}) \). Thus, using (3.16) and the definition in (1.1), we have

\[
0 \in L_f(x, x_+) = f(x) + f'(x)(x_+ - x) + \mathbf{F}(x_+).
\]

After some simple manipulations in the last inequality and taking into account (3.11), we obtain that

\[
0 \in -E_f(x, x_+) + f(x_+) + f'(x_+)(x_+ - x_+) + \mathbf{F}(x_+).
\]

Using (1.1), we conclude that the last inclusion is equivalent to \( E_f(x, x_+) \in L_f(x_+, x_+) \), which implies that

\[
x_+ \in L_f(x_+, E_f(x, x_+))^{-1} \cap B(x_1, r_{x_1}).
\]

Proposition 3.1.6 For each \( 0 \leq t < t_* \) we have \( K(t) \subset B(x_0, t_*) \) and \( N_{f+\mathbf{F}}(K(t)) \subset K(n_\psi(t)) \). As a consequence, \( K \subseteq B(x_0, t_*) \) and \( N_{f+\mathbf{F}}(K) \subset K \).

Proof. The first inclusion follows trivially from (3.17). Take \( x \in K(t) \). From (3.17) and (3.8), we have

\[
\|x - x_0\| \leq t, \quad \|L_f(x, 0)^{-1} \cap B(x_1, r_{x_1}) - x\| \leq -\frac{\psi(t)}{\psi'(t)}, \quad t < n_\psi(t) < t_.*
\]

Definition of Newton iteration mapping in (3.15) implies that, for all \( x \in K(t) \), there holds

\[
\|N_{f+\mathbf{F}}(x) - x_0\| \leq \|x - x_0\| + \|N_{f+\mathbf{F}}(x) - x\| = \|x - x_0\| + \|L_f(x, 0)^{-1} \cap B(x_1, r_{x_1}) - x\|,
\]

and consequently, by using (3.8) and (3.19), the last inequality becomes

\[
\|N_{f+\mathbf{F}}(x) - x_0\| \leq t - \frac{\psi(t)}{\psi'(t)} = n_\psi(t) < t_*.
\]

For the sake of simplicity with regard to the notations, let \( x_+ = N_{f+\mathbf{F}}(x) \in B(x_1, r_{x_1}) \). Thus, using (3.16) and the definition in (1.1), we have

\[
0 \in L_f(x, x_+) = f(x) + f'(x)(x_+ - x) + \mathbf{F}(x_+).
\]

After some simple manipulations in the last inequality and taking into account (3.11), we obtain that

\[
0 \in -E_f(x, x_+) + f(x_+) + f'(x_+)(x_+ - x_+) + \mathbf{F}(x_+).
\]

Using (1.1), we conclude that the last inclusion is equivalent to \( E_f(x, x_+) \in L_f(x_+, x_+) \), which implies that

\[
x_+ \in L_f(x_+, E_f(x, x_+))^{-1} \cap B(x_1, r_{x_1}).
\]
Since \( \psi \) has the smallest root \( t_* \in (0, R) \), thus (3.20) implies that \( x_+ \in B[x_0, t_*] \). Now, we prove that
\[
E_f(x, x_+) \in B[0, r_0].
\] (3.22)

Since \( x \in K(t) \), definitions (3.8) and (3.15) together with (3.19) imply that \( t < n_\psi(t) \) and \( \|x_+ - x\| \leq n_\psi(t) - t \). Thus, applying the second inequality in Lemma 3.1.5 with \( y = x_+ \) and \( v = n_\psi(t) \), we conclude that
\[
\lambda \|E_f(x, x_+)\| \leq \frac{1}{2} \psi''(n_\psi(t))(n_\psi(t) - t)^2.
\]

On the other hand, \( \textbf{a}2 \) implies that \( \psi'' \) is increasing and Lemma 3.1.2 together with \( \textbf{a}1 \) gives \( n_\psi(t) - t = -\psi(t)/\psi'(t) \leq -\psi(0)/\psi'(0) = \psi(0) \). Thus, the above inequality becomes
\[
\lambda \|E_f(x, x_+)\| \leq \psi''(t_*) \psi(0)^2/2.
\]

Hence, using (3.3) we have (3.22). Since \( x_+ \in B[x_0, t_*] \), combining (3.21) with (3.22) and the first part of Proposition 3.1.4, we obtain \( x_+ = L_f(x_+, E_f(x, x_+))^{-1} \cap B(x_1, r_{x_1}) \). Thus, the second part of Proposition 3.1.4 implies that
\[
\|L_f(x_+, 0)^{-1} \cap B(x_1, r_{x_1}) - x_+\| \leq \frac{\lambda}{\psi'(|x_+ - x_0|)} \|E_f(x_+, x_+)|
\]

Through \( x_+ = N_{f+F}(x) \), we have, from (3.20) that \( \|x_+ - x_0\| \leq n_\psi(t) \). Then, taking into account that \( \psi' \) is increasing and negative, it follows from the above inequality, Lemma 3.1.5, (3.15), and (3.19) that
\[
\|L_f(x_+, 0)^{-1} \cap B(x_1, r_{x_1}) - x_+\| \leq -\frac{\lambda}{\psi(n_\psi(t))} \|E_f(x_+, x_+)| \leq -\frac{\psi(t, n_\psi(t))}{\psi'(n_\psi(t))} \|x_+ - x\|^2.
\]

On the other hand, using the definitions (3.8) and (3.12), after some manipulations, we conclude that
\[
\psi(n_\psi(t)) = \psi(n_\psi(t)) - [\psi(t) + \psi'(t)(n_\psi(t) - t)] = e_\psi(t, n_\psi(t)).
\]

Since \( x_+ = N_{f+F}(x) \), (3.8), and (3.19) imply that \( \|x - x_+\| \leq n_\psi(t) - t \), the latter inequality becomes
\[
\|L_f(x_+, 0)^{-1} \cap B(x_1, r_{x_1}) - x_+\| \leq -\frac{\psi(n_\psi(t))}{\psi'(n_\psi(t))}.
\]

Therefore, since (3.20) implies that \( \|x_+ - x_0\| \leq n_\psi(t) \) the inclusion \( N_{f+F}(K(t)) \subset K(n_\psi(t)) \) follows. The inclusion \( K \subseteq B(x_0, t_*) \) follows from (3.17) and (3.18). To prove the last inclusion, consider \( x \in K \). Thus, \( x \in K(t) \) for some \( t \in [0, t_*] \). Since \( N_{f+F}(K(t)) \subset K(n_\psi(t)) \), we have \( N_{f+F}(x) \in K(n_\psi(t)) \). Since \( n_\psi(t) \in [0, t_*] \) and using (3.18) we conclude the proof. \( \blacksquare \)
3.1.2 Convergence analysis

To prove the convergence results, which are consequences of the above results, first, we note that the definition (3.15) implies that the sequence \( \{x_k\} \) defined in (3.4), can be formally stated as

\[
x_{k+1} = N_{f+F}(x_k), \quad k = 0, 1, \ldots, \tag{3.23}
\]

or equivalently as,

\[
0 \in f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1}), \quad x_{k+1} \in B(x_1, r_{x_1}), \quad k = 0, 1, \ldots. \tag{3.24}
\]

First, we show that the sequence \( \{x_k\} \) generated by Newton’s method converges to \( x^* \in B[x_0, t_*] \), a solution of the generalized equation (2), and is well behaved with respect to the set defined in (3.17).

**Corollary 3.1.7** The sequence \( \{x_k\} \) is well defined, is contained in \( B[x_0, t_*] \), converges to a point \( x^* \in B[x_0, t_*] \) satisfying \( 0 \in f(x^*) + F(x^*) \). Moreover, \( x_k \in K(t_k) \) and \( \|x^* - x_k\| \leq t_* - t_k, \) for all \( k = 0, 1, \ldots \)

**Proof.** Since the mapping \( x \mapsto L_f(x_0, x) \) is strongly regular at \( x_1 \) for 0, Corollary 1.1.4 implies that \( x_1 = L_f(x_0, 0)^{-1} \cap B(x_1, r_{x_1}) \) and the first Newton iterate is well defined. Thus, \( a_1, (3.2), (3.17) \) and (3.18) yield

\[
\{x_0\} = K(0) \subset K. \tag{3.25}
\]

We know from Proposition 3.1.6 that \( N_{f+F}(K) \subset K \). Thus, using (3.25) and (3.23) we conclude that the sequence \( \{x_k\} \) is well defined and rests in \( K \). From the first inclusion in the second part of Proposition 3.1.6, we have that \( \{x_k\} \subset B(x_0, t_*). \) To prove the convergence, first we prove by induction that

\[
x_k \in K(t_k), \quad k = 0, 1, \ldots \tag{3.26}
\]

The above inclusion, for \( k = 0 \), follows from (3.25). Assume now that \( x_k \in K(t_k) \). Then combining Proposition 3.1.6, (3.23), and (3.8), we conclude that \( x_{k+1} \in K(t_{k+1}) \), which completes the induction proof. Now, using (3.26) and (3.17), combined with (3.23), (3.15), and (3.4), we have

\[
\|x_{k+1} - x_k\| = \|L_f(x_k, 0)^{-1} \cap B(x_1, r_{x_1}) - x_k\| \leq -\frac{\psi(t_k)}{\psi'(t_k)} = t_{k+1} - t_k, \quad k = 0, 1, \ldots, \tag{3.27}
\]

Taking into account that \( \{t_k\} \) converges to \( t_* \), we can easily conclude from the above inequality that

\[
\sum_{k=k_0}^{\infty} \|x_{k+1} - x_k\| \leq \sum_{k=k_0}^{\infty} t_{k+1} - t_k = t_* - t_{k_0} < +\infty.
\]
for any $k_0 \in \mathbb{N}$. Hence, we conclude that $\{x_k\}$ is a Cauchy sequence in $B(x_0, t_*)$ and thus it converges to some $x_\ast \in B[x_0, t_*]$. Therefore, using (3.27) again, we conclude that the inequality in the corollary holds. Now, we show that $x_\ast$ is a solution of $f(x) + F(x) \geq 0$. From (3.24), we conclude

$$(x_{k+1}, -f(x_k) - f'(x_k)(x_{k+1} - x_k)) \in \text{gph } F, \quad k = 0, 1, \ldots.$$ 

Since $f$ and $f'$ are continuous in $\Omega$, $B[x_0, t_*] \subset \Omega$ and $F$ has a closed graph, the last inclusion implies that

$$(x_\ast, -f(x_\ast)) = \lim_{k \to \infty} ((x_{k+1}, -f(x_k) - f'(x_k)(x_{k+1} - x_k)) \in \text{gph } F,$$

which implies that $f(x_\ast) + F(x_\ast) \geq 0$ and the proof is complete. \hfill \blacksquare

Now, we prove that $\{x_k\}$ converges $Q$-linearly to $x^\ast$, the unique solution of (2) in $B[x_0, t_*] \cap B[x_1, r_{x_1}]$, and under $a4$ its $Q$-quadratic convergence. For that, we require the result.

**Lemma 3.1.8** Take $x, y \in B(x_0, R)$ and $0 \leq \psi(0) \leq t < R$. If

$$t < t^*, \quad \|x - x_0\| \leq t, \quad \|y - x_1\| \leq r_{x_1}, \quad \|y - x\| \leq t_* - t, \quad 0 \in f(y) + F(y), \quad (3.28)$$

then the following inequality holds

$$\|y - N_{f+F}(x)\| \leq [t_* - n_\psi(t)]\frac{\|y - x\|^2}{(t_* - t)^2}.$$ 

**Proof.** Since $0 \in f(y) + F(y)$, using (3.11) and (1.1), after some simple manipulations, we obtain that

$$0 \in f(y) + F(y) = E_f(x, y) + L_f(x, y),$$

which implies that $y \in L_f(x, -E_f(x, y))^{-1}$. Now, we prove the following inclusion

$$E_f(x, y) \in B(0, r_0). \quad (3.29)$$

Applying Lemma 3.1.5 with $v = t_*$, and using that $0 \leq \psi(0) \leq t < t_*$ we have

$$\lambda\|E_f(x, y)\| \leq \frac{1}{2} \psi''(t_*)(t_* - t)^2 \leq \frac{1}{2} \psi''(t_*)(t_* - \psi(0))^2.$$ 

On the other hand, Lemma 3.1.2 gives us $t_* - n_\psi(0) \leq t_*/2$, which implies that $t_* - n_\psi(0) \leq n_\psi(0) = \psi(0)$. Therefore, the above equation becomes $\lambda\|E_f(x, y)\| \leq \psi''(t_*)\psi(0)^2/2$, which under assumption in (3.3) gives the desired inclusion in (3.29). Since Proposition 3.1.4 implies that for any $x \in B(x_0, t^*)$, the mapping $z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1})$ is single-valued.
from $B(0, r_0)$ to $B(x_1, r_{x_1})$. Thus, taking into account the third inequality in (3.28), (3.29), and that $y \in L_f(x, -E_f(x, y))^{-1}$, we have $y = L_f(x, -E_f(x, y))^{-1} \cap B(x_1, r_{x_1})$. Therefore, combining (3.15) with the second part of Proposition 3.1.4 we conclude

$$\|y - N_{f+F}(x)\| = \|L_f(x, -E_f(x, y))^{-1} \cap B(x_1, r_{x_1}) - L_f(x, 0)^{-1} \cap B(x_1, r_{x_1})\| \leq -\frac{\lambda \|E_f(x, y)\|}{\psi(t)},$$

and since $t < t^*$, $\|x - x_0\| \leq t$ and $\|y - x\| \leq t_* - t$, we can apply Lemma 3.1.5 with $v = t_*$ to obtain

$$\|y - N_{f+F}(x)\| \leq -\frac{e_\psi(t, t_*) \|y - x\|^2}{\psi'(t)}.$$  

However, owing to $0 \leq t < t_*$ and $\psi'(t) < 0$, using (3.12), (3.8), and $\psi(t_*) = 0$, we have

$$-\frac{e_\psi(t, t_*)}{\psi'(t)} = t_* - t + \frac{\psi(t)}{\psi'(t)} - \frac{\psi(t_*)}{\psi'(t)} = t_* - t + \frac{\psi(t)}{\psi'(t)} = t_* - n_\psi(t),$$

which combined with the last inequality yields the desired result.  

**Corollary 3.1.9** The sequences $\{x_k\}$ and $\{t_k\}$ satisfy the following inequality

$$\|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2} \|x_* - x_k\|^2, \quad k = 0, 1, \ldots \tag{3.30}$$

As a consequence, the sequence $\{x_k\}$ converges $Q$-linearly to the solution $x^*$ as follows

$$\|x_* - x_{k+1}\| \leq \frac{1}{2} \|x_* - x_k\|, \quad k = 0, 1, \ldots \tag{3.31}$$

Additionally, if $\psi$ satisfies $a_4$ then the sequence $\{x_k\}$ converges $Q$-quadratically to $x_*$ as follows

$$\|x_* - x_{k+1}\| \leq \frac{\psi''(t_*)}{2\psi'(t_*)} \|x_* - x_k\|^2, \quad k = 0, 1, \ldots \tag{3.32}$$

**Proof.** We know, from Corollary 3.1.7, that $\{x_k\}$ is well defined, converges to $x_*$, $\|x_k - x_0\| \leq t_k$ and $\|x_* - x_k\| \leq t_* - t_k$, for $k = 0, 1, \ldots$. Since $\{x_k\}$ is well defined, it follows from (3.4) that $x_k \in B(x_1, r_{x_1})$ for $k = 1, 2, \ldots$. Hence, $x_* \in B[x_1, r_{x_1}]$, i.e., $\|x_* - x_1\| \leq r_{x_1}$. Hence, since $a_1$ implies that $t_1 = n_\psi(0) = \psi(0)$ and $\{t_k\}$ is strictly increasing, we can apply Lemma 3.1.8 with $x = x_k$, $y = x_*$ and $t = t_k$ to obtain

$$\|x_* - N_{f+F}(x_k)\| \leq \frac{\|x_* - x_k\|^2}{(t_* - t_k)^2}. $$

Thus, inequality (3.30) follows from the above inequality, (3.23) and (3.9). From the first part in Lemma 3.1.2, (3.9) and Corollary 3.1.7, we have

$$\frac{t_* - t_{k+1}}{t_* - t_k} \leq \frac{1}{2}, \quad \frac{\|x_* - x_k\|}{t_* - t_k} \leq 1.$$
Combining these inequalities with (3.30), we obtain (3.31). Now, assume that \( a_1 \) holds. Then, by Corollary 3.1.3, the second inequality on (3.7) holds, which combined with (3.30) implies that (3.32).

**Corollary 3.1.10** The limit \( x_* \) of the sequence \( \{x_k\} \) is the unique solution of the generalized equation \( f(x) + F(x) \ni 0 \) in \( B[x_0, t_*] \cap B[x_1, r_{x_1}] \).

**Proof.** Corollary 3.1.7 implies that \( \{x_k\} \) is well defined and \( \{x_k\} \) is contained in \( B(x_0, t_*) \), thus it follows from (3.4) that \( x_k \in B(x_0, t_*) \cap B(x_1, r_{x_1}) \) for \( k = 1, 2, \ldots \). Hence \( x_* \in B[x_0, t_*] \cap B[x_1, r_{x_1}] \). Suppose there exists \( y_* \in B[x_0, t_*] \cap B[x_1, r_{x_1}] \) such that \( y_* \) is a solution of \( f(x) + F(x) \ni 0 \). We prove by induction that

**(3.33)**

\[
\|y_* - x_k\| \leq t_* - t_k, \quad k = 0, 1, \ldots
\]

The case \( k = 0 \) is trivial, because \( t_0 = 0 \) and \( y_* \in B[x_0, t_*] \). We assume that the inequality holds for some \( k \). First note that Corollary 3.1.7 implies that \( x_k \in K(t_k) \), for \( k = 0, 1, \ldots \). Thus, from the definition of \( K(t_k) \) we conclude that \( \|x_k - x_0\| \leq t_k \), for \( k = 0, 1, \ldots \). Since \( a_1 \) implies that \( t_1 = n_\psi(0) = \psi(0) \), \( \{t_k\} \) is strictly increasing and \( \|x_k - x_0\| \leq t_k \), we may apply Lemma 3.1.8 with \( x = x_k \), \( y = y_* \) and \( t = t_k \) to obtain

\[
\|y_* - N_{f(x) + F(x)}\| \leq [t_* - n_\psi(t_k)]\|y_* - x_k\|^2 / (t_* - t_k)^2, \quad k = 1, 2, \ldots
\]

Using inductive hypothesis, (3.23) and (3.9), we obtain, from the latter inequality, that (3.33) holds for \( k + 1 \). Since \( x_k \) converges to \( x_* \) and \( t_k \) converges to \( t_* \), from (3.33) we conclude that \( y_* = x_* \). Therefore, \( x_* \) is the unique solution of \( f(x) + F(x) \ni 0 \) in \( B[x_0, t_*] \cap B[x_1, r_{x_1}] \).

**3.2 Special cases**

In this section, we study some special cases of Theorem 3.1.1. There exist some classes of well-known functions \( f \), for which it is possible to obtain \( \psi \) satisfying (1.8). For instance, the classes of functions in Examples 1.2.8, 1.2.12, 1.2.10, and 1.2.11. In this sense, the results obtained in Theorem 3.1.1 unify the convergence analysis of Newton’s method for the classes of generalized equations involving these classes of functions. We begin remarking that Theorem 3.1.1 for \( F \equiv \{0\} \) and \( f \) in the classes of functions in Example 1.2.8, up to some minor adjustments, simplify to the classical Kantorovich’s theorem (see [37], see also [38]). For \( F = N_C \), the normal cone of a convex set \( C \), and \( f \) in the classes of functions in Example 1.2.8, Theorem 3.1.1 is equivalent to [36, Theorem 1]. Finally, under the condition that \( C \) is a polyhedral convex set we can also prove that [13, Theorem 2] follows
from Theorem 3.1.1. In this case, \( \psi : [0, 1/L) \to \mathbb{R} \), defined by \( \psi(t) := (\lambda L/2)t^2 - t + b \), is the function associated to \( f \) satisfying (1.8) at \( x_0 \). Another important convergence result on Newton’s method is the \( \alpha - \) theorem for analytic functions due to S. Smale, see [59] (see also [6] and [1] for a version of this theorem for metrically regular mapping). Following the ideas of [2], then we present a version of this theorem for solving (2).

**Theorem 3.2.1** Let \( \Omega \subset \mathbb{R}^n \), \( x_0 \in \Omega \), \( \lambda > 0 \), and \( f \in \mathcal{S}_\lambda \), where \( \mathcal{S}_\lambda \) is defined in Example 1.2.10. Let \( F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n \) be a set-valued mapping with a closed graph. Suppose that \( L_f(x_0, \cdot) : \Omega \rightrightarrows \mathbb{R}^n \) at \( x_0 \), is strongly regular at \( x_1 \in \Omega \) for 0 with modulus \( \lambda > 0 \), \( B(x_0, 1/\gamma) \subseteq \Omega \) and there exists \( b > 0 \) such that \( \|x_1 - x_0\| \leq b \) and \( b\gamma \leq 3 - 2\sqrt{2} \). Additionally, suppose that for \( r_0 \) and \( r_x \) fixed in (3.1), the conditions

\[
t_* \leq r_{x_0}, \quad \frac{4^3 b^2}{\lambda \left(3 - b\gamma + \sqrt{(b\gamma + 1)^2 - 8b\gamma}\right)^3} < r_0,
\]

hold, where \( t_* = (b\gamma + 1 - \sqrt{(b\gamma + 1)^2 - 8b\gamma})/4\gamma \). Then, the sequence generated by Newton’s method for solving \( f(x) + F(x) \ni 0 \) with starting point \( x_0 \), \( x_{k+1} := L_f(x_k, 0)^{-1} \cap B(x_1, r_{x_1}) \), for all \( k = 0, 1, \ldots \), is well defined, \( \{x_k\} \) is contained in \( B(x_0, t_*) \), and converges to the point \( x_* \), which is the unique solution of \( f(x) + F(x) \ni 0 \) in \( B[x_0, t_*] \cap B[x_1, r_{x_1}] \), where \( r_{x_1} \) is fixed in (3.1). Moreover, \( \{x_k\} \) converges \( Q \)-linearly as follows \( \|x_* - x_{k+1}\| \leq \|x_* - x_k\|/2 \), for all \( k = 0, 1, \ldots \). Additionally, if \( b\gamma < 3 - 2\sqrt{2} \), then \( \{x_k\} \) converges \( Q \)-quadratically as follows

\[
\|x_* - x_{k+1}\| \leq \frac{\gamma}{(1 - \gamma t_*)[2(1 - \gamma t_*)^2 - 1]}\|x_* - x_k\|^2, \quad k = 0, 1, \ldots.
\]

**Proof.** Consider \( \psi : [0, 1/\gamma) \to \mathbb{R} \) defined by \( \psi(t) = t/(1 - \gamma t) - 2t + b \). Note that \( \psi(0) = b > 0 \), \( \psi'(t) = 1/(1 - \gamma t)^2 - 2 \), \( \psi'(0) = -1 \), \( \psi''(t) = 2\gamma/(1 - \gamma t)^3 \), \( \psi(t_*) = 0 \). Moreover, from Example 1.2.10, \( \psi \) satisfies \( a1, a2, a3, a4 \), and (3.3). Therefore, the result follows from the Theorem 3.1.1. \( \blacksquare \)

Following the ideas of [2], with some adjustments, in the next theorem we present a version of Theorem 3.1.1 for \( f \) associated to self-concordant function, see Example 1.2.11.

**Theorem 3.2.2** Let \( \Omega \subset \mathbb{R}^n \) be a convex set, \( x_0 \in \Omega \), \( f \in \mathcal{A}_\lambda \), where \( \mathcal{A}_\lambda \) is defined in Example 1.2.11. Let \( F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n \) be a set-valued mapping with a closed graph. Suppose that \( L_f(x_0, \cdot) : \Omega \rightrightarrows \mathbb{R}^n \) at \( x_0 \), is strongly regular at \( x_1 \in \Omega \) for 0 with modulus \( \lambda > 0 \), \( W_1(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| < \delta \} \subseteq \Omega \) and there exists \( b > 0 \) such that \( \|x_1 - x_0\| \leq b \) and \( b\gamma \leq 3 - 2\sqrt{2} \). Additionally, suppose that for \( r_0 \) and \( r_x \) fixed in (3.1) the conditions

\[
t_* \leq r_{x_0}, \quad \frac{4^3 b^2}{\lambda \left(3 - b + \sqrt{(b + 1)^2 - 8b}\right)^3} < r_0,
\]
hold, where \( t_* = (b + 1 - \sqrt{(b + 1)^2 - 8b})/4 \). Then, the sequence generated by Newton’s method for solving \( f(x) + F(x) \ni 0 \) with starting point \( x_0, \) \( x_{k+1} := L_f(x_k, 0)^{-1} \cap B(x_1, r_{x_1}) \), for all \( k = 0, 1, \ldots, \) is well defined, \( \{x_k\} \) is contained in \( B(x_0, t_*) \), and converges to the point \( x_* \), which is the unique solution of \( f(x) + F(x) \ni 0 \) in \( B[x_0, t_*] \cap B[x_1, r_{x_1}] \), where \( r_{x_1} \) is fixed in (3.1). Moreover, \( \{x_k\} \) converges \( Q \)-linearly as follows

\[
\|x_* - x_{k+1}\| \leq \|x_* - x_k\| / 2, \quad k = 0, 1, \ldots
\]

Additionally, if \( b \gamma < 3 - 2\sqrt{2} \), then \( \{x_k\} \) converges \( Q \)-quadratically as follows

\[
\|x_* - x_{k+1}\| \leq \frac{1}{(1 - t_*)[2(1 - t_*^2) - 1]} \|x_* - x_k\|^2, \quad k = 0, 1, \ldots
\]

Proof. Consider \( \psi : [0, 1) \to \mathbb{R} \) defined by \( \psi(t) = t/(1 - t) - 2t + 2 \). Note that \( \psi(0) = b > 0, \psi'(t) = 1/(1 - t)^2 - 2, \psi''(0) = -1, \psi'''(t) = 2/(1 - t)^3, \psi(t_*) = 0 \). Moreover, from Example 1.2.11, \( \psi \) satisfies \( a_1, a_2, a_3, a_4 \) and (3.3). Therefore, the result follows by applying the Theorem 3.1.1.

We end this section by noting that above theorem can be applied for minimizing self-concordant functions constrained by a convex set \( C \subset \Omega \).
Chapter 4

Inexact Newton method for nonlinear functions with values in a cone

The inexact Newton method, for solving the nonlinear equation (1) was introduced in [11] for denoting any method which, given an initial point $x_0$, generates a sequence $\{x_k\}$ as follows:

$$\|f(x_k) + f'(x_k)(x_{k+1} - x_k)\| \leq \eta_k \|f(x_k)\|, \quad k = 0, 1, \ldots,$$

(4.1)

and $\{\eta_k\}$ is a sequence of forcing terms such that $0 \leq \eta_k < 1$; for others variants of this method see [5,18,28]. In [11] was proven, under suitable assumptions, that $\{x_k\}$ is convergent to a solution with super-linear rate. In [39] numerical issues about this method are discussed.

In this chapter, we extend the inexact Newton method (4.1) for solving the nonlinear inclusion $f(x) \in C$, as any method which, given an initial point $x_0$, generates a sequence $\{x_k\}$ satisfying

$$x_{k+1} = x_k + d_k, \quad d_k \in \arg\min_{d \in \mathcal{X}} \left\{ \|d\| : f(x_k) + f'(x_k)d + r_k \in C \right\}, \quad (4.2)$$

$$\max_{w \in \{-r_k, r_k\}} \left\| T^{-1}_{x_0} w \right\| \leq \theta \left\| T^{-1}_{x_0}[-f(x_k)] \right\|, \quad (4.3)$$

for $k = 0, 1, \ldots$, $0 \leq \theta < 1$ is a fixed suitable tolerance, and

$$T^{-1}_{x_0}(y) := \left\{ d \in \mathcal{X} : f'(x_0)d - y \in C \right\},$$

for $y \in \mathcal{Y}$.

We point out that, if $\theta = 0$ then (4.2)-(4.3) reduces to extended Newton method (11) for solving (3) and, in the case, $C = \{0\}$ it reduces to affine invariant version of (4.1), which was also studied in [28].

It is worth noting that (3) is a particular instance of the following generalized equation

$$f(x) + T(x) \ni 0,$$

(4.4)
when $T(x) \equiv -C$ and $T : \mathbb{X} \rightharpoonup \mathbb{Y}$ is a set-valued mapping. In [18] (see also [7]), Dontchev and Rockafellar proposed the following inexact Newton method for solving (4.4):

$$
(f(x_k) + f'(x_k)(x_{k+1} - x_k) + T(x_{k+1})) \cap R_k(x_k, x_{k+1}) \neq \emptyset, \quad k = 0, 1, \ldots, \quad (4.5)
$$

where $R_k : \mathbb{X} \times \mathbb{X} \rightharpoonup \mathbb{Y}$ is a sequence of set-valued mappings with closed graphs. Note that, in the case, when $T \equiv 0$, and

$$
R_k(x_k, x_{k+1}) \equiv B_{\eta_k \|f(x_k)\|}(0),
$$

where we denote by $B_r(x)$ the closed ball centered at $x$ with radius $r$, the iteration (4.5) reduces to (4.1). We also remark that, in the particular case $T(x) \equiv -C$, the iteration (4.5) has (4.2)-(4.3) as a minimal norm affine invariant version. Therefore, in some sense, our method is a particular case of [18]. However, the analysis presented in [18] is local, i.e., it is made assumption at a solution, while in our analysis we not assume existence of solution. In fact, our aim is to prove a robust Kantorovich’s theorem for (4.2)-(4.3), under assumption $\textbf{H1}$ and an affine invariant majorant condition generalizing $\textbf{H2}$, which in particular, prove existence of solution for (2). Moreover, the analysis presented show that the robust analysis of the inexact Newton’s method for solving nonlinear inclusion problems, under affine Lipschitz-like and affine Smale’s conditions, can be obtained as a special case of the general theory. Besides, for the degenerate cone, which the nonlinear inclusion becomes a nonlinear equation, our analysis retrieves the classical results on semi-local analysis of inexact Newton method; [28].

The analysis of this method, under Lipschitz’s condition and Smale’s condition, are provided as special case.

### 4.1 Convex Process

A set-valued mapping $T : \mathbb{X} \rightharpoonup \mathbb{Y}$ is called \textit{sublinear} or \textit{convex process} when its graph is a convex cone, i.e.,

$$
0 \in T(0), \quad T(\lambda x) = \lambda T(x), \quad \lambda > 0, \quad T(x + x') \supseteq T(x) + T(x'), \quad x, x' \in \mathbb{X}, \quad (4.6)
$$

(sublinear mapping has been extensively studied in [16, 52, 54, 55]). The \textit{domain} and \textit{range} of a sublinear mapping $T$ are defined, respectively, by $\text{dom } T := \{d \in \mathbb{X} : Td \neq \emptyset\}$, and $\text{rge } T := \{y \in \mathbb{Y} : y \in T(x) \text{ for some } x \in \mathbb{X}\}$. The \textit{norm} (or inner norm as is called in [16]) of a sublinear mapping $T$ is defined by

$$
\|T\| := \sup \{\|Td\| : d \in \text{dom } T, \|d\| \leq 1\}, \quad (4.7)
$$

43
where \( \|Td\| := \inf \{\|v\| : v \in Td\} \) for \( Td \neq \emptyset \). We use the convention \( \|Td\| = +\infty \) for \( Td = \emptyset \), it will be also convenient to use the convention \( Td + \emptyset = \emptyset \) for all \( d \in X \). Let \( S, T : X \to Y \) and \( U : Y \to Z \) be sublinear mappings. The scalar multiplication, addition and composition of sublinear mappings are sublinear mappings defined, respectively, by \( (\alpha S)(x) := \alpha S(x) \), \( (S + T)(x) := S(x) + T(x) \), and \( UT(x) := \bigcup \{U(y) : y \in T(x)\} \), for all \( x \in X \) and \( \alpha > 0 \) and the following norm properties there hold \( \|\alpha S\| = |\alpha|\|S\| \), \( \|S + T\| \leq \|S\| + \|T\| \) and \( \|UT\| \leq \|U\|\|T\| \).

**Remark 4.1.1** Note that definition of the norm in (4.7) implies that if \( \text{dom} \ T = X \) and \( A \) is a linear mapping from \( Z \) to \( X \) then \( \|T(-A)\| = \|TA\| \).

Let \( \Omega \subseteq X \) be an open set and \( f : \Omega \to Y \) be a continuously Fréchet differentiable function. The linear map \( f'(x) : X \to Y \) denotes the Fréchet derivative of \( f : \Omega \to Y \) at \( x \in \Omega \). Let \( C \subset Y \) be a nonempty closed convex cone, \( z \in \Omega \) and \( T_z : X \rightharpoonup Y \) a mapping defined as

\[
T_zd := f'(z)d - C. \tag{4.8}
\]

It is well-known that the mappings \( T_z \) and \( T_z^{-1} \) are sublinear with closed graph, \( \text{dom} \ T_z = X \), \( \|T_z\| < +\infty \) and, moreover, \( \text{rge} \ T_z = Y \) if and only if \( \|T_z^{-1}\| < +\infty \) (see Lemma 3 of [19] and Corollary 4A.7, Corollary 5C.2 and Example 5C.4 of [16] ). Note that

\[
T_z^{-1}y := \{d \in X : f'(z)d - y \in C\}, \quad z \in \Omega, \ y \in Y. \tag{4.9}
\]

**Lemma 4.1.2** There holds \( T_z^{-1}f'(v)T_v^{-1}w \subseteq T_z^{-1}w \), for all \( v, z \in \Omega, \ w \in Y \). As a consequence,

\[
\|T_z^{-1}[f'(y) - f'(x)]\| \leq \|T_z^{-1}f'(v)T_v^{-1}[f'(y) - f'(x)]\|, \quad v, x, y, z \in \Omega.
\]

**Proof.** See [19]. \( \blacksquare \)

### 4.2 Inexact Newton method

Our goal here is to state and prove a robust semi-local affine invariant theorem for the inexact Newton method to solve a nonlinear inclusion of the form (3). Some definitions are required to state this theorem.

Let \( X, Y \) be Banach spaces, \( X \) reflexive, \( \Omega \subseteq X \) an open set, and \( f : \Omega \to Y \) be a continuously Fréchet differentiable function. The function \( f \) satisfies Robinson’s Condition at \( x_0 \in \Omega \) if

\[
\text{rge} \ T_{x_0} = Y,
\]

44
where $T_{x_0}: X \Rightarrow Y$ is a sublinear mapping as defined in (4.8). Let $R > 0$ be a scalar constant. A continuously differentiable function $\psi:[0,R) \rightarrow \mathbb{R}$ is a majorant function at a point $x_0 \in \Omega$ for $f$ if

$$B(x_0,R) \subseteq \Omega, \quad \|T^{-1}_{x_0}[f'(y) - f'(x)]\| \leq \psi'(\|x-x_0\| + \|y-x\|) - \psi'(\|x-x_0\|), \quad (4.10)$$

for all $x,y \in B(x_0,R)$ such that $\|x-x_0\| + \|y-x\| < R$ and satisfies the conditions a1, a2 e a3. We also need the following condition on the majorant condition $\psi$, which is considered to hold only when explicitly stated.

\[a5\] \[\psi(t) < 0\] for some $t \in (0,R)$.

Note that the condition $a5$ implies the condition $a3$.

The sequence $\{z_k\}$ generated by the inexact Newton method for solving the inclusion $f(x) \in C$ with starting point $z_0$ and the residual relative error tolerance $0 \leq \theta < 1$ is defined by:

$$z_{k+1} := z_k + d_k, \quad d_k \in \arg\min_{d \in X} \{\|d\| : f(z_k) + f'(z_k)d + r_k \in C\}, \quad \max_{w \in \{-r_k, r_k\}} \|T^{-1}_{x_0}w\| \leq \theta \|T^{-1}_{x_0}[-f(z_k)]\|,$$

for $k = 0, 1, \ldots$. The statement of the main result in this chapter is:

**Theorem 4.2.1** Let $C \subset Y$ be a nonempty closed convex cone, $R > 0$. Suppose that $x_0 \in \Omega$, $f$ satisfies Robinson’s condition at $x_0$, $\psi$ is a majorant function for $f$ at $x_0$, and

$$\|T^{-1}_{x_0}[-f(x_0)]\| \leq f(0). \quad (4.11)$$

Let $\beta := \sup\{-\psi(t) : t \in [0,R)\}$. Take $0 \leq \rho < \beta/2$ and define the constants

$$\kappa_{\rho} := \sup_{\rho < t < R} \frac{-(\psi(t) + 2\rho)}{|\psi'(\rho)| (t - \rho)}, \quad \lambda_{\rho} := \sup\{t \in [\rho,R) : \kappa_{\rho} + \psi'(t) < 0\}, \quad \tilde{\theta}_{\rho} := \frac{\kappa_{\rho}}{2 - \kappa_{\rho}}. \quad (4.12)$$

Then, for any $\theta \in [0, \tilde{\theta}_{\rho}]$ and $z_0 \in B(x_0,\rho)$, the sequence $\{z_k\}$, is well defined, for any particular choice of each $d_k$,

$$\|T^{-1}_{x_0}[-f(z_k)]\| \leq \left(\frac{1 + \theta^2}{2}\right)^k [\psi(0) + 2\rho], \quad (4.13)$$

$\{z_k\}$ is contained in $B(z_0, \lambda_{\rho})$ and converges to a point $x_\star \in B[x_0, \lambda_{\rho}]$ such that $f(x_\star) \in C$. Moreover, if

\[a6\] $\lambda_{\rho} < R - \rho$,  

45
then the sequence \( \{z_k\} \) satisfies, for \( k = 0, 1, \ldots \),
\[
\|z_k - z_{k+1}\| \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta}{2} \frac{D^\prime(\lambda + \rho)}{|\psi(\lambda + \rho)|} \|z_k - z_{k-1}\| + \theta \frac{2|\psi'(\rho)| + \psi'(\lambda + \rho)}{|\psi'(\lambda + \rho)|} \right] \|z_k - z_{k-1}\|.
\]
(4.14)

If, additionally, \( 0 \leq \theta < \frac{1 - (\kappa + \rho)}{\sqrt{4(\kappa + \rho)^2 + \kappa(4 + \kappa)}} \left[ \frac{1}{4 + \kappa} \right] \), then \( \{z_k\} \) converges Q-linearly as follows
\[
\limsup_{k \to \infty} \frac{\|x^* - z_{k+1}\|}{\|x^* - z_k\|} \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta}{2} + \frac{2\theta}{\kappa} \right], \quad k = 0, 1, \ldots \tag{4.15}
\]

\textbf{Remark 4.2.2} In Theorem 4.2.1 if \( \theta = 0 \) we obtain the exact Newton method as in [19] and its convergence properties. Now, taking \( \theta = \theta_k \) in each iteration and letting \( \theta_k \) approach zero as \( k \) approaches infinity, inequality (4.14) implies that the sequence \( \{z_k\} \) converges to the solution of (3) at an asymptotic superlinear rate. If \( C = \{0\} \) we obtain the inexact Newton method as in [28] and its convergence properties are similar.

Henceforth, we assume that the assumption on Theorem 4.2.1 holds, except for \( a6 \), which is considered to hold only when explicitly stated.

\subsection*{4.2.1 Preliminary results}

We first prove Theorem 4.2.1 for the case \( \rho = 0 \) and \( z_0 = x_0 \). In order to simplify the notation in the case \( \rho = 0 \), we use \( \kappa, \lambda \) and \( \theta \) instead of \( \kappa_0, \lambda_0, \) and \( \tilde{\theta}_0 \), respectively:
\[
\kappa := \sup_{0 < t < R} \frac{-f(t)}{t}, \quad \lambda := \sup\{t \in [0, R) : \kappa + f'(t) < 0\}, \quad \tilde{\theta} := \frac{\kappa}{2 - \kappa}. \tag{4.16}
\]

\textbf{Majorant function}

In this section we prove the main results of the majorant function. Define
\[
t_* = \min \psi^{-1}(\{0\}), \quad \bar{t} = \sup\{t \in [0, R) : \psi'(t) < 0\}.
\]

Then we have the following remark about the above constants. This remark was proven in [28, Proposition 2.4]:

\textbf{Remark 4.2.3} For \( \kappa, \lambda, \theta \) as in (4.16) it holds that \( 0 < \kappa < 1, 0 < \theta < 1 \) and \( t_* < \lambda \leq \bar{t} \). Moreover, \( \psi'(t) + \kappa < 0 \), for \( t \in [0, \lambda) \) and \( \inf_{0 \leq t \leq R}(\psi(t) + \kappa t) = \lim_{t \to \lambda-}(\psi(t) + \kappa t) = 0 \).

Next remark was proven in [28, Propositions 2.3 and 5.2] and [27, Proposition 3].
Remark 4.2.4 If \( \psi \) satisfies a5 then \( \psi'(t) < 0 \) for any \( t \in [0, \bar{t}] \), \( 0 < t_* < \bar{t} \leq R \), \( \beta = -\lim_{t \to \bar{t}} \psi(t) \), \( 0 < \beta < \bar{t} \) and if \( 0 \leq \rho < \beta/2 \) then \( \rho < \bar{t}/2 < \bar{t} \) and \( \psi'(\rho) < 0 \).

Take \( 0 \leq \theta \) and \( 0 \leq \varepsilon \). We need the following auxiliary mapping, which is associated with the inexact Newton iteration applied to the majorant function, \( n_\theta : [0, \bar{t}] \times [0, \infty) \to \mathbb{R} \times \mathbb{R} \),

\[
n_\theta(t, \varepsilon) := \left( t - (1 + \theta)\frac{\psi(t) + \varepsilon}{\psi'(t)}, \varepsilon + 2\theta(\psi(t) + \varepsilon) \right), \tag{4.17}
\]

The following auxiliary set is important for establishing the convergence of the inexact Newton sequence associated with the majorant function

\[
\mathcal{A} := \{(t, \varepsilon) \in \mathbb{R} \times \mathbb{R} : 0 \leq t < \lambda, 0 \leq \varepsilon \leq \kappa t, 0 < \psi(t) + \varepsilon\}. \tag{4.18}
\]

The following lemma was proven in [28, Lemma 4.2].

Lemma 4.2.5 If \( 0 \leq \theta \leq \tilde{\theta} \), \( (t, \varepsilon) \in \mathcal{A} \) and \( (t_+, \varepsilon_+) := n_\theta(t, \varepsilon) \), that is, \( t_+ := t - (1 + \theta)(\psi(t) + \varepsilon)/\psi'(t) \) and \( \varepsilon_+ := \varepsilon + 2\theta(\psi(t) + \varepsilon) \), then \( n_\theta(t, \varepsilon) \in \mathcal{A} \), \( t < t_+ \) and \( \varepsilon \leq \varepsilon_+ \). Moreover, \( \psi(t_+) + \varepsilon_+ < [(1 + \theta^2)/2](\psi(t) + \varepsilon) \).

We need the following result relating to the linearization error, for proving it see [28, Lemma 3.3].

Lemma 4.2.6 If \( 0 \leq b \leq t \), \( 0 \leq a \leq s \) and \( t + s < R \), then there holds:

\[
e_\psi(a + b, b) \leq \max \left\{ e_\psi(t + s, t), \frac{1}{2} \frac{\psi'(t + s) - \psi'(t)}{s} a^2 \right\}, \quad s \neq 0.
\]

Relationships between the majorant and nonlinear functions

In this section, we present the main relationships between the majorant function \( \psi \) and the nonlinear function \( f \) we need for proving Theorem 4.2.1. Note that Robinson’s condition, namely, \( \text{rge} T_{x_0} = \mathbb{Y} \) implies that \( \text{dom} T_{x_0}^{-1} = \mathbb{Y} \).

Proposition 4.2.7 If \( \|x - x_0\| \leq t < \bar{t} \) then \( \text{dom}[T_x^{-1}f'(x_0)] = \mathbb{X} \) and there holds \( \|T_x^{-1}f'(x_0)\| \leq -1/\psi'(t) \). As a consequence, \( \text{rge} T_x = \mathbb{Y} \).

Proof. See [19, Proposition 12].

Newton’s iteration at a point \( x \in \Omega \) happens to be a solution of the linearization of the inclusion \( f(y) \in C \) at such a point, namely, a solution of the linear inclusion \( f(x) + f'(x)(x - y) \in C \). Thus, we study the linearization error of \( f \) at a point in \( \Omega \). We bound the error \( E_f(y, x) \) by the error \( e_\psi(v, t) \), i.e., the error in the linearization on the majorant function associated to \( f \).
Lemma 4.2.8 If $x, y \in X$ and $\|x - x_0\| + \|y - x\| < R$ then
$$\|T_{x_0}^{-1}E_f(y, x)\| \leq e_\psi(\|x - x_0\| + \|y - x\|, \|x - x_0\|).$$

Proof. As $x, y \in B(x_0, R)$ and the ball is convex $x + \tau(y - x) \in B(x_0, R)$, for all $\tau \in [0, 1]$. Since, by assumption, $\text{rge } T_{x_0} = Y$ we obtain that $\text{dom } T_{x_0}^{-1} = Y$. Thus, using that $f'(z)$ is a linear mapping for each $z \in X$, we conclude
$$\|T_{x_0}^{-1}([f'(x + \tau(y - x)) - f'(x)](y - x)) \| \leq \|T_{x_0}^{-1}[f'(x + \tau(y - x)) - f'(x)]\| \|y - x\|,$$
for all $\tau \in [0, 1]$. Hence, as $\psi$ is a majorant function for $f$ at $x_0$, using (4.10) and the last inequality we have
$$\|T_{x_0}^{-1}([f'(x + \tau(y - x)) - f'(x)](y - x)) \| \leq [\psi'(\|x - x_0\| + \tau \|y - x\|) - \psi'(\|x - x_0\|)] \|y - x\|,$$
for all $\tau \in [0, 1]$. Thus, since $\text{dom } T_{x_0}^{-1} = Y$, we apply Lemma 2.1 of [41] with $U = T_{x_0}^{-1}$ and the functions $G(\tau)$ and $g(\tau)$ equal the expressions in the last inequality, in parentheses on the left and right hand sides, respectively, obtaining
$$\|T_{x_0}^{-1} \int_0^1 [f'(x + \tau(y - x)) - f'(x)](y - x) \, d\tau \| \leq \int_0^1 [\psi'(\|x - x_0\| + \tau \|y - x\|) - \psi'(\|x - x_0\|)] \|y - x\| \, d\tau,$$
which, after integrating the right hand side, taking into account the definition of $e_\psi(v, t)$ and that the error $E_f(y, x)$ is equivalent to
$$E_f(y, x) = \int_0^1 [f'(x + \tau(y - x)) - f'(x)](y - x) \, d\tau,$$
yields the desired inequality. $\blacksquare$

Lemma 4.2.9 If $x, y \in X$ and $\|x - x_0\| + \|y - x\| < R$ then
$$\|T_{x_0}^{-1}[-E_f(y, x)]\| \leq e_\psi(\|x - x_0\| + \|y - x\|, \|x - x_0\|).$$

Proof. To prove this lemma we follow the same arguments used in the proof of Lemma 4.2.8, by taking into account Remark 4.1.1. $\blacksquare$

Corollary 4.2.10 If $x, y \in X$, $\|x - x_0\| \leq t$, $0 < \|y - x\| \leq s$ and $s + t < R$ then
$$\max \{\|T_{x_0}^{-1}[-E_f(y, x)]\|, \|T_{x_0}^{-1}E_f(y, x)\|\} \leq \max \left\{e_\psi(t + s, t), \frac{1}{2} \psi'(s + t) - \psi'(t), \frac{s}{s} \|y - x\|^2 \right\}.$$
Proof. The results follow by direct combination of the Lemmas 4.2.8, 4.2.9 and 4.2.6 by taking \( b = \| x - x_0 \| \) and \( a = \| y - x \| \).

Lemma 4.2.11 If \( x \in \mathbb{X} \) and \( \| x - x_0 \| \leq t < R \) then \( \| T_{x_0}^{-1}f'(x) \| \leq 2 + \psi'(t) \).

Proof. First, we use the definition of sublinear mapping in (4.6) to obtain

\[
T_{x_0}^{-1}f'(x) \supseteq T_{x_0}^{-1}[f'(x) - f'(x_0)] + T_{x_0}^{-1}f'(x_0).
\]

Hence, taking into account properties of the norm, the above inclusion enables us to conclude that

\[
\| T_{x_0}^{-1}f'(x) \| \leq \| T_{x_0}^{-1}[f'(x) - f'(x_0)] \| + \| T_{x_0}^{-1}f'(x_0) \|.
\]

Since \( T_{x_0}^{-1}f'(x_0) \supseteq f'(x_0)^{-1}f'(x_0) \) we have \( \| T_{x_0}^{-1}f'(x_0) \| \leq 1 \). Thus, using assumption (4.10), the last inequality becomes

\[
\| T_{x_0}^{-1}f'(x) \| \leq \psi'(\| x - x_0 \|) - \psi'(0) + 1.
\]

Therefore, assumptions \( a_1 \), \( a_2 \) and the last inequality imply the statement of the lemma.

The next result is used to show that the inexact Newton method is robust with respect to the initial iterate, and its proof can be found in [19, Proposition 16].

Proposition 4.2.12 If \( y \in B(x_0, R) \) then \( \| T_{x_0}^{-1}[-f(y)] \| \leq \psi(\| y - x_0 \|) + 2\| y - x_0 \| \).

4.3 Convergence analysis of the inexact Newton Method

In this section we prove Theorem 4.2.1. Prior to this, we need to study the iteration of the inexact Newton method, associated to the function \( f \), and prove Theorem 4.2.1 for the case \( \rho = 0 \) and \( z_0 = x_0 \).

4.3.1 The inexact Newton iteration

The outcome of an inexact Newton iteration is any point satisfying some error tolerance. Hence, instead of a mapping for the inexact Newton iteration, we are concerned with a family of functions, describing all possible inexact iterations. Before defining the inexact
Newton iteration mapping, we need to define the mapping of the inexact Newton step, \( D_{f,C,\theta} : B(x_0, \bar{t}) \to \mathbb{X} \),

\[
D_{f,C,\theta}(x) := \arg\min_{d \in \mathbb{X}} \{ \|d\| : f(x) + f'(x)d + r \in C \}; \quad \max_{w \in \{-r, r\}} \| T_{x_0}^{-1} w \| \leq \theta \| T_{x_0}^{-1}[-f(x)] \| ,
\]

(4.19)

associated to \( f \), \( C \) and \( \theta \). Since \( \mathbb{X} \) is reflexive, the second part of Proposition 3.1.4 guarantees, in particular, that the exact Newton step \( D_{f,C,0}(x) \) is nonempty, for each \( x \in B(x_0, \bar{t}) \). Since \( D_{f,C,0}(x) \subseteq D_{f,C,\theta}(x) \), we conclude \( D_{f,C,\theta}(x) \neq \emptyset \), for \( x \in B(x_0, \bar{t}) \). Therefore, for \( 0 \leq \theta \leq \tilde{\theta} \), we can define \( N_\theta \) the family of inexact Newton iteration mappings, \( N_{f,C,\theta} : B(x_0, \bar{t}) \to \mathbb{X} \),

\[
N_{f,C,\theta}(x) := x + D_{f,C,\theta}(x).
\]

(4.20)

A single Newton iteration can be applied to any \( x \in B(x_0, \bar{t}) \) to obtain the set \( N_{f,C,\theta}(x) \), which may not be contained in \( B(x_0, \bar{t}) \), or even may not be in the domain of \( f \). Therefore, this is sufficient to guarantee the well–definedness of only one iteration. To ensure that the inexact Newton iteration mapping may be repeated indefinitely, we need some additional results. First, define some subsets of \( B(x_0, \bar{t}) \) in which, as we prove, inexact Newton iteration mappings (4.20) are “well behaved”. Define

\[
K(t, \varepsilon) := \{ x \in \mathbb{X} : \|x - x_0\| \leq t, \| T_{x_0}^{-1}[-f(x)] \| \leq \psi(t) + \varepsilon \},
\]

(4.21)

and

\[
\mathcal{K} := \bigcup_{(t, \varepsilon) \in \mathcal{A}} K(t, \varepsilon).
\]

(4.22)

**Proposition 4.3.1** Take \( 0 \leq \theta \leq \tilde{\theta} \) and \( N_{f,C,\theta} \in \mathcal{N}_\theta \). Then, for any \((t, \varepsilon) \in \mathcal{A} \) and \( x \in K(t, \varepsilon) \)

\[
\|y - x\| \leq t_+ - t,
\]

(4.23)

where \( y \in N_{f,C,\theta}(x) \) and \( t_+ \) is the first component of the function \( n_\theta(t, \varepsilon) \) defined in (4.17). Moreover,

\[
N_{f,C,\theta}(K(t, \varepsilon)) \subset K(n_\theta(t, \varepsilon)).
\]

(4.24)

As a consequence,

\[
n_\theta(\mathcal{A}) \subset \mathcal{A}, \quad N_{f,C,\theta}(\mathcal{K}) \subset \mathcal{K}.
\]

(4.25)

**Proof.** Take \( 0 \leq \theta \), \((t, \varepsilon) \in \mathcal{A} \) and \( x \in K(t, \varepsilon) \). Thus, the definitions of the sets \( \mathcal{A} \) in (4.18), \( K(t, \varepsilon) \) in (4.21) together with Lemma 4.2.5 imply that

\[
\|x - x_0\| \leq t < \bar{t}, \quad \| T_{x_0}^{-1}[-f(x)] \| \leq \psi(t) + \varepsilon, \quad t - (1 + \theta) \frac{\psi(t) + \varepsilon}{\psi'(t)} < \lambda \leq R.
\]

(4.26)
Take $y \in N_{f,C,\theta}(x)$ and $r$ as in (4.19). Using the third property of the convex process in (4.6), we have

$$T_x^{-1}[-f(x) - r] \supseteq T_x^{-1}[-f(x)] + T_x^{-1}[-r].$$

Applying Lemma 4.1.2 to each term on the right hand side of the last inclusion, one with $w = -r$, $z = x$, and $v = x_0$, and the other one with $w = -f(x)$, $z = x$, and $v = x_0$, we obtain

$$T_x^{-1}[-f(x) - r] \supseteq T_x^{-1}f'(x_0)T_{x_0}^{-1}[-f(x)] + T_x^{-1}f'(x_0)T_{x_0}^{-1}[-r].$$

Hence, taking the norm on both sides of the last inclusion and using the properties of the norm yields

$$\|T_x^{-1}[-f(x) - r]\| \leq \|T_x^{-1}f'(x_0)\| \|T_{x_0}^{-1}[-f(x)]\| + \|T_x^{-1}f'(x_0)\| \|T_{x_0}^{-1}[-r]\|.$$

Considering that $y - x \in D_{f,C,\theta}(x)$, we obtain that $\|y - x\| = \|T_x^{-1}[-f(x) - r]\|$. Thus, combining the last inequality with Proposition 4.2.7 and the third inequality in (4.26), and after some manipulation taking into account (4.19), we have

$$\|y - x\| \leq -(1 + \theta) \frac{\psi(t) + \varepsilon}{\psi'(t)}, \quad (4.27)$$

which, using definition of $t_+$, is equivalent to (4.23).

Since $\|y - x_0\| \leq \|y - x\| + \|x - x_0\|$, thus (4.27), the first and the last inequality in (4.26) give

$$\|y - x_0\| \leq t - (1 + \theta) \frac{\psi(t) + \varepsilon}{\psi'(t)} < \lambda \leq R. \quad (4.28)$$

On the other hand, the linearization error of $E_f(y, x)$ and the third property of the convex process in (4.6) imply

$$T_{x_0}^{-1}[-f(y)] \supseteq T_{x_0}^{-1}[-E_f(y, x)] + T_{x_0}^{-1}[-f(x) - f'(x)(y - x)].$$

Thus, taking the norm on both sides of the last inclusion and using the triangular inequality we obtain

$$\|T_{x_0}^{-1}[-f(y)]\| \leq \|T_{x_0}^{-1}[-E_f(y, x)]\| + \|T_{x_0}^{-1}[-f(x) - f'(x)(y - x)]\|.$$

Since $y \in N_{f,C,\theta}(x)$ we have $T_{x_0}^{-1}[r] \subset T_{x_0}^{-1}[-f(x) - f'(x)(y - x)]$, where $r$ satisfies $f(x) + f'(x)(y - x) + r \in C$ and (4.19). Then, the last inequality implies

$$\|T_{x_0}^{-1}[-f(y)]\| \leq \|T_{x_0}^{-1}[-E_f(y, x)]\| + \theta \|T_{x_0}^{-1}[-f(x)]\|.$$

The second term on the right hand side of the last inequality is bound by the third inequality in (4.26). Thus, letting $s = -(1 + \theta)(\psi(t) + \varepsilon)/\psi'(t)$, using (4.27), the first and last inequality (4.26), we can apply Corollary 4.2.10 to conclude that

$$\|T_{x_0}^{-1}[-f(y)]\| \leq e_f \left( t - (1 + \theta) \frac{\psi(t) + \varepsilon}{\psi'(t)}, t \right) + \theta(\psi(t) + \varepsilon).$$
Therefore, combining the last inequality with the definition of $E_\psi(v,t)$, we easily obtain that
\[
\|T_{x_0}^{-1}[-f(y)]\| \leq \psi \left( t - (1 + \theta) \frac{\psi(t) + \varepsilon}{\psi'(t)} \right) + \varepsilon + 2\theta(\psi(t) + \varepsilon).
\]
Finally, (4.28), the last inequality, definitions (4.17) and (4.21) prove that the inclusion (4.24) holds.

The inclusions in (4.25) are an immediate consequence of Lemma 4.2.5, (4.24) and the definitions in (4.18) and (4.22). Thus, the proof of the proposition is concluded.

\[ \blacksquare \]

### 4.3.2 Convergence analysis

In this section we prove Theorem 4.2.1. First, we show that the sequence generated by the inexact Newton method is well behaved with respect to the set defined in (4.21).

**Theorem 4.3.2** Take $0 \leq \theta \leq \tilde{\theta}$ and $N_{f,C,\theta} \in N$. For any $(t_0,\varepsilon_0) \in A$ and $y_0 \in K(t_0,\varepsilon_0)$ the sequences
\[
y_{k+1} \in N_{f,C,\theta}(y_k), \quad (t_{k+1},\varepsilon_{k+1}) = n_\theta(t_k,\varepsilon_k), \quad k = 0,1,\ldots, \quad (4.29)
\]
are well defined,
\[
y_k \in K(t_k,\varepsilon_k), \quad (t_k,\varepsilon_k) \in A \quad k = 0,1,\ldots, \quad (4.30)
\]
the sequence \( \{t_k\} \) is strictly increasing and converges to some $\tilde{t} \in (0,\lambda]$, the sequence $\{\varepsilon_k\}$ is non-decreasing and converges to some $\tilde{\varepsilon} \in [0,\kappa\lambda]$,\n\[
\|T_{x_0}^{-1}[-f(y_k)]\| \leq \psi(t_k) + \varepsilon_k \leq \left( \frac{1 + \theta^2}{2} \right)^k (\psi(t_0) + \varepsilon_0), \quad k = 0,1,\ldots, \quad (4.31)
\]
\( \{y_k\} \) is contained in $B(x_0,\lambda)$, converges to a point $x_* \in B[x_0,\lambda]$ such that $f(x_*) \in C$, and satisfies\n\[
\|y_{k+1} - y_k\| \leq t_{k+1} - t_k, \quad \|x_* - y_k\| \leq \tilde{t} - t_k, \quad k = 0,1,\ldots. \quad (4.32)
\]
Moreover, if
\[ a6') \lambda < R, \]
then the sequence $\{y_k\}$ satisfies\n\[
\|y_k - y_{k+1}\| \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta D^{-1}(\lambda)}{2|\psi'(\lambda)|} \|y_k - y_{k-1}\| + \theta \frac{2 + \psi'(\lambda)}{|\psi'(\lambda)|}\right] \|y_k - y_{k-1}\|, \quad k = 0,1,\ldots. \quad (4.33)
\]
If, additionally, $0 \leq \theta < -2(\kappa + 1) + \sqrt{4(\kappa + 1)^2 + \kappa(4 + \kappa)}/(4 + \kappa)$ then $\{y_k\}$ converges $Q$-linearly as follows
\[
\limsup_{k \to \infty} \frac{\|x_* - y_{k+1}\|}{\|x_* - y_k\|} \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta}{2} + \frac{2\theta}{\kappa} \right], \quad k = 0,1,\ldots. \quad (4.34)
\]
Proof. Since \(0 \leq \theta \leq \tilde{\theta}, (t_0, \varepsilon_0) \in \mathcal{A}\) and \(y_0 \in K(t_0, \varepsilon_0)\), an appropriate definition of the sequences \(\{t_k, \varepsilon_k\}\) and \(\{y_k\}\), as defined in (4.29), follows from the last two inclusions (4.25) in Proposition 4.3.1. Moreover, since (4.30) holds for \(k = 0\), using the first inclusion in Proposition 4.3.1, the first inclusion in (4.25) and induction on \(k\), we conclude that (4.30) holds for all \(k\). The first inequality in (4.32) follows from (4.23) in Proposition 4.3.1, (4.29) and (4.30), whereas the first inequality in (4.31) follows from (4.30) and the definition of \(K(t, \varepsilon)\) in (4.21).

The definition of \(\mathcal{A}\) in (4.18) implies \(\mathcal{A} \subset [0, \lambda) \times [0, \kappa\lambda)\). Therefore, using (4.30) and the definition of \(K(t, \varepsilon)\) we have

\[
t_k \in [0, \lambda), \quad \varepsilon_k \in [0, \kappa\lambda), \quad y_k \in B(x_0, \lambda), \quad k = 0, 1, \ldots.
\]

Using (4.18) and Lemma 4.2.5 we conclude that \(\{t_k\}\) is strictly increasing, \(\{\varepsilon_k\}\) is non-decreasing, and the second equality in (4.31) holds for all \(k\). Therefore, in view of the first two above inclusions, \(\{t_k\}\) and \(\{\varepsilon_k\}\) converge, respectively, to some \(t \in (0, \lambda]\) and \(\varepsilon \in [0, \kappa\lambda)\). The convergence of \(\{t_k\}\) to \(t\), together with the first inequality in (4.32) and the inclusion \(y_k \in B(x_0, \lambda)\) implies that \(y_k\) converges to some \(x_\ast \in B[x_0, \lambda]\) and that the second inequality on (4.32) holds for all \(k\). Moreover, taking the limit in (4.31), as \(k\) approaches \(+\infty\), we conclude that

\[
\lim_{k \to +\infty} \|T_{x_0}^{-1}[-f(y_k)]\| = 0.
\]

Thus, there exists \(\{d_k\} \subset X\) such that \(d_k \in T_{x_0}^{-1}[-f(y_k)]\), for all \(k = 0, 1, \ldots\), with \(\lim_{k \to +\infty} d_k = 0\). Since \(d_k \in T_{x_0}^{-1}[-f(y_k)]\), for all \(k = 0, 1, \ldots\), the Definition 4.9 implies that \(f'(x_0)d_k + f(y_k) \in C'\), for all \(k = 0, 1, \ldots\). Hence, letting \(k\) approach \(+\infty\) in the last inclusion and taking into account that \(C\) is closed and \(\{y_k\}\) converges to \(x_\ast\), we conclude that \(f(x_\ast) \in C\).

We next prove (4.33). Since \(y_{k+1} \in N_{f,C,\theta}(y_k)\), for \(k = 0, 1, \ldots\), we have

\[
\|y_{k+1} - y_k\| = \|T_{y_k}^{-1}[-f(y_k) - r_k]\|,
\]

\[
\max_{w \in \{-r_k, r_k\}} \|T_{x_0}^{-1}w\| \leq \theta \|T_{x_0}^{-1}[-f(y_k)]\|,
\]

(4.35)

The third property in (4.6) implies \(T_{y_k}^{-1}[-f(y_k) - r_k] \supseteq T_{y_k}^{-1}[-f(y_k)] + T_{y_k}^{-1}[-r_k]\). Then applying Lemma 4.1.2 twice, once with \(z = y_k, v = x_0,\) and \(w = -f(y_k)\) and, the second time with \(z = y_k, v = x_0,\) and \(w = -r_k\), we obtain that

\[
T_{y_k}^{-1}[-f(y_k) - r_k] \supseteq T_{y_k}^{-1}f'(x_0)T_{x_0}^{-1}[-f(y_k)] + T_{y_k}^{-1}f'(x_0)T_{x_0}^{-1}[-r_k].
\]

Combining the last inclusion with (4.35) and the properties of the norm, we conclude, after some algebra, that

\[
\|y_{k+1} - y_k\| \leq (1 + \theta) \|T_{y_k}^{-1}f'(x_0)\| \|T_{x_0}^{-1}[-f(y_k)]\|.
\]

(4.36)
Using (??), the third property in (4.6) and triangular inequality, after some manipulation, we have
\[ \|T_{x_0}^{-1}[-f(y_k)]\| \leq \|T_{x_0}^{-1}[-E_f(y_k, y_{k-1})]\| + \|T_{x_0}^{-1}[-f'(y_{k-1})(y_k - y_{k-1})]\|. \tag{4.37} \]

On the other hand, because \( y_k \in N_{f,C,\theta}(y_{k-1}) \) we have \( T_{x_0}^{-1}[r_{k-1}] \subset T_{x_0}^{-1}[-f(y_{k-1}) - f'(y_{k-1})(y_k - y_{k-1})] \), where \( r_{k-1} \) satisfies
\[ \|T_{x_0}^{-1}r_{k-1}\| \leq \theta \|T_{x_0}^{-1}[-f(y_{k-1})]\|. \]

Therefore, we have
\[ \|T_{x_0}^{-1}[-f(y_{k-1}) - f'(y_{k-1})(y_k - y_{k-1})]\| \leq \theta \|T_{x_0}^{-1}[-f(y_{k-1})]\|, \tag{4.38} \]
which combined, with the inequalities in (4.36) and (4.37), yields
\[ \|y_{k+1} - y_k\| \leq (1 + \theta) \|T_{y_k}^{-1}f'(x_0)\| \left[ \|T_{x_0}^{-1}[-E_f(y_k, y_{k-1})]\| + \theta \|T_{x_0}^{-1}[-f(y_{k-1})]\| \right]. \tag{4.39} \]

Again using (??), the third property in (4.6), and the triangular inequality, we obtain after some algebra that
\[ \|T_{x_0}^{-1}[-f(y_{k-1})]\| \leq \|T_{x_0}^{-1}E_f(y_k, y_{k-1})\| + \|T_{x_0}^{-1}[-f(y_k)]\| + \|T_{x_0}^{-1}f'(y_{g-1})(y_k - y_{k-1})\|. \]

Combining the last inequality with the inequalities in (4.37) and (4.38) we conclude that
\[ \|T_{x_0}^{-1}[-f(y_{k-1})]\| \leq \frac{1}{1 - \theta} \left[ \|T_{x_0}^{-1}[E_f(y_k, y_{k-1})]\| + \theta \|T_{x_0}^{-1}[-E_f(y_k, y_{k-1})]\| + \|T_{x_0}^{-1}f'(y_{g-1})(y_k - y_{k-1})\| \right]. \]

The inequality in (4.39) combined with the last inequality becomes
\[ \|y_{k+1} - y_k\| \leq \frac{1 + \theta}{1 - \theta} \|T_{y_k}^{-1}f'(x_0)\| \left[ \|T_{x_0}^{-1}[-E_f(y_k, y_{k-1})]\| + \theta \left( \|T_{x_0}^{-1}[E_f(y_k, y_{k-1})]\| + \|T_{x_0}^{-1}f'(y_{g-1})(y_k - y_{k-1})\| \right) \right]. \]

Therefore, combining the last inequality with Proposition 4.2.7, Lemma 4.2.11, and Corollary 4.2.10 with \( x = y_{k-1}, \ y = y_k, \ s = t_k - t_{k-1} \) and \( t = t_{k-1} \), we have
\[ \|y_k - y_{k+1}\| \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta \psi'(t_k) - \psi'(t_{k-1})}{t_k - t_{k-1}} \|y_{k-1} - y_k\| + \theta[2 + \psi'(t_{k-1})] \right] \|y_{k-1} - y_k\|, \tag{4.40} \]
for \( k = 0, 1, \ldots \). Since \( \|y_{k-1} - y_k\| \leq t_k - t_{k-1} \), see (4.32), \( \psi' < -\kappa < 0 \) in \([0, \lambda]\), (4.34) follows from the last inequality. Using a6’ and Theorem 4.1.1 on p. 21 of [33] and taking
into account that \(|\psi'|\) is decreasing in \([0, \lambda]\), \(\psi'\) is increasing in \([0, \lambda]\) and \(\{t_k\} \subset [0, \lambda]\), we obtain that (4.33) follows from the above inequality.

To conclude the proof, it remains to prove that \(\{y_k\}\) converges \(Q\)-linearly as in (4.34). First note that \(\|y_{k-1} - y_k\| \leq t_k - t_{k-1}\) and \(\psi'(t_{k-1}) \leq \psi'(t_k) < 0\). Thus, we conclude from (4.40) that
\[
\|y_k - y_{k+1}\| \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta}{2} + \frac{2\theta}{\kappa} \right] \|y_{k-1} - y_k\|, \quad k = 0, 1, \ldots. \tag{4.41}
\]
which, from Proposition 2 of [23], implies that (4.34) holds. Since \(0 \leq \theta < -2(\kappa + 1) + \sqrt{4(\kappa + 1)^2 + \kappa(4 + \kappa)/(4 + \kappa)}\), the quantity on the right hand side of (4.34) is less than one. Hence, \(\{y_k\}\) converges \(Q\)-linearly, which concludes the proof. 

**Proposition 4.3.3** Let \(R > 0\) and \(\psi : [0, R) \to \mathbb{R}\) be a continuously differentiable function. Suppose that \(x_0 \in \Omega\), \(\psi\) is a majorant function for \(f\) at \(x_0\) and satisfies \(a5\). If \(0 \leq \rho < \beta/2\), then for any \(z_0 \in B(x_0, \rho)\) the scalar function \(g : [0, R - \rho) \to \mathbb{R}\), defined by
\[
g(t) := \frac{-1}{\psi'(\rho)}[\psi(t + \rho) + 2\rho], \tag{4.42}
\]
is a majorant function for \(f\) at \(z_0\) and also satisfies condition \(a5\).

**Proof.** For the prove, see Proposition 17 of [19].

**[Proof of Theorem 4.2.1]** First, we prove Theorem 4.2.1 with \(\rho = 0\) and \(z_0 = x_0\). Note that, from the definition in (4.16), we have
\[
\kappa_0 = \kappa, \quad \lambda_0 = \lambda, \quad \tilde{\theta}_0 = \tilde{\theta}.
\]
The assumption (4.11) implies that \(x_0 \in K(0, 0)\). Since \((t_0, \varepsilon_0) = (0, 0) \in \mathcal{A}\) and \(y_0 = x_0 \in K(0, 0)\), we apply Theorem 4.3.2 with \(z_k = y_k\), for \(k = 0, 1, \ldots\), to conclude that Theorem 4.2.1 holds for \(\rho = 0\) and \(z_0 = x_0\).

We next prove the general case. From Proposition 4.2.4 we have \(\rho < \bar{\ell}\), which implies that \(\|z_0 - x_0\| < \rho < \bar{\ell}\). Thus, we can apply Proposition 4.2.7 to obtain
\[
\|T_{z_0}^{-1}f'(x_0)\| \leq \frac{-1}{\psi'(\rho)}. \tag{4.43}
\]
Moreover, the point \(z_0\) satisfies Robinson's condition, namely,
\[
rge T_{z_0} = \overline{Y}.
\]
Then, using Lemma 4.1.2, the property of the norm, (4.43), and Proposition 4.2.12 with \(y = z_0\) we have
\[
\|T_{z_0}^{-1}[-f(z_0)]\| \leq \|T_{z_0}^{-1}f'(x_0)\|\|T_{z_0}^{-1}[-f(z_0)]\| \leq \frac{-1}{\psi'(\rho)}[\psi(\|z_0 - x_0\|) + 2\|z_0 - x_0\|].
\]

55
Since $\psi' \geq -1$, the function $t \mapsto \psi(t) + 2t$ is (strictly) increasing. Thus, combining this fact with the last inequality, the inequality $\|z_0 - x_0\| < \rho$, and (4.42) we conclude that

$$\|T_{z_0}^{-1}[-f'(z_0)]\| \leq g(0).$$

Proposition 4.3.3 implies that $g$, defined in (4.42), is a majorant function for $f$ at point $z_0$ and also satisfies condition $\textbf{a5}$. Moreover, (4.42) and $\kappa_\rho$, $\lambda_\rho$, and $\tilde{\theta}_\rho$ as defined in (4.12) imply

$$\kappa_\rho = \sup_{0 < t < R - \rho} \frac{-g(t)}{t}, \quad \lambda_\rho = \sup\{t \in [0, R - \rho) : \kappa_\rho + g'(t) < 0\}, \quad \tilde{\theta}_\rho = \frac{\kappa_\rho}{2 - \kappa_\rho},$$

which are the same as (4.12) with $g$ instead of $\psi$, then we can apply Theorem 4.3.2 for $f$ and the majorant function $g$ at point $z_0$ and $\rho = 0$, to conclude that the sequence $\{z_k\}$ is well defined, remains in $B(z_0, \lambda_\rho)$, satisfies (4.13), and converges to some $x_\ast \in B[z_0, \lambda_\rho]$ with $f(x_\ast) \in C$. Furthermore, since

$$g'(t) = \psi'(t + \rho)/|\psi'(\rho)|, \quad D^- g'(t) = D^- \psi'(t + \rho)/|\psi'(\rho)|, \quad t \in [0, R - \rho),$$

after some algebra, we conclude that inequalities (4.14) and (4.15) also hold. Therefore, this concludes the proof of the theorem.

### 4.4 Special cases

In this section we use Theorem 4.2.1 to analyze the convergence of the inexact Newton method for cone inclusion problems under the affine invariant Lipschitz condition and in the setting of Smale’s $\alpha$-theory. To the best of our knowledge, this is the first time the inexact Newton method for cone inclusion problems with a relative error tolerance under Lipschitz’s condition and Smale’s condition are analyzed.

#### 4.4.1 Under affine invariant Lipschitz condition

In this section we present the convergence analysis of the inexact Newton method for cone inclusion problems under affine invariant Lipschitz condition. Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces, $\mathbb{X}$ be reflexive, $\Omega \subseteq \mathbb{X}$ be an open set, $x_0 \in \Omega$, and $L > 0$. A continuously Fréchet differentiable function $f : \Omega \to \mathbb{Y}$ satisfies the affine invariant Lipschitz condition with constant $L$ at $x_0$, if $B(x_0, 1/L) \subseteq \Omega$ and

$$\|T_{x_0}^{-1} [f'(y) - f'(x)]\| \leq L\|x - y\|, \quad x, y \in B(x_0, 1/L).$$
Theorem 4.4.1 Let $C \subseteq Y$ be a nonempty closed convex cone. Suppose that $x_0 \in \Omega$ and $f$ satisfies Robinson’s and the affine invariant Lipschitz condition with constant $L > 0$ at $x_0$ and

$$\|T_{x_0}^{-1}f(x_0)\| \leq b, \quad 0 \leq \theta \leq (1 - \sqrt{2bL})/(1 + \sqrt{2bL}).$$

Then, $\{x_k\}$ generated by the inexact Newton method for solving $f(x) \in C$ with starting point $x_0$ and residual relative error tolerance $\theta$: $x_{k+1} := x_k + d_k,$

$$d_k \in \arg\min_{d \in X} \{\|d\| : f(x_k) + f'(x_k)d + r_k \in C\}, \quad \max_{w \in \{-r_k, r_k\}} \|T_{x_0}^{-1}w\| \leq \theta \|T_{x_0}^{-1}[-f(x_0)]\|,$$

for all $k = 0, 1, ..., \text{is well defined, for any particular choice of each } d_k, \|T_{x_0}^{-1}[-f(x_0)]\| \leq [(1 + \theta^2)/2]b,$ for all $k = 0, 1, ..., \{x_k\} \text{ is contained in } B(x_0, \lambda), \text{converges to a point } x_s \in B[x_0, \lambda],$ where $\lambda := \sqrt{2bL}/L.$ Moreover, \{x_k\} satisfies

$$\|x_k - x_{k+1}\| \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta}{2} \frac{L}{1 - \sqrt{2bL}} \|x_{k-1} - x_k\| + \theta \frac{1 + \sqrt{2bL}}{1 - \sqrt{2bL}} \|x_{k-1} - x_k\| \right], \quad k = 0, 1, \ldots$$

If, additionally, $0 \leq \theta < \left(-2(2 - \sqrt{2bL}) + \sqrt{10bL - 14\sqrt{2bL} + 21}\right)/(5 - \sqrt{2bL})$ then \{x_k\} converges $Q$-linearly as follows

$$\limsup_{k \to \infty} \frac{\|x_s - x_{k+1}\|}{\|x_s - x_k\|} \leq \frac{1 + \theta}{1 - \theta} \left[ \frac{1 + \theta}{2} + \frac{2\theta}{1 - \sqrt{2bL}} \right], \quad k = 0, 1, \ldots.$$

Proof. Take $\bar{\theta} = (1 - \sqrt{2bL})/(1 + \sqrt{2bL})$. Since $\psi : [0, 1/L] \to \mathbb{R},$ defined by $\psi(t) := (L/2)t^2 - t + b,$ is a majorant function for $f$ at point $x_0$, all results follow from Theorem 4.2.1, applied to this particular context. \hfill \blacksquare

Remark 4.4.2 In Theorem 4.4.1, if $\theta = 0$ and $C = \{0\}$ then we obtain, [19, Theorem 18] for the exact Newton method and [28, Theorem 6.3] for the inexact Newton method, respectively.

### 4.4.2 Under affine invariant Smale’s condition

In this section we present the convergence analysis of the inexact Newton method for cone inclusion problems under the affine invariant Smale’s condition.

Let $X$ and $Y$ be Banach spaces, $\Omega \subseteq X$ and $x_0 \in \Omega.$ A continuous function $f : \Omega \to Y$ and analytic in $\text{int} (\Omega)$ satisfies the affine invariant Smale’s condition with constant $\gamma$ at $x_0,$ if $B(x_0, 1/\gamma) \subset \Omega$ and

$$\gamma := \sup_{n > 1} \left\| \frac{T_{x_0}^{-1}f^{(n)}(x_0)}{n!} \right\|^{1/(n-1)} < +\infty.$$
Theorem 4.4.3 Let $C \subset \mathbb{Y}$ a nonempty closed convex cone. Suppose that $x_0 \in \Omega$ and $f$ satisfies Robinson’s and the affine invariant Smale’s condition with constant $\gamma$ at $x_0$ and there exists $b > 0$ such that

$$\|T_{x_0}^{-1}[-f(x_0)]\| \leq b, \quad b\gamma < 3 - 2\sqrt{2}, \quad 0 \leq \theta \leq \frac{1 - 2\sqrt{\gamma b} - \gamma b}{[1 + 2\sqrt{\gamma b} + \gamma b]}.$$  

Then, $\{x_k\}$ generated by the inexact Newton method for solving $f(x) \in C$ with starting point $x_0$ and residual relative error tolerance $\theta$: $x_{k+1} = x_k + d_k$,

$$d_k \in \text{argmin}_{d \in X} \{\|d\| : f(x_k) + f'(x_k)d + r_k \in C\}, \quad \max_{w \in \{-r_k, r_k\}} \|T_{x_0}^{-1}w\| \leq \theta \|T_{x_0}^{-1}[-f(x_k)]\|,$$

for all $k = 0, 1, \ldots$, is well defined, for any particular choice of each $d_k$, $\|T_{x_0}^{-1}[-f(x_k)]\| \leq [(1 + \theta^2)/2]k$, for all $k = 0, 1, \ldots$, $\{x_k\}$ is contained in $B(x_0, \lambda)$ and converges to a point $x_*$ in $B[x_0, \lambda]$ such that $f(x_*) \in C$, where $\lambda := b/\sqrt{\gamma b} + \gamma b$. Moreover, letting $\psi : [0, 1/\gamma) \to \mathbb{R}$ be defined by $\psi(t) = t/(1 - \gamma t) - 2t + b$, the sequence $\{x_k\}$ satisfies

$$\|x_k - x_{k+1}\| \leq \frac{1 + \theta}{1 - \theta} \left[\frac{1 + \theta D^-\psi'(\lambda)}{2|\psi'(\lambda)|}\|x_{k-1} - x_k\| + \theta \frac{2 + \psi'(\lambda)}{|\psi'(\lambda)|}\|x_{k-1} - x_k\|\right], \quad k = 0, 1, \ldots.$$  

If, additionally,

$$0 \leq \theta < \left(-2(2 - 2\sqrt{\gamma b} - \gamma b) + \sqrt{5\gamma^2b^2 - 44\gamma\sqrt{\gamma b} + 20\gamma b\sqrt{\gamma b} - 2\gamma b + 21}\right)/(5-2\sqrt{\gamma b} - \gamma b)$$

then $\{x_k\}$ converges $Q$-linearly as follows

$$\limsup_{k \to \infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|} \leq \frac{1 + \theta}{1 - \theta} \left[\frac{1 + \theta}{2} + \frac{2\theta}{1 - 2\sqrt{\gamma b} - \gamma b}\right], \quad k = 0, 1, \ldots.$$  

Proof. Take $\tilde{\theta} = (1 - 2\sqrt{\gamma b} - \gamma b)/(1 + 2\sqrt{\gamma b} + \gamma b)$. Use Lemma 20 of [19] to prove that $\psi : [0, 1/\gamma) \to \mathbb{R}$ defined by $\psi(t) = t/(1 - \gamma t) - 2t + b$, is a majorant function for $f$ in $x_0$, see [27]. Therefore, all results follow from Theorem 4.2.1, applied to this particular context.

Remark 4.4.4 In Theorem 3.2.1, if $\theta = 0$ and $C = \{0\}$ then we obtain, in the setting of Smale’s $\alpha$-theory, [19, Theorem 21] for the exact Newton method and [28, Theorem 6.1] for the inexact Newton method, respectively.
Chapter 5

Final remarks

In this thesis, we addressed the problem of finding a point satisfying the generalized equation (2) using the majorant condition. Our approach is based on the Banach Perturbation Lemma obtained by S. M. Robinson in [53, Theorem 2.4]. The majorant condition allows us to unify several convergence results pertaining to Newton’s method.

We obtained local as well as semi local convergence results for Newton’s method for solving (2). It is worth pointing out that Corollary 1.1.4 played a key role in the proof of the main results in Chapter 2 and Chapter 3. More specifically, it was used to ensure that the Newton iteration is well–defined; see Lemma 2.2.4 and Proposition 3.1.4.

We remark that the largest radius of convergence $r_x$ in Theorem 2.3.1 and Theorem 2.3.3 is inversely proportional to the Hölder/Lipschitz constant $K$ and Smale’s constant $\gamma$, respectively. Consequently, even though we may not know the solution $\bar{x}$ of the generalized equation, an estimation of these constants, which, in some cases, are independent of the solutions, provides information on the size of the convergence region.

In future work, we aim to study Newton-type methods by taking the approach used in this thesis under an assumption weaker than strong regularity, namely, the regularity metric or strong metric subregularity; see [16, 17]. We expect that with such an assumption, we can prove similar results to Theorem 2.1.1 and Theorem 3.1.1. However, we cannot ensure that the generated sequence in Theorem 3.1.1 is unique. Indeed, in this case, the mapping $L_f(x, \cdot)^{-1} \cap B(x_1, r_{x_1})$ is a set–valued mapping from $B(0, r_0)$ to $B(x_1, r_{x_1})$. It is well–known that the inexact analysis supports efficient computational implementations of the exact one. Therefore, following the idea of this thesis, we propose to study the inexact Newton method for solving problem (2), described by

$$(f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1})) \cap R_k(x_k, x_{k+1}) \neq \emptyset, \quad k = 0, 1, \ldots,$$

where $R_k : \mathbb{X} \times \mathbb{X} \rightrightarrows \mathbb{Y}$ is a sequence of set–valued mappings with closed graphs, in order to
support computational implementations of the method. Recently, this method has been the subject of a great deal of research; see [7, 16–18, 47]. Furthermore, it would be interesting to study both of the above mentioned methods under a majorant condition and metric regularity; see [16].
Bibliography

[1] 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.

[2] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton’s method in Riemannian manifolds. *Found. Comput. Math.*, 8(2):197–226, 2008.

[3] 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.

[4] 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.

[5] I. K. Argyros and S. Hilout. Inexact Newton-type methods. *J. Complexity*, 26(6):577–590, 2010.

[6] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.

[7] R. Cibulka, A. Dontchev, and M. H. Geoffroy. Inexact Newton Methods and Dennis–Moré Theorems for Nonsmooth Generalized Equations. *SIAM J. Control Optim.*, 53(2):1003–1019, 2015.

[8] R. Cibulka, A. Dontchev, J. Preininger, T. Roubal, and V. Veliov. Kantorovich-type theorems for generalized equations. *Research Report 2015-16, ORCOS, Vienna University of Technology*, pages 1–26, 2015.

[9] J. W. Daniel. Newton’s method for nonlinear inequalities. *Numer. Math.*, 21:381–387, 1973.

[10] J.-P. Dedieu, P. Priouret, and G. Malajovich. Newton’s method on Riemannian manifolds: covariant alpha theory. *IMA J. Numer. Anal.*, 23(3):395–419, 2003.
[11] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact Newton methods. *SIAM J. Numer. Anal.*, 19(2):400–408, 1982.

[12] S. P. Dokov and A. L. Dontchev. Robinson’s strong regularity implies robust local convergence of Newton’s method. In *Optimal control (Gainesville, FL, 1997)*, volume 15 of *Appl. Optim.*, pages 116–129. Kluwer Acad. Publ., Dordrecht, 1998.

[13] 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.

[14] A. L. Dontchev. Local convergence of the Newton method for generalized equations. *C. R. Acad. Sci. Paris Sér. I Math.*, 322(4):327–331, 1996.

[15] A. L. Dontchev and R. T. Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J. Optim.*, pages 1087–1105, 1996.

[16] A. L. Dontchev and R. T. Rockafellar. *Implicit functions and solution mappings. A view from variational analysis*. Springer Monographs in Mathematics. Springer, Dordrecht, 2009.

[17] A. L. Dontchev and R. T. Rockafellar. Newton’s method for generalized equations: a sequential implicit function theorem. *Math. Program.*, 123(1, Ser. B):139–159, 2010.

[18] 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.

[19] O. Ferreira. A robust semi–local convergence analysis of Newton’s method for cone inclusion problems in Banach spaces under affine invariant majorant condition. *Journal of Computational and Applied Mathematics*, 279(0):318 − 335, 2015.

[20] O. P. Ferreira. Local convergence of Newton’s method in Banach space from the viewpoint of the majorant principle. *IMA J. Numer. Anal.*, 29(3):746–759, 2009.

[21] O. P. Ferreira. Local convergence of Newton’s method under majorant condition. *J. Comput. Appl. Math.*, 235(5):1515–1522, 2011.

[22] O. P. Ferreira, M. L. N. Gonçalves, and P. R. Oliveira. Local convergence analysis of the Gauss–Newton method under a majorant condition. *J. Complexity*, 27(1):111–125, 2011.
[23] O. P. Ferreira, M. L. N. Gonçalves, and P. R. Oliveira. Convergence of the Gauss–Newton method for convex composite optimization under a majorant condition. *SIAM J. Optim.*, 23(3):1757–1783, 2013.

[24] O. P. Ferreira and G. N. Silva. Inexact Newton’s method to nonlinear functions with values in a cone. *Submitted*, 2016.

[25] O. P. Ferreira and G. N. Silva. Local convergence analysis of Newton’s method for solving strongly regular generalized equations. *Submitted*, 2016.

[26] O. P. Ferreira and G. N. Silva. Kantorovich’s theorem on Newton’s method for solving strongly regular generalized equation. *SIAM Journal on Optimization, Accepted for publication*, 2017.

[27] O. P. Ferreira and B. F. Svaiter. Kantorovich’s majorants principle for Newton’s method. *Comput. Optim. Appl.*, 42(2):213–229, 2009.

[28] O. P. Ferreira and B. F. Svaiter. A robust Kantorovich’s theorem on the inexact Newton method with relative residual error tolerance. *J. Complexity*, 28(3):346–363, 2012.

[29] M. L. N. Gonçalves. Local convergence of the Gauss–Newton method for injective–overdetermined systems of equations under a majorant condition. *Computers and Mathematics with Applications*, 66(4):490 – 499, 2013.

[30] M. L. N. Gonçalves. Inexact Gauss–Newton like methods for injective-overdetermined systems of equations under a majorant condition. *Numerical Algorithms*, 72(2):377–392, 2016.

[31] M. L. Gonçalves and J. G. Melo. A Newton conditional gradient method for constrained nonlinear systems. *Journal of Computational and Applied Mathematics*, 311:473 – 483, 2016.

[32] J. M. Gutierrez and M. A. Hernández. Newton’s method under weak Kantorovich conditions. *IMA Journal of Numerical Analysis*, 20(4):521–532, 2000.

[33] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms. I*, volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.

[34] Z. Huang. The convergence ball of Newton’s method and the uniqueness ball of equations under Hölder-type continuous derivatives. *Comput. Math. Appl.*, 47(2-3):247–251, 2004.
[35] A. F. Izmailov and M. V. Solodov. *Newton-type methods for optimization and variational problems*. Springer Series in Operations Research and Financial Engineering, Springer, Cham, 2014.

[36] N. Josephy. *Newton’s Method for Generalized Equations and the PIES Energy Model*. University of Wisconsin–Madison., 1979.

[37] L. V. Kantorovich. On Newton’s method for functional equations. *Doklady Akad. Nauk SSSR (N.S.)*, 59:1237–1240, 1948.

[38] L. V. Kantorovich and G. P. Akilov. *Functional analysis in normed spaces*. The Macmillan Co., New York, 1964.

[39] C. Kelley. *Solving Nonlinear Equations with Newton’s Method*. Society for Industrial and Applied Mathematics, 2003.

[40] C. Li and K. Ng. Extended Newton methods for conic inequalities: Approximate solutions and the extended Smale $\alpha$–theory. *J. Math. Anal. Appl.*, 440(2):636 – 660, 2016.

[41] C. Li and K. F. Ng. Convergence analysis of the Gauss-Newton method for convex inclusion and convex-composite optimization problems. *J. Math. Anal. Appl.*, 389(1):469–485, 2012.

[42] C. Li and K. F. Ng. Convergence analysis of the Gauss–Newton method for convex inclusion problems and convex composite optimization. *Preprint.*, pages 1–29, 2013.

[43] C. Li and W. Shen. Local convergence of inexact methods under the Hölder condition. *J. Comput. Appl. Math.*, 222(2):544–560, 2008.

[44] I. Moret. A Kantorovich-type theorem for inexact Newton methods. *Numer. Funct. Anal. Optim.*, 10(3-4):351–365, 1989.

[45] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.

[46] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*, volume 30 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.

[47] A. Pietrus and C. Jean-Alexis. Newton-secant method for functions with values in a cone. *Serdica Math. J.*, 39(3-4):271–286, 2013.
[48] B. T. Polyak. Newton’s method and its use in optimization. *European J. Oper. Res.*, 181(3):1086–1096, 2007.

[49] F. A. Potra. The Kantorovich theorem and interior point methods. *Mathematical Programming*, 102(1):47–70, 2005.

[50] L. B. Rall. A note on the convergence of Newton’s method. *SIAM Journal on Numerical Analysis*, 11(1):pp. 34–36, 1974.

[51] S. M. Robinson. Extension of Newton’s method to nonlinear functions with values in a cone. *Numer. Math.*, 19:341–347, 1972.

[52] S. M. Robinson. Normed convex processes. *Trans. Amer. Math. Soc.*, 174:127–140, 1972.

[53] S. M. Robinson. Strongly regular generalized equations. *Math. Oper. Res.*, 5(1):43–62, 1980.

[54] R. T. Rockafellar. *Monotone processes of convex and concave type*. Memoirs of the American Mathematical Society, No. 77. American Mathematical Society, Providence, R.I., 1967.

[55] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[56] W. Shen and C. Li. Kantorovich-type convergence criterion for inexact Newton methods. *Appl. Numer. Math.*, 59(7):1599–1611, 2009.

[57] G. N. Silva. Kantorovich’s theorem on Newton’s method for solving generalized equations under the majorant condition. *Applied Mathematics and Computation*, 286:178–188, 2016.

[58] G. N. Silva. Local convergence of Newton’s method for solving generalized equations with monotone operator. *Applicable Analysis*, pages 1–12, 2017.

[59] S. Smale. Newtons method estimates from data at one point. In R. Ewing, K. Gross, and C. Martin, editors, *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics*, pages 185–196. Springer New York, 1986.

[60] J. F. Traub and H. Woźniakowski. Convergence and complexity of Newton iteration for operator equations. *J. Assoc. Comput. Mach.*, 26(2):250–258, 1979.

[61] X. Wang. Convergence of Newton’s method and inverse function theorem in Banach space. *Math. Comp.*, 68(225):169–186, 1999.
[62] P. P. Zabrejko and D. F. Nguen. The majorant method in the theory of
Newton-Kantorovich approximations and the Pták error estimates. Numer. Funct. Anal. Optim., 9(5-6):671–684, 1987.

[63] Y. Zhang, J. Wang, and S.-M. Guu. Convergence criteria of the generalized Newton
method and uniqueness of solution for generalized equations. J. Nonlinear Convex Anal.,
16(7):1485–1499, 2015.