Nonlinear Functional Analysis and Applications
Vol. 17, No. 4 (2012), pp. 539-563

http://nfaa.kyungnam.ac.kr/jour-nfaa.htm

A FINER DISCRETIZATION AND MESH INDEPENDENCE OF NEWTON'S METHOD FOR SOLVING GENERALIZED EQUATIONS

I. K. Argyros[1] and S. K. Khattri[2]

[1]Department of Mathematical Sciences
Cameron University, Lawton, Oklahoma 73505-6377, USA
e-mail: iargyros@cameron.edu

[2]Department of Engineering
Stord Haugesund University College, Norway
e-mail: sanjay.khattri@hsh.no

Abstract. We further refine the mesh independence principle, proposed by Alt et al. (see [14]), for solving generalized equations in a Banach space setting by using Newton's method.

1. Introduction

Generalized equations can be used to solve optimization, optimal control and other problems [1–37]. Solution of such equations usually involves an iterative method. The most popular method for solving generalized equations is undoubtedly the generalized Newton's method (**GNM**) [3, 14, 15, 33, 34, 35]. Robinson provided local and semilocal results for generalized equations in [35]. Josephy [24] used the results in [32] to provide a convergence analysis of Newton type methods for variational inequalities in finite dimensional spaces. Applications of these results to optimization and nonlinear control problems were given by Alt et al. [10]-[14], Dontcher et al. [16]-[18] and Malanowski [28, 29]. Argyros [1]-[8] introduced the center-Lipschitz condition and more precise majorizing sequences to show convergence of Newton's method with the following advantages over the works in [9]: **Semilocal case -** a) Weaker sufficient convergence conditions. b) Tighter error bounds on the distances

[0]Received November 23, 2012. Revised December 7, 2012.
[0]2000 Mathematics Subject Classification: 65G99, 65J15, 47H17, 49M15.
[0]Keywords: Newton's method, generalized equation, discretization, mesh independence principle, Banach space, Fréchet-derivative.

involved. c) At least as precise information on location of the solution. **Local case -** d) Large convergence ball. e) Tighter error bounds. These advantages were obtained under the same computational cost. Only in special cases, generalized equations in infinite dimensional spaces can be solved analytically. Discretized generalized equations are usually used to approximate solutions of original equations using (**GNM**). For operator equations Allgower et al. [9] have provided relations between the original and the discretized equations. Moreover, they showed that if certain consistency and stability conditions are satisfied then the local behavior of the discretized (**GNM**) iterations is asymptotically the same as that for the original equation. The mesh independence principle was extended in [14] to generalized equations. Further, Argyros et al. improved the results in [8, 9].

In this study, we improve the results in [14]. We show as in the case of nonlinear equations and under the same computational cost that the advantages (a)-(e) hold. Moreover, the number of iterations required to have the same error tolerance between (**GNM**) and discretized (**GNM**) iterations can be reduced.

The rest of the paper is organized as follows. The local convergence of (**GNM**) is given in the Section 2. The mesh independence principle, for the generalized equations, is shown in Section 3.

2. Newton's method for generalized equations

Let $\mathbf{Y}$ be a normed space, $\mathcal{Z}$ be a Banach space, $\mathcal{D}$ be an open subset of $\mathcal{Z}$. Furthermore, let $\mathcal{F} : \mathcal{D} \to \mathbf{Y}$ be a mapping and $\mathbf{T} : \mathcal{D} \to \mathbf{Y}$ be a multi-valued mapping. Then we consider the problem

$$\text{Find} \quad z \in \mathcal{D} \quad \text{such that} \quad \mathcal{F}(x) \in \mathbf{T}(z). \tag{2.1}$$

Given that the mapping $\mathcal{F}$ is Fréchet differentiable on $\mathcal{D}$ then Newton's method for operator equations can be extended to the generalized equation (2.1) as follows.

(GNM):. Select an initial point $z_0 \in \mathcal{D}$, z_{k+1} is computed as the solution of the generalized equation

$$\mathcal{F}(z_k) + \mathcal{F}'(z_k)(z - z_k) \in \mathbf{T}(z). \tag{2.2}$$

For (**GNM**), we prove local convergence and semi-local convergence theorems. These theorems generalize the well-known Newton-Kantorovich theorem [8]. In the next section, these convergence results are employed to analyze discretization of (**GNM**). Following the ideas presented in [14, 33, 35], the linearized equations (2.2) are considered as perturbations of the equation (2.1). Let $w \in \mathcal{Z}$ then a family of perturbed generalized equations, depending on the parameter w, is given as

$$\mathcal{F}(w) + \mathcal{F}'(w)(z - w) \in \mathbf{T}(z). \tag{2.3}$$

During the k^{th} iteration of (**GNM**), the above equation is solved for $w = z_k$. To perform convergence analysis of (**GNM**), a somewhat generalized version of an implicit-function theorem due to Robinson [28] is used. The implicit-function theorem, due to Robinson [28], requires the concept of strong regularity and which generalizes the assumption of invertibility of $\mathcal{F}'(z^\star)$ in cases of operator equations.

Definition 2.1. Let the mapping $\mathcal{F}$ be Fréchet differentiable on $\mathcal{D}$. We say that (2.1) is strongly regular at $z^\star \in \mathcal{D}$ if there exist $r_Y(z^\star) > 0$ and $l_L(z^\star) > 0$ such that for all $y \in U_Y(0_Y, r_Y(z^\star))$ the linearized system

$$\mathcal{F}(z^\star) + \mathcal{F}'(z^\star)(z - z^\star) + y \in \mathbf{T}(z), \tag{2.4}$$

has a unique solution $S_L(z^\star, y)$ and the mapping $S_L(z^\star, \cdot) : U_Y(0_Y, r_Y(z^\star)) \to \mathcal{Z}$ is Lipschitz continuous with modulus $l_L(z^\star)$.

Particularly, strong regularity of (2.1) at $z^\star$ requires that $z^\star$ is the unique solution of (2.4) for $y = 0_Y$. Extensions of this concept and applications to optimal control can be found in Hager et al. [15] and Dontchev et al. [9, 10]. We need the following assumptions for Lipschitz continuity of $\mathcal{F}'$.

A. Let $z^\star \in \mathcal{D}$. There exist parameters $r_1(z^\star) > 0$ and $l_{\mathcal{F}}(z^\star) > 0$ such that $U_z(z^\star, r_1(z^\star)) \subseteq \mathcal{D}$. If $\mathcal{F} : \mathcal{D} \subseteq \mathcal{Z} \to \mathbf{Y}$ is Fréchet differentiable then

$$\|\mathcal{F}'(u) - \mathcal{F}'(v)\|_{\mathcal{Z}\to\mathbf{Y}} \le l_F(z^\star)\|u - v\|_{\mathcal{Z}} \tag{2.5}$$

for all $u, v \in U_z(z^\star, r_1(z^\star))$. The multi-valued mapping $\mathbf{T}$ has closed graph.

A′. It follows from hypotheses (**A**) that there exists $\alpha_{\mathcal{F}(z^\star)} \in (0, 1]$ such that

$$\|\mathcal{F}'(u) - \mathcal{F}'(z^\star)\|_{\mathcal{Z}\to\mathbf{Y}} \le \alpha_{\mathcal{F}(z^\star)} l_F(z^\star)\|u - z^\star\|_{\mathcal{Z}} \tag{2.6}$$

for all $u \in U_z(z^\star, r_1(z^\star))$. Note that (**A′**) is not an additional hypothesis to (**A**).

For some $z^\star \in \mathcal{D}$, we define mapping $l(z^\star; \cdot) : \mathcal{D} \times \mathcal{D} \to \mathbf{Y}$ by

$$l(z^\star; z, w) = \mathcal{F}(w) + \mathcal{F}'(w)(z - w) - \mathcal{F}(z^\star) - \mathcal{F}'(z^\star) \tag{2.7}$$

(see [25]-[28]). We need the following results on the local convergence of (**GNM**).

Lemma 2.2. *Under the hypotheses* (**A**) the following assertions hold

1. For all $z, w \in U_z(z^\star, r_1(z^\star))$

$$\|l(z^\star; z, w)\| \leq l_{\mathcal{F}(z^\star)} \left[\frac{\|w - z^\star\|_{\mathcal{Z}}}{z} + \alpha_{\mathcal{F}(z^\star)} \|z - z^\star\|_{\mathcal{Z}} \right] \|w - z^\star\|. \tag{2.8}$$

2. For fixed $r_Y > 0$ and

$$\phi_{\mathcal{F}}(z^\star) = \frac{1 + 2\alpha_{\mathcal{F}}(z^\star)}{3} \tag{2.9}$$

define

$$r_2(z^\star) = \min \left\{ r_1(z^\star), \sqrt{\frac{2r_Y}{3\phi_F(z^\star) l_{\mathcal{F}}(z^\star)}} \right\}. \tag{2.10}$$

Then, for all $z, w \in U_z(z^\star, r_1(z^\star))$

$$\|l(z^\star; z, w)\|_Y \leq r_Y. \tag{2.11}$$

3. For fixed $l^\star > 0$, define

$$r_3(z^\star) = \min \left\{ r_1(z^\star), \frac{2}{3\alpha_F(z^\star) l_{\mathcal{F}}(z^\star) l^\star} \right\}. \tag{2.12}$$

Then for all $u, v \in U_z(z^\star, r_1(z^\star))$ and for all $w \in U_z(z^\star, r_3(z^\star))$

$$\|l(z^\star; u, w) - l(z^\star; v, w)\|_Y \leq \frac{2}{3l^\star} \|u - v\|_Z. \tag{2.13}$$

Proof.

1. Mapping l can be written as

$$\begin{aligned} l(z^\star; z, w) &= \left[\mathcal{F}(w) - \mathcal{F}(z^\star) - \mathcal{F}'(w)(w - z^\star)\right] \\ &\quad + (\mathcal{F}'(w) - \mathcal{F}'(z^\star))(z - z^\star) \\ &= \int_0^1 \left[\mathcal{F}'(z^\star + \theta(w - z^\star)) - \mathcal{F}'(w)\right] (w - z^\star)\, d\theta \\ &\quad + (\mathcal{F}'(w) - \mathcal{F}'(z^\star))(z - z^\star) \end{aligned} \tag{2.14}$$

Using (2.5) in the first term and (2.6) in the second term we get (2.8).

2. It follows from (2.8) that its right hand side is bounded from above by

$$\frac{l_F(z^\star)}{2} r_1^2(z^\star) + \phi_F(z^\star) l_F(z^\star) r_1^2(z^\star) \leq r_Y \tag{2.15}$$

which is true by the choice of $r_2(z^\star)$ and ϕ.

3. Using (2.7), we get

$$L(z^\star; u, w) - L(z^\star; v, w) = (\mathcal{F}'(w) - \mathcal{F}'(z^\star))(u - v) \tag{2.16}$$

The result now follows from (2.6) and (2.16). That completes the proof of the Lemma.

□

Assume z̃ is a solution of the generalized equation (2.1). Next we provide sufficient conditions for the quadratic convergence of (**GNM**) to $\tilde{z}$.

Lemma 2.3. *Suppose*

$\mathbf{A}_1$. *There exists a solution* $\tilde{z}$ *of generalized equation* (2.1) *and*
$\mathbf{A}_2$. *Hypotheses* $(\mathbf{A}_1)$ is satisfied for $z^\star = \tilde{z}$.

Then for all $w \in U_z(\tilde{z}, r_1(\tilde{z}))$ *the following assertions hold*

$$\|l(\tilde{z}; \tilde{z}, w)\|_{\mathbf{Y}} \leq \frac{l_{\mathcal{F}}(\tilde{z})}{2} \|w - \tilde{z}\|^2_{\mathcal{Z}} \tag{2.17}$$

and

$$\|l(\tilde{z}; \tilde{z}, w)\|_{\mathbf{Y}} \leq \frac{3\alpha_{\mathcal{F}}(z^\star) l_{\mathcal{F}}(\tilde{z})}{2} \|w - \tilde{z}\|^2_{\mathcal{F}}. \tag{2.18}$$

Proof. Set $z^\star = \tilde{z} = z$ in (2.8) to obtain (2.17). From (2.7), we also have

$$\begin{aligned} l(z; \tilde{z}, w) &= \mathcal{F}(w) - \mathcal{F}(\tilde{z}) - \mathcal{F}'(w)(w - \tilde{z}) \\ &= \int_0^1 \left[\mathcal{F}'(\tilde{z} + \theta(w - \tilde{z})) - \mathcal{F}'(\tilde{z})\right](w - \tilde{z})\, d\theta \\ &\quad + (\mathcal{F}'(\tilde{z}) - \mathcal{F}'(w))(w - \tilde{z}). \end{aligned} \tag{2.19}$$

Estimate (2.18) follows from (2.6) for $\tilde{z} = z^\star$ and (2.19). That completes the proof of the Lemma. □

Remark 2.4. If $(\mathbf{A}_2)$ is satisfied and problem (2.1) is strongly regular at $\tilde{z}$ then in view of Lemma 2.2 (i) mapping $S_L(\tilde{z}, l(\tilde{z}; z, w))$ is well defined for all $z, w \in U_z(\tilde{z}, r_2(\tilde{z}))$. Therefore for each $w \in U_z(\tilde{z}, r_2(\tilde{z}))$ we define a mapping

$$\begin{aligned} &S_w : U_z(\tilde{z}, r_2(\tilde{z}) \longrightarrow \mathcal{Z}, \\ &\mathcal{Z} \to S_L(\tilde{z}, L(\tilde{z}; z, w)). \end{aligned} \tag{2.20}$$

It follows from $L(\tilde{z}; z, w) = O_{\mathbf{X}}$ that $S_{\tilde{z}}(\tilde{z}) = S_L(\tilde{z}, O_{\mathbf{Y}})$. That is if (2.1) is strongly regular at $\tilde{z}$ then $\tilde{z} = S_L(\tilde{z}, O_Y)$ which means that $\tilde{z}$ is a fixed point of $S_{\tilde{z}}$. Moreover, z_w is a fixed point of S_w iff z_w is a solution of (2.3). It also follows that z_{k+1} is a fixed point of S_{z_k}.

We show the following local result for quadratic convergence of (**GNM**).

Theorem 2.5. *Suppose that hypotheses* $(\mathbf{A}_1)$, $(\mathbf{A}_2)$ are satisfied and generalized equation (2.1) is strongly regular at $\tilde{z}$. Define

$$\rho = \rho(\tilde{z}) = \begin{cases} m_1 & \text{if} \quad \alpha_{\mathcal{F}}(\tilde{z}) < \dfrac{1}{3}, \\ m_2 & \text{if} \quad \alpha_{\mathcal{F}}(\tilde{z}) \geq \dfrac{1}{3} \end{cases} \tag{2.21}$$

where,

$$m_1 = \min\left\{ r_1(\tilde{z}), \sqrt{\frac{2r_y(\tilde{z})}{3\phi_F(\tilde{z})l_{\mathcal{F}}(\tilde{z})}}, \frac{2}{3l_{\mathcal{F}}(\tilde{z})l_L(\tilde{z})} \right\}, \tag{2.22}$$

$$m_2 = \min\left\{ r_1(\tilde{z}), \sqrt{\frac{2r_y(\tilde{z})}{3\phi_F(\tilde{z})l_F(\tilde{z})}}, \frac{2}{9\alpha_{\mathcal{F}}(\tilde{z})l_F(\tilde{z})l_L(\tilde{z})} \right\}. \tag{2.23}$$

Then, there exists a single valued mapping $S : U_z(\tilde{z},\rho) \to U_z(\tilde{z},\rho)$ such that for each $w \in U_z(\tilde{z},\rho), \rho(w)$ is the unique solution in $U_z(\tilde{z},\rho)$ of (2.3). Moreover, the following assertions hold

$$\|S(w) - \tilde{z}\|_{\mathcal{Z}} \le 3l_L(\tilde{z})\|l(\tilde{z};\tilde{z},w)\|_{\mathbf{Y}} \le \frac{3}{2}l_L(\tilde{z})l_F(\tilde{z})\|w-\tilde{z}\|^2_{\mathcal{Z}} \tag{2.24}$$

and

$$\|S(w) - \tilde{z}\| \le 3l_L(\tilde{z})\|l(\tilde{z};\tilde{z},w)\|_{\mathbf{Y}} \le \frac{9\alpha}{2}l_L(\tilde{z})l_{\mathcal{F}}(\tilde{z})\|w-\tilde{z}\|^2_{\mathcal{Z}} \tag{2.25}$$

In the special case when $\alpha_{\mathcal{F}}(\tilde{z}) = 1$ (i.e. $\phi_{\mathcal{F}}(\tilde{z}) = 1$), the proof of (2.24) was given in [10] [Theorem 2.4] and in the Appendix of [14]. The proof of (2.25) follows from (2.24) and (2.18).

Next we state another local result for (**GNM**).

Theorem 2.6. *Suppose that the hypotheses* ($\mathbf{A}_2$), ($\mathbf{A}_3$) are satisfied and generalized equation (2.1) is strongly regular at $\tilde{z}$. Let $\rho(\tilde{z})$ be given by (2.21). Define

$$\tilde{\rho}(\tilde{z}) = \begin{cases} \frac{3}{8}l_{\mathcal{F}}(\tilde{z})l_L(\tilde{z})\rho^2(\tilde{z}), \\ \frac{9\alpha_{\mathcal{F}}(\tilde{z})}{8}l_{\mathcal{F}}(\tilde{z})l_L(\tilde{z})\rho^2(\tilde{z}) \quad \text{if} \quad \alpha < \frac{1}{3}, \end{cases} \tag{2.26}$$

$$\delta = \begin{cases} \frac{3}{4}l_F(\tilde{z})l_L(\tilde{z})\rho(\tilde{z}), \\ \frac{9\alpha_{\mathcal{F}}(\tilde{z})}{4}l_{\mathcal{F}}(\tilde{z})l_L(\tilde{z})\rho(\tilde{z}) \quad \text{if} \quad \alpha < \frac{1}{3}. \end{cases} \tag{2.27}$$

Then, sequence $\{z_n\}$ generated by (**GNM**) for $z_0 \in U_z(\tilde{z},\tilde{\rho}(\tilde{z}))$ converges to $\tilde{z}$. Moreover, the following assertions hold

$$\|z_{k+1} - \tilde{z}\|_{\mathcal{Z}} \le \frac{3}{2}l_{\mathcal{F}}(\tilde{z})l_L(\tilde{z})\|z_k-\tilde{z}\|^2_{\mathcal{Z}} \le \frac{1}{2}\tilde{\rho}(\tilde{z})\delta^{2^{k+1}-1}, \tag{2.28}$$

$$\|z_{k+1} - \tilde{z}\|_{\mathcal{Z}} \le \frac{9\alpha_{\mathcal{F}}(\tilde{z})}{2}l_{\mathcal{F}}(\tilde{z})l_L(\tilde{z})\|z_k-\tilde{z}\|^2_{\mathcal{Z}} \le \frac{1}{2}\tilde{\rho}(\tilde{z})\delta^{2^{k+1}-1}, \tag{2.29}$$

$$\|z_{k+1} - \tilde{z}\|_{\mathcal{Z}} \le \frac{3}{2}l_{\mathcal{F}}(\tilde{z})l_L(\tilde{z})\|z_k-\tilde{z}\|^2_{\mathcal{Z}} \le \frac{1}{2}\|z_k-\tilde{z}\|_{\mathcal{Z}} \tag{2.30}$$

and

$$\|z_{k+1} - \tilde{z}\|_{\mathcal{Z}} \leq \frac{9\alpha_{\mathcal{F}}(\tilde{z})}{2} l_{\mathcal{F}}(\tilde{z}) l_L(\tilde{z}) \|z_k - \tilde{z}\|_{\mathcal{Z}}^2 \leq \frac{1}{2}\|z_k - \tilde{z}\|_{\mathcal{Z}}. \tag{2.31}$$

The (2.28) is proved in [14][Theorem 3.3 and Theorem 2.6] for $\phi_F(\tilde{z}) = 1$ by using (2.24). The proof of (2.29) is given in an analogous way but by using (2.25). Finally estimates (2.30) and (2.31) are obtained from the definitions of $\tilde{\rho}(\tilde{z})$, $y(\tilde{z})$ and (2.29), (2.30), respectively.

Remark 2.7. If $\alpha_{\mathcal{F}}(\tilde{z}) = 1$, all the preceding results reduce to the corresponding ones in [14]. Otherwise (i.e. if $\alpha \in (0, 1)$), they constitute an improvement since the convergence balls are at least as large and error estimates are tighter. Note also that if $\alpha \in (0, 1/3)$, estimates (2.25) and (2.29) are tighter than (2.24) and (2.28), respectively.

In many problems, the solution $\tilde{z}$ of the generalized equation (2.1) and the corresponding iterates $\{z_k\}$ have better smoothness properties than the elements of $\mathcal{Z}$. It encourages us to consider a subset $\mathcal{Z}_R \subset \mathcal{Z}$ such that

$$\tilde{z} \in Z_R, \quad z_k \in Z_R, \quad z_k - \tilde{z} \in Z_R, \quad z_{k+1} - z_k \in Z_R, \quad k = 0, 1, \ldots. \tag{2.32}$$

To show that $\{z_k\}_{k\in\mathbb{N}} \subset Z_R$ if $z_0, \tilde{z} \in Z_R$, we consider the assumptions

- **A3**$_1$. Let there be closed and convex subsets $\tilde{\mathcal{Z}}_R \subset \mathcal{Z}_R$ with $\tilde{z} \in \tilde{\mathcal{Z}}_R$, $\mathbf{Y}_R \subset \mathbf{Y}$ with $O_{\mathbf{Y}} \in \mathbf{Y}_R$ and constants $r_{\mathbf{Y}}(\tilde{z}) > 0$ and $l_L(\tilde{z}) > 0$, such that for all $y \in \mathbf{Y}_R \cap U_{\mathbf{Y}}(0_{\mathbf{Y}}, r_{\mathbf{Y}}(\tilde{z}))$ the linearized system
$$\mathcal{F}(\tilde{z}) + \mathcal{F}'(z - \tilde{z}) + y \in \mathbf{T}(z), \tag{2.33}$$
has a unique solution $S_L(\tilde{z}, y) \in \tilde{\mathcal{Z}}_R$ and the mapping $S_L(\tilde{z}, \cdot) : \mathbf{Y}_R \cap U_{\mathbf{Y}}(O_{\mathbf{Y}}, r_{\mathbf{Y}}) \to \tilde{\mathcal{Z}}_R$ is Lipschitz continuous with modulus $l_L(\tilde{z})$.
- **A3**$_2$. There exists $r_R > 0$ such that $l(\tilde{z}, z, w) \in \mathbf{Y}_R \cap U_{\mathbf{Y}}(0_{\mathbf{Y}}, r_{\mathbf{Y}}(\tilde{z}))$ for all $z, w \in \tilde{\mathcal{Z}}_R \cap B_{\mathcal{Z}}(\tilde{z}, r_R)$.

The preceding assumptions reflect strong regularity. The assumptions (**A3**) have long been familiar in stability and sensitivity analysis of optimal control problems (cf. [20, 21]). From Theorem 2.6, we obtain

Corollary 2.8. *Let us presume that Assumptions* (**A1**)-(**A3**) *are satisfied and let* $\tilde{\rho}(\tilde{z})$ *be defined by* (2.26) *and* (2.27). *Then for any starting point* $z_0 \in \tilde{\mathcal{Z}}_R \cap U_{\mathcal{Z}}(\tilde{z}, \tilde{r})$ *with* $\tilde{r} = \min\{\tilde{\rho}(\tilde{z}, r_R)\}$ *the* (**GNM**) *generates a unique sequence* $\{z_k\} \subset \tilde{Z}_R$ *convergent to* $\tilde{z} \in \tilde{Z}_R$ *satisfying* (2.29) *for* $k \geq 1$.

Proof. It can be easily seen from the proof of Theorem 2.5 that (**A3**) implies $S(z) \in \tilde{\mathcal{Z}}_R$ for $z \in \tilde{\mathcal{Z}}_R \cap U_{\mathcal{Z}}(\tilde{z}, \tilde{r})$. Thus, if $\{z_k\}$ is the sequence defined by (**GNM**) and $z_0 \in \tilde{\mathcal{Z}}_R \cap U_{\mathcal{Z}}(\tilde{z}, \tilde{r})$ then $\{z_k\} \subset \tilde{\mathcal{Z}}_R$. $\square$

Remark 2.9. Let the hypothesis of Corollary 2.8 be satisfied and $\mathcal{Z}_R \subset \mathcal{Z}$ be any set with the property $\mathcal{Z}_R \supset \tilde{\mathcal{Z}}_R - \tilde{\mathcal{Z}}_R$. Then

$$z_k - \tilde{z}_k \in \mathcal{Z}_R, \quad z_{k+1} - z_k \in \mathcal{Z}_R, \quad k = 0, 1, \ldots,$$

that is (2.32) is satisfied. Later, in Section 4, we will use this fact and apply the abstract theory to optimal control problems.

Next we derive a result on semi-local convergence of (**GNM**) which generalizes the well-known Newton-Kantorvich theorem [8]. Similar results have been obtained [19][Theorem 7.1], [26][Theorem 5.1] (see also [17]). For the starting point z_0 of (**GNM**), we assume that the following assumptions hold :

B$_1$**.** $z_0 \in \mathcal{D}$.

B$_2$**.** Assumption (**A**) is satisfied with $z^\star = z_0$.

Suppose that the generalized equation (2.1) is strongly regular at z_0. Then for all $y \in U_{\mathbf{Y}}(0_{\mathbf{Y}}, r_{\mathbf{Y}}(z_0))$ the linear generalized equation

$$\mathcal{F}(z_0) + \mathcal{F}'(z_0)(z - z_0) + y \in \mathbf{T}(z), \tag{2.34}$$

has a unique solution $S_L(z_0, y) \in \mathcal{Z}$. For $y = 0_{\mathbf{Y}}$, let us denote the unique solution of (2.34) by $z_1 = S_L(z_0, 0_Y)$. Since z_0 is the initial point for (**GNM**) and z_1 is the first iterate. It follows from the definition (2.5) of l that $l(z_0; z_1, z_0) = 0_{\mathbf{Y}}$ and hence

$$z_1 = S_L(z_0, 0_{\mathbf{Y}}) = S_L(z_0, l(z_0; z_1, z_0)). \tag{2.35}$$

By Lemma 2.2(b), $S_L(z_0, l(z_0; z, w)$ is well-defined for all $z, w \in U_{\mathcal{Z}}(z_0, r_2(z_0))$ where $r_2(z_0)$ is defined by (2.10). Hence for $w \in U_{\mathcal{Z}}(z_0, r_2(z_0))$ we can define a mapping

$$S_w : U_{\mathcal{Z}}(z_0, r_2(z_0)) \longrightarrow \mathcal{Z}, \quad z \mapsto S_L(z_0, l(z_0; z, w).$$

For $w = z_0$ we obtain from (2.14)

$$S_{z_0}(z_1) = S_L(z_0, l(z_0; z_1, z_0) = S_L(z_0, O_{\mathbf{Y}}) = z_1,$$

that is z_1 is a fixed point of the mapping S_{z_0}. Analogous to Theorem 2.5 one can show that S_w has a unique fixed point provided that $\|z_1 - z_0\|_{\mathcal{Z}}$ is sufficiently small (see [14][Theorem 2.4]).

Theorem 2.10. *Let the assumptions* (**B**$_1$)-(**B**$_2$) hold and let

$$\rho = \rho(z_0) = \min\left\{\frac{1}{2} r_2(z_0), \frac{2}{9 l_F(z_0) l_L(z_0)}\right\}. \tag{2.36}$$

Furthermore assume that $z_1 \in U_{\mathcal{Z}}(z_0, \rho)$. Then there exists a single valued function $S : U_{\mathcal{Z}}(z_0, \rho) \to U_{\mathcal{Z}}(z_1, \rho)$ such that for each $w \in U_{\mathcal{Z}}(z_0, \rho)$, $S(w)$ is the unique fixed point in $U_{\mathcal{Z}}(z_1, \rho)$ of S_w and

$$\|S(w) - S(v)\|_{\mathcal{Z}} \leq 3 l_L(z_0) \|l(z_0; S(v), w) - l(z_0; S(v), v)\|_{\mathbf{Y}} \tag{2.37}$$

for all $v, w \in U_{\mathcal{Z}}(z_0, \rho)$.

Remark 2.11. Let $\{z_k\}$ be the sequence defined by (**GNM**). It follows from the definition of l that z_{k+1} is the solution of

$$\mathcal{F}(z_0) + \mathcal{F}'(z_0)(z - z_0) + l(z_0; z_{k+1}, z_k) \in \mathbf{T}(z).$$

If $l(z_0; z_{k+1}, z_k) \in U_{\mathbf{Y}}(0_{\mathbf{Y}}, r_{\mathbf{Y}}(z_0))$ this implies that

$$z_{k+1} = S_L(z_0, l(z_0; z_{k+1}, z_k)) \Leftrightarrow z_{k+1} = S_{z_k}(z_{k+1}).$$

that is z_{k+1} is a fixed point of S_{z_k} and hence $z_{k+1} = S(z_k)$. This fact will be used in the proof of Theorem 2.14.

We need some results to find upper bounds on the right hand side of estimate (2.37).

Lemma 2.12. *Suppose assumptions* ($\mathbf{B}_1$)-($\mathbf{B}_2$) are satisfied. Then the following assertions hold

$$\|l(z_0; w, w) - l(z_0; w, v)\|_{\mathbf{Y}} \leq \frac{l_F(z_0)}{2}\|w - v\|^2_{\mathcal{Z}}$$

for all $w, v \in U_{\mathcal{Z}}(z_0, r_1(z_0))$ and

$$\|l(z_0; w, w) - l(z_0; w, z_0)\|_{\mathbf{Y}} \leq \frac{\alpha_{\mathcal{F}}(z_0) l_{\mathcal{Z}}(z_0)}{2}\|w - z_0\|^2_{\mathcal{Z}}$$

for all $w \in U_{\mathcal{Z}}(z_0, r_1(z_0))$.

Proof. By the definition of l we have

$$l(z_0; w, w) - l(z_0; w, v) = \mathcal{F}(w) - \mathcal{F}(v) - \mathcal{F}'(v)(w - v).$$

The first assertion holds from assumption ($\mathbf{B}_2$) (i.e. (**A**)) where as the second assertion holds from ($\mathbf{B}_2$) but using ($\mathbf{A}'$) and $v = z_0$. That completes the proof of the Lemma. □

Remark 2.13. If $\tilde{z} \in U_z(z_0, \rho(z_0))$ then by Lemma 2.2(b)

$$l(z_0; \tilde{z}, \tilde{z}) = \mathcal{F}(\tilde{z}) - \mathcal{F}(z_0) - \mathcal{F}'(z_0)(\tilde{z} - z_0) \in U_{\mathbf{Y}}(0_{\mathbf{Y}}, r_{\mathbf{Y}}(z_0)).$$

Hence, $\tilde{z}$ is a fixed point of S iff

$$\begin{aligned}
&\tilde{z} = S_{\tilde{z}}(\tilde{z}) \Leftrightarrow \tilde{z} = S_L(z_0, l(z_0; \tilde{z}, \tilde{z})),\\
&\Leftrightarrow \mathcal{F}(z_0) + \mathcal{F}'(z_0)(z - z_0) + l(z_0; \tilde{z}, \tilde{z}) \in \mathbf{T}(z),\\
&\Leftrightarrow \mathcal{F}(\tilde{z}) \in \mathbf{T}(\tilde{z}).
\end{aligned}$$

That is $\tilde{z}$ is a solution of the generalized equation (2.1). We need the following result on majorizing sequences for (**GNM**). The proof is obtained from the [9][Section 1.1] by simply setting $L_0 = \alpha L$, $\alpha \in (0, 1]$.

Lemma 2.14. *Let $l > 0$, $\eta > 0$ and $\alpha \in (0,1]$ be given parameters. Define function $\gamma^\alpha : (0,1] \to (0,1]$ by*

$$\gamma^\alpha(t) = \frac{1}{8}\left(4t + \sqrt{t} + \sqrt{t + 8t^2}\right).$$

Let us denote

$$\gamma = \gamma^\alpha(\alpha).$$

Suppose

$$h_\alpha = \gamma l \eta \le \frac{1}{2}. \tag{2.38}$$

Define function $g^\alpha : (0,1] \to (0,+\infty)$ by

$$g^\alpha(t) = \begin{cases} 1 + \dfrac{(1+\sqrt{1+8\alpha})^2}{16\alpha}\dfrac{t}{1-t}, & \alpha \ne 1 \\ \dfrac{2}{1+\sqrt{1-2t}}, & \alpha = 1. \end{cases}$$

Then, scalar sequence $\{t_n^\alpha\}$ $(n \ge 0)$ given by

$$t_0^\alpha = 0,\ \ t_1^\alpha = \eta,\ \ t_2^\alpha = \eta + \frac{\alpha l \eta^2}{2(1-\alpha l \eta)},\ \ t_{n+2}^\alpha = t_{n+1}^\alpha + \frac{l(t_{n+1}^\alpha - t_n^\alpha)^2}{2(1-\alpha l t_{n+1}^\alpha)} \tag{2.39}$$

is well defined, increasing, bounded from above by

$$t_\star^\alpha = g^\alpha(\alpha l \eta)\eta \tag{2.40}$$

and converges to its unique least upper bound t_∞^α satisfying

$$t_\infty^\alpha \le t_\star^\alpha. \tag{2.41}$$

Moreover, the following assertions hold

$$h_\alpha \le h_1, \tag{2.42}$$

$$t_n^\alpha \le t_n^1 \tag{2.43}$$

$$t_{n+1}^\alpha - t_n^\alpha \le t_{n+1}^1 - t_n^1 \tag{2.44}$$

$$t_\infty^\alpha \le t_\infty^1 = g^1(l\eta)\eta = \frac{2\eta}{1+\sqrt{1-2l\eta}} \le 2\eta, \tag{2.45}$$

$$t_\star^\alpha \le \lambda^\alpha \eta \le \frac{1}{\alpha l} \tag{2.46}$$

and

$$0 < r_\alpha^\star = r^\star := t_\infty^\alpha - \eta \le t_\star^\alpha - \eta = r_1^\star = r_1^\star(\alpha) \tag{2.47}$$

where,

$$\lambda = \lambda^\alpha = \begin{cases} g^\alpha(\alpha l \eta_0) = g^\alpha(\frac{\alpha}{2\gamma}) \\ = \frac{\sqrt{\alpha}}{\alpha}(\sqrt{\alpha} + \sqrt{1+8\alpha} - 1) \in (l, 3) & \text{for} \quad \alpha \neq 1 \\ 2, & \text{if} \quad \alpha = 1 \end{cases} \tag{2.48}$$

and

$$\eta_0 = \frac{1}{2\gamma l}.$$

Furthermore, estimates (2.42)-(2.44) hold as strict inequalities if $\alpha \in (0,1)$ and $n \geq 2$ for (2.43), $n \geq 1$ for (2.44). It follows from (2.48) that $\lambda^\alpha < \lambda^1 = 2$ for $\alpha \in (0, 0.81]$. In particular, if $\alpha = 0.81$, $\lambda^\alpha = 1.99697662$. That is $t_\star^\alpha < t_\star^1$ for $\alpha \in (0, 0.8)$.

Remark 2.15. Iteration $\{t_n^1\}$ was used in [14] as a majorizing sequence for $\{z_n\}$. However, if $\alpha \in (0,1)$ then $\{t_n^\alpha\}$ is a tighter majorizing sequence for $\{t_n\}$ than $\{t_n^1\}$.

Let us define

$$l_G = 3 l_{\mathcal{F}}(z_0) l_{\mathcal{Z}}(z_0) \tag{2.49}$$

and assume that

$$(\mathbf{B}_3) \quad : \quad \|z_1 - z_0\|_{\mathcal{Z}} \leq \eta \leq \frac{3}{2} l_{\mathcal{F}}(z_0) l_L(z_0) \rho(z_0)$$

Then, we have

$$\eta \leq \frac{2}{9 l_G} < \frac{1}{2\gamma l_G} \tag{2.50}$$

which implies that $h_\alpha < 1/2$ where $l = l_G$ and $\alpha = \alpha_{\mathcal{F}}(z_0)$. For simplicity we shall still use the same notation for $\{t_n^\alpha\}$.

We can show the following result for the semilocal convergence of (**GNM**).

Theorem 2.16. *Suppose*

1. *assumptions* ($\mathbf{B}_1$) - ($\mathbf{B}_3$) *are satisfied and*
2. *generalized equation* (2.1) *is strongly regular at* z_0 *where* $\rho(z_0)$ *is defined by* (2.36).

Then the sequence $\{z_k\}$ *generated by (**GNM**) is well-defined, remains in* $U_{\mathcal{Z}}(r_1, r^\star)$ *for all* $k \geq 0$ *and converges to a solution* $\tilde{z} \in U_{\mathcal{Z}}(z_1, r^\star)$ *of the generalized equation* (2.1) *where* $r^\star$ *is given in the Lemma* 2.14. *Moreover, the solution* $\tilde{z}$ *is unique in* $U_z(z_0, \rho(z_0))$ *and*

$$\|\tilde{z} - z_0\|_{\mathcal{Z}} \leq \lambda \|z_1 - z_0\|_{\mathcal{Z}} \tag{2.51}$$

where λ *is given in the Lemma* 2.14.

Proof. Let $\rho = \rho(z_0)$ and set $\delta = 3/2 l_{\mathcal{F}}(z_0) l_L(z_0)\rho$. Then, we have $\delta \le 1/3$ and $\eta \le \delta\rho \le 1/3\rho$. We shall show by induction that

$$\begin{aligned} \|z_{k+1} - z_k\|_{\mathcal{Z}} &\le \rho\delta^{2^{k+1}-1}, \quad \|z_{k+1} - z_0\| &\le \frac{2}{3}\rho \\ \|z_{k+1} - z_k\|_{\mathcal{Z}} &\le t_{k+1}^{\alpha} - t_k^{\alpha}, \quad \|z_k - z_1\| &\le r^{\star} \end{aligned} \tag{2.52}$$

for all $k = 1, 2, \ldots$. It follows from assumption ($\mathbf{B}_3$) that $\|z_1 - z_0\|_{\mathcal{Z}} \le \delta\rho$. Then by Theorem 2.11 and Remark 2.12 $z_2 = S(z_1) \in U_{\mathcal{Z}}(z_0, \rho)$ exists and

$$\begin{aligned} \|z_2 - z_1\|_{\mathcal{Z}} &= \|S(z_1) - S(z_0)\|_{\mathcal{Z}} \\ &\le 3 l_L(z_0) \|l(z_0; S(z_0), z_1) - l(z_0; S(z_0), z_0)\|_{\mathbf{Y}} \\ &= 3 l_L(z_0) \|l(z_0; z_1, z_1) - l(z_0; z_1, z_0)\|_{\mathbf{Y}}. \end{aligned}$$

Using the second assertion in Lemma 2.13 and the definition of δ, we get in turn

$$\|z_2 - z_1\|_{\mathcal{Z}} \le \frac{3}{2}\alpha_{\mathcal{F}}(z_0) l_{\mathcal{F}}(z_0) l_L(z_0) \|z_1 - z_0\|_{\mathcal{Z}}^2 \tag{2.53}$$

and

$$\|z_2 - z_1\|_{\mathcal{Z}} \le \frac{3}{2} l_F(z_0) l_L(z_0) (\delta\rho)^2 = \rho\delta^3 < \delta\rho \le \frac{1}{3}\rho. \tag{2.54}$$

Then we deduce that

$$\|z_2 - z_0\|_{\mathcal{Z}} \le \|z_2 - z_1\| + \|z_1 - z_0\|_{\mathcal{Z}} \le \frac{2}{3}\rho.$$

We also have by (2.53) and the definition of l_G that

$$\begin{aligned} \|z_2 - z_1\|_{\mathcal{Z}} &\le \frac{1}{2}\alpha_{\mathcal{F}}(z_0) l_G \|z_1 - z_0\|_{\mathcal{Z}}^2 \\ &\le \frac{\alpha_{\mathcal{F}}(z_0) l_G \|z_1 - z_0\|_{\mathcal{Z}}^2}{2(1 - \alpha_{\mathcal{F}}(z_0) l_G t_1^{\alpha})} \\ &\le \frac{\alpha_{\mathcal{F}}(z_0) l_G (t_1 - t_0)^2}{2(1 - \alpha_{\mathcal{F}}(z_0) l_G t_1^{\alpha})} = t_2^{\alpha} - t_1^{\alpha} \le t_{\star}^{\alpha} - t_1^{\alpha} = r^{\star}, \end{aligned}$$

since $0 \le \alpha_{\mathcal{F}}(z_0) l_G t_n^{\alpha} < 1$ and $\|z_1 - z_0\|_{\mathcal{Z}} \le \eta = t_1^{\alpha} - t_0^{\alpha}$. This shows that (2.52) holds for $k = 2$. We also have $\|z_2 - z_1\|_{\mathcal{Z}} \le \delta\rho$. Hence, $z_3 = S(z_2) \in U(z_0, \rho)$ exists and

$$\begin{aligned} \|z_3 - z_2\|_{\mathcal{Z}} &= \|S(z_2) - S(z_1)\|_{\mathcal{Z}} \\ &\le l_{\mathcal{F}}(z_0) \|l(z_1; S(z_1), z_2) - l(z_1; S(z_1), z_1)\|_{\mathbf{Y}} \\ &= 3 l_{\mathcal{F}}(z_0) \|l(z_1; z_2, z_2) - l(z_1; z_2, z_1)\|_{\mathbf{Y}}. \end{aligned}$$

It then follows from the assertion 1 of Lemma 2.12 that

$$\begin{aligned}\|z_3 - z_2\|_{\mathcal{Z}} &\leq \frac{3}{2} l_{\mathcal{F}}(z_0) l_L(z_0) \|z_1 - z_0\|_{\mathcal{Z}}^2 \\ &\leq \frac{3}{2} l_{\mathcal{F}}(z_0) l_{\mathcal{F}}(z_0) (\delta\rho)^3 \\ &< \delta\rho \leq \frac{1}{3}\rho.\end{aligned}$$

Thus, we get

$$\|z_3 - z_0\|_{\mathcal{Z}} \leq \|z_3 - z_2\|_{\mathcal{Z}} + \|z_2 - z_1\|_{\mathcal{Z}} \leq \frac{2}{3}\rho.$$

It follows from (2.52) and the definition of l_G that

$$\begin{aligned}\|z_3 - z_2\|_{\mathcal{Z}} &\leq \frac{1}{2} l_G \|z_2 - z_1\|_{\mathcal{Z}}^2 \\ &\leq \frac{l_G (t_2 - t_1)^2}{2(1 - l_G t_2)} = t_3^\alpha - t_2^\alpha\end{aligned}$$

and

$$\begin{aligned}\|z_3 - z_1\|_{\mathcal{Z}} &\leq \|z_3 - z_2\|_{\mathcal{Z}} + \|z_2 - z_1\|_{\mathcal{Z}} \\ &\leq t_3^\alpha - t_2^\alpha + t_2^\alpha - t_1^\alpha = t_3^\alpha - t_1^\alpha \\ &\leq t_\star^\alpha - t_1^\alpha = r_\alpha^\star = r^\star.\end{aligned}$$

This shows that (2.52) holds for $k = 3$. Suppose that (2.52) holds for $k \leq n$. Then, for $k = n$

$$\begin{aligned}\|z_n - z_1\|_{\mathcal{Z}} &\leq \sum_{j=2}^{n} \|z_j - z_{j-1}\|_{\mathcal{Z}} \\ &\leq \sum_{j=2}^{n} (t_j^\alpha - t_{j-1}^\alpha) = t_n^\alpha - t_1^\alpha \leq t_\star^\alpha - t_1^\alpha = t_\star^\alpha - \eta = r^\star.\end{aligned}$$

We have $z_n \in U_{\mathcal{Z}}(z_0, 2/3\rho)$, thus $z_{n+1} = S(z_n)$ exists and

$$\begin{aligned}\|z_{n+1} - z_n\|_{\mathcal{Z}} &= \|S(z_n) - S(z_{n-1})\|_{\mathcal{Z}} \\ &\leq 3 l_L(z_0) \|l(z_0; S(z_{n-1}), z_n) - l(z_0; S(z_{n-1}), z_{n-1})\| \\ &\leq \frac{3}{2} l_{\mathcal{F}}(z_0) l_L(z_0) \|z_n - z_{n-1}\|_{\mathcal{Z}}^2.\end{aligned}$$

As for $k = 3$, we show that (2.52) holds for $k = n + 1$. It follows from the Lemma 2.14 that sequence $\{t_n^\alpha\}$ is a complete sequence. By (2.52) $\{z_k\}$ is a complete sequence too in a Banach space and as such it converges to some $\tilde{z} \in U_z(z_1, r^\star)$. By the definition of (**GNM**)

$$\mathcal{F}(z_k) + \mathcal{F}'(z_k)(z_{k+1} - z_k) \in \mathbf{T}(z_{k+1})$$

and by assumption (**A**), **T** has a closed graph. Hence, we get

$$\mathcal{F}(\tilde{z}) \in \mathbf{T}(\tilde{z}) \quad \text{as} \quad k \to \infty.$$

We also have $\tilde{z} \in U(z_0, \rho)$ by (2.52). To show uniqueness let $\tilde{y} \in U_{\mathcal{Z}}(z_0, \rho)$ be a solution of generalized equation (2.1). We get in turn

$$\begin{aligned}\|\tilde{y} - \tilde{z}\|_{\mathcal{Z}} &= \|S(\tilde{y}) - S(\tilde{z})\|_{\mathcal{Z}} \\ &\le 3l_L(z_0)\|l(z_0; S(\tilde{y}), \tilde{z}) - l(z_0; S(\tilde{y}), \tilde{y})\|_{\mathbf{Y}} \\ &= 3l_L(z_0)\|l(z_0; \tilde{y}, \tilde{z}) - l(z_0; \tilde{y}, \tilde{y})\|_{\mathbf{Y}}\end{aligned}$$

We also have

$$\begin{aligned}\|\tilde{y} - \tilde{z}\|_{\mathcal{Z}} &\le \|\tilde{y} - z_0\|_{\mathcal{Z}} + \|z_0 - \tilde{z}\|_{\mathcal{Z}} \\ &\le 2\rho \le \frac{4}{9l_{\mathcal{F}}(z_0)l_L(z_0)},\end{aligned}$$

so

$$\begin{aligned}\|\tilde{y} - \tilde{z}\|_{\mathcal{Z}} &\le \frac{3}{2}l_{\mathcal{F}}(z_0)l_L(z_0)\|\tilde{y} - \tilde{z}\|_{\mathcal{Z}}^2 \\ &\le \frac{2}{3}\|\tilde{y} - \tilde{z}\|_{\mathcal{Z}}\end{aligned}$$

which implies $\tilde{y} = \tilde{z}$. We have

$$\begin{aligned}\|\tilde{z} - z_0\|_{\mathcal{Z}} &\le \|\tilde{z} - z_1\|_{\mathcal{Z}} + \|z_1 - z_0\|_{\mathcal{Z}} \\ &\le t_\star^\alpha - t_1^\alpha + t_1^\alpha = t_\star^\alpha.\end{aligned}$$

Estimate (2.51) has been shown in the Lemma 2.14. That completes the proof of the Theorem. □

Remark 2.17.

C_1. If $\alpha = 1$, our Theorem 2.16 reduces to 2.14 in [14]. Otherwise (i.e. if $\alpha \in (0,1)$) according to Remark 2.15 it is an improvement under the same hypotheses and computational cost.

C_2. Similar improvements can be found if $\alpha < 1/3$ and the corresponding estimates are used in the proof.

3. Mesh Independence

The (**GNM**) can rarely be solved in infinite-dimensional spaces thus for practical purposes (2.1) is replaced by a family of discretized equations

$$z \in \mathcal{D}_N, \quad \mathcal{F}_N(z) \in \mathbf{T}_N(z), \tag{3.1}$$

where $N \in \mathbb{N}$, $N \ge \tilde{N}$ for some $\tilde{N} \in \mathbb{N}$. Here, $\mathcal{Z}_N$, $\mathbf{Y}_N$ are finite-dimensional spaces, $\mathcal{D}_N$ is an open subset of $\mathcal{Z}_N$, $\mathcal{F} : \mathcal{D}_N \to \mathbf{Y}_N$ is a mapping and $\mathbf{T}_N : \mathcal{D}_N \to \mathbf{Y}_N$ a multi-valued mapping. Analogous to the discretization methods for operator equations examined by Allgower et al. [9] and Argyros et al. [1]-[8], we also consider the discretization methods in the following form

$$(\mathcal{F}_N, \mathbf{T}_N, h_N, \Delta_N, \tilde{\Delta}_N), \quad N \geq \tilde{N}, \tag{3.2}$$

where $\{h_N\}$ is a sequence of mesh sizes with

$$\lim_{N\to\infty} h_N = 0, \tag{3.3}$$

and $\Delta_N : \mathcal{Z} \to \mathcal{Z}_N$, $\tilde{\Delta} : \mathbf{Y} \to \mathbf{Y}_N$ are bounded linear discretization operators. Applying (**GNM**) to the discrete generalized equations (3.1), we obtain the discrete process

3.1. $(\mathbf{GNM})_N$**:** Select a starting point x_0, $N \in \mathcal{D}_N$. After computing z_k, N, we compute z_{k+1}, N as the solution of the following generalized equation

$$\mathcal{F}_N(z_k, N) + \mathcal{F}'_N(z_k, N)(z - z_k, N) \in \mathbf{T}_N(z). \tag{3.4}$$

We shall examine convergence of these discrete processes and their relations to the infinite ones (**GNM**). Let $\mathcal{Z}_R \subset \mathcal{Z}$ and $\mathbf{Y}_R \subset \mathbf{Y}$ (compare Assumption $(\mathbf{A}_3)$). We use the following assumptions for the discretized generalized equations. Similar assumptions were used in [1].

$\mathbf{D}_1$**.** The mappings $\mathcal{F}_N$ are Fréchet differentiable on $\mathcal{D}_N$, the multi-valued mappings $\mathbf{T}_N$ have closed graph and there exists $r_0 > 0$ such that

$$\Delta_N(U_Z(\tilde{z}, r_0) \cap \mathcal{Z}_R) \subset \mathcal{D}_N, \quad N \geq \tilde{N}.$$

$\mathbf{D}_2$**.** The discretization (3.2) is Lipschitz uniform. That is there exist constants $r_1 > 0$ and $l_{\mathcal{F}} > 0$ such that

$$U_{\mathcal{Z}_N}(\Delta_N(\tilde{z}, r_1)) \subset \mathcal{D}_N, \quad N \geq \tilde{N},$$

and

$$\|\mathcal{F}'_N(z_1) - \mathcal{F}'_N(z_2)\|_{\mathcal{Z}_N \to \mathbf{Y}_N} \leq l_{\mathcal{F}} \|z_1 - z_2\|_{\mathcal{Z}_N}$$

for all $z_1, z_2 \in U_{\mathcal{Z}_N}(\Delta_N(\tilde{z}), r_1)$ and for all $N \geq \tilde{N}$.

$\mathbf{D}_3$**.** The discretization (3.2) is bounded. That is there exist $l_B > 0$ such that

$$\|\Delta_N z\|_{\mathcal{Z}_N} \leq l_B \|z\|$$

for all $z \in \mathcal{Z}_R$ and for all $N \geq \tilde{N}$.

$\mathbf{D}_4$**.** The discretization (3.2) is stable. That is the generalized equation (3.1), $N \geq \tilde{N}$, are uniformly strongly regular at $\Delta_N \tilde{z}$ which requires that for $N \geq \tilde{N}$ the following holds with the constants $r_{\mathbf{Y}}$ and l_L from $(\mathbf{D}_1)$: For each $y \in U_{\mathbf{Y}_N}(0, r_{\mathbf{Y}})$ the linearized system

$$\mathcal{F}_N(\Delta_N \tilde{z}) + \mathcal{F}'_N(\Delta_N \tilde{z})(z - \Delta_N \tilde{z}) + y \in \mathbf{T}_N(z).$$

has a unique solution $S_{L,N}(\Delta_N \tilde{z})$ and $S_{L,N}(\Delta_N \tilde{z}, \cdot) : U_{\mathbf{Y}_N}(0, r_{\mathbf{Y}}) \to \mathcal{Z}_N$ is Lipschitz continuous with modulus l_L.

$\mathbf{D}_5$. The discretization (3.2) is consistent of order p. That is there are constants $c_0, c_1 > 0$ such that

$$\|\tilde{\Delta}_N \mathcal{F}(z) - \mathcal{F}_N(\Delta_N z)\|_{\mathbf{Y}_N} \leq c_0 h_N^p$$

for all $z \in \mathcal{Z}_R \cap U_{\mathcal{Z}}(\tilde{z}, r_1)$, $N \geq \tilde{N}$ and

$$\|\tilde{\Delta}_N(\mathcal{F}'(u)v) - \mathcal{F}'_N(\Delta_N u)(\Delta_N v)\|_{\mathbf{Y}_N} \leq c_1 h_N^p \|v\|_{\mathcal{Z}}$$

for all $u \in \mathcal{Z}_R \cap U_{\mathcal{Z}}(\tilde{z}, r_1)$, $v \in \mathcal{Z}_R$, $N \geq \tilde{N}$.

Remark 3.1. For operator equations, stability requires that the linearized equation in ($\mathbf{D}_4$) has a unique solution for any $y \in \mathbf{Y}_N$. In the more general context considered here this property is required only for $y \in U_{\mathbf{Y}_N}(0, r_Y)$, where $r_{\mathbf{Y}}$ is independent of N. However, in some applications, we can choose $r_{\mathbf{Y}} \to \infty$. ($\mathbf{D}_5$) is the usual definition of consistency for operator equations (see [1]-[9]). For generalized equations, we need the additional assumption ($\mathbf{D}_6$) which is stated below. And which is always satisfied for stable and consistent operator equations.

$\mathbf{D}_6$. There exist constants $c_2, c_3 > 0$ such that the following holds: If $y \in \mathbf{Y}_R \cap U_{\mathbf{Y}}(0, r_{\mathbf{Y}})$ and $\tilde{z} \in \mathcal{Z}_R$ is the solution of the linear generalized equation (2.12) the for each $N \geq \tilde{N}$ there exist $z_N \in \mathcal{Z}_N$ and $y_N \in \mathbf{Y}_N$ such that

$$\|z_N - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \leq c_2 h_N^p, \quad \|y_N - \tilde{\Delta}_N y\|_{\mathbf{Y}_N} \leq c_3 h_N^p,$$

and z_N is the solution of the linear generalized equation

$$\mathcal{F}_N(\Delta_N \tilde{z}) + \mathcal{F}'_N(\Delta_N \tilde{z})(z - \Delta_N \tilde{z}) + y_N \in \mathbf{T}_N(z). \tag{3.5}$$

Using Theorem 2.14 on semilocal convergence of (**GNM**), we first prove existence of solutions of the discretized generalized equations (3.1) and error estimates for the stable and consistent discretizations.

Theorem 3.2. *Let $\tilde{z} \in \mathcal{Z}_R \subset \mathcal{Z}$ be a solution of* (2.1) *and let a discretization be defined by* (3.2) *which satisfies Assumptions* ($\mathbf{D}_1$), ($\mathbf{D}_2$), ($\mathbf{D}_4$) and ($\mathbf{D}_6$) with $\mathbf{Y}_R = \{0_{\mathbf{Y}}\}$. Then there exists $N_1 \geq \tilde{N}$ such that for all $N \geq N_1$ the generalized equation (3.1) has a locally unique solution $\tilde{z}_N$ and

$$\|\tilde{z}_N - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \leq \lambda(c_2 + l_L c_3) h_N^p, \tag{3.6}$$

where $\lambda = \lambda^\alpha$ and is given in the Lemma 2.14.

Proof. Let $N \geq \tilde{N}$. We apply Theorem 2.16 to (3.1) with the starting point $z_{0,N} = \Delta_N \tilde{z}$. The sequence generated by (**GNM**) is denoted by $\{z_k, N\}$. By ($\mathbf{D}_1$) and ($\mathbf{D}_2$) Assumptions ($\mathbf{B}_1$) and ($\mathbf{B}_2$) are satisfied. By Assumption ($\mathbf{D}_4$), (3.1) is strongly regular at $z_{0,N}$. Since $\tilde{z}$ is the solution of the linear generalized equation

$$\mathcal{F}(\tilde{z}) + \mathcal{F}'(\tilde{z})(z - \tilde{z}) + 0_{\mathbf{Y}} \in \mathbf{T}(z),$$

and $\tilde{\Delta}_N 0 = 0$, ($\mathbf{D}_6$) with $\mathbf{Y}_R = \{0_{\mathbf{Y}}\}$ implies that there exist $z_N \in \mathcal{Z}_N$ and $y_N \in \mathbf{Y}_N$ such that

$$\|z_N - \Delta_N z\|_{\mathcal{Z}_N} \leq c_2 h_N^p, \quad \|y_N\|_{\mathbf{Y}_N} \leq c_3 h_N^p, \tag{3.7}$$

and z_N is the solution of the linear generalized equation

$$\mathcal{F}_N(\Delta_N \tilde{z}) + \mathcal{F}'_N(\Delta_N \tilde{z})(z - \Delta_N \tilde{z}) + y_N \in \mathbf{T}_N(z).$$

Since by the definition of (**GNM**), $z_{1,N}$ is the solution of

$$\mathcal{F}_N(\Delta_N \tilde{z}) + \mathcal{F}'_N(\Delta_N \tilde{z})(z - \Delta_N \tilde{z}) + 0 \in \mathbf{T}_N(z).$$

It follows by ($\mathbf{D}_4$) and (3.7) that if $c_3 h_N^p \leq r_{\mathbf{Y}}$ then

$$\|z_{1,N} - z_N\|_{\mathcal{Z}_N} = \|S_{L,N}(\Delta_N \tilde{z}; 0) - S_{L,N}(\Delta_N \tilde{z}; y_N)\|_{\mathcal{Z}_N} \leq l_L \|y\|_{\mathbf{Y}_N} \leq l_L c_3 h_N^p.$$

By (3.7) we further obtain

$$\|z_{1,N} - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \leq \|z_{1,N} - z_N\|_{\mathcal{Z}_N} + \|z_1 - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \leq (c_2 + l_L c_3) h_N^p. \tag{3.8}$$

In order to satisfy Assumption ($\mathbf{B}_3$) we must have

$$b_N = \|z_{1,N} - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \leq \frac{3}{2} l_F l_L \rho^2$$

where

$$\rho = \min \left\{ \frac{1}{2} r_1, \frac{1}{2} \sqrt{\frac{2 r_Y}{\phi_F(\tilde{z}) l_F}}, \frac{2}{9 l_F\, l_L} \right\}.$$

Hence, by (3.8), ($\mathbf{B}_3$), is satisfied if

$$h_N^p \leq \frac{3 l_F l_L \rho^2}{2(c_2 + l_L c_3)}, \quad c_3 h_N^p \leq r_{\mathbf{Y}}. \tag{3.9}$$

By (3.9) there exists $N_1 \geq \tilde{N}$ such that for $N \geq N_1$, (3.9) holds. Therefore if $N \geq N_1$, Theorem 2.16 implies the existence of a solution $\tilde{z}_N$ of (3.1). This solution is unique on $U_{\mathcal{Z}_N}(\Delta_N \tilde{z}, \rho)$ and

$$\|\tilde{z}_N - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \leq \lambda \|z_{1,N} - \Delta_N \tilde{z}\|_{\mathcal{Z}_N}.$$

Together with (3.8) this implies (3.6). □

Based on the local convergence result of (**GNM**), stated in Theorem 2.6, we now investigate relations between the infinite process (**GNM**) and the discrete processes $(\mathbf{GNM})_N$. For stable and consistent discretizations, we show that the local behavior of the discrete Newton iterations is asymptotically the same as that for the original iteration.

Theorem 3.3. *Let the hypotheses of Theorem 2.6 be satisfied. And, let $\mathcal{Z}_R \subset \mathcal{Z}$ be such that* (2.32) *is satisfied for $z_0 \in U_{\mathcal{Z}}(\tilde{z}, \tilde{\rho}(\tilde{z}))$ and $\mathbf{Y}_R \subset \mathbf{Y}$ such that ($\boldsymbol{A}_3$) is satisfied. Further let a discretization be defined by* (3.2) *which satisfies Assumptions ($\boldsymbol{D}_1$)-($\boldsymbol{D}_6$). Then there exists $N_2 \geq N_1$, $\rho_2 > 0$ such that the sequence $\{z_{N,k}\}_{k\in\mathbb{N}}$ generated by $(\boldsymbol{GNM})_N$ with starting point $\Delta_N z_0$ converges to $\tilde{z}_N$ and that*

$$\|\tilde{z}_N - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \leq \tilde{c}_1\, h_N^p, \quad k = 0, 1, 2, \ldots, \tag{3.10}$$

for all $N \geq N_2$ and all $z_0 \in \mathcal{Z}_R \cap U_{\mathcal{Z}}(\tilde{z}, \rho_2)$.

Proof. Without loss of generality, we may assume that $r_1(\tilde{z}) = r_1$, $l_F(\tilde{z}) = l_F$, $l_L(\tilde{z}) = l_L$, $\phi_F(\tilde{z}) = \phi_F$ and $r_{\mathbf{Y}}(\tilde{z}) = r_{\mathbf{Y}}$. Let $N \geq N_1$. By Theorem 3.2 the generalized equation (3.1) has a locally unique solution $\tilde{z}_N$. We define

$$\rho = \min\left\{r_1, \sqrt{\frac{2r_{\mathbf{Y}}}{3\phi_F l_F}}, \frac{2}{3l_L l_F}\right\}, \quad \tilde{\rho} = \frac{3}{8} l_F l_L \rho^2, \quad \rho_1 = \min\left\{\frac{\tilde{\rho}}{2}, \frac{\tilde{\rho}}{2l_B}\right\}.$$

By Theorem 2.6 the sequence $\{z_{N,k}\}_{k\in\mathbb{N}}$ generated by $(\mathbf{GNM})_N$ with starting point $\Delta_N z_0$ converges to $\tilde{z}_N$, if

$$\|\Delta_N z_0 - \tilde{z}_N\|_{\mathcal{Z}} \leq \tilde{\rho}. \tag{3.11}$$

By (3.3) there exists $M_1 \geq N_1$ such that

$$\lambda(c_2 + l_L c_3) h_N^p \leq \rho_1 \quad \text{for all} \quad N \geq M_1.$$

Then by ($\mathbf{D}_3$) and Theorem 3.2 we obtain

$$\begin{aligned}\|\Delta_N z_0 - \tilde{z}_N\|_{\mathcal{Z}} &\leq \|\Delta_N z_0 - \Delta_N \tilde{z}\|_{\mathcal{Z}} + \|\Delta_N \tilde{z} - \tilde{z}_N\|_{\mathcal{Z}} \\ &\leq l_B \|z_0 - \tilde{z}\|_{\mathcal{Z}} + \lambda(c_2 + l_L c_3) h_N^p \leq \tilde{\rho}.\end{aligned}$$

Therefore, (3.11) is satisfied, if $z_0 \in U_{\mathcal{Z}}(\tilde{z}, \rho_1)$ and if $N \geq M_1$. In order to prove (3.10), we define

$$\rho_2 = \min\left\{\rho_1, \frac{1}{18 l_F l_L l_B}\right\} \tag{3.12}$$

and

$$\tilde{c}_1 = 24 l_L(c_0 + c_1\rho_2) + 2\lambda(c_2 + l_L c_3). \tag{3.13}$$

Further, we choose $N_2 \geq M_1$ such that

$$l_F \tilde{c}_1 h_N^p \leq \frac{1}{6l_L} \quad \text{for} \quad N \geq N_2. \tag{3.14}$$

Now let $N \geq N_2$ and $z_0 \in U_{\mathcal{Z}}(\tilde{z}, \rho_2)$ be given. For $k = 0$ we have $z_{0,N} = \Delta_N(z_0)$. Hence (3.10) is satisfied. Suppose that (3.10) holds for $k = 0, 1, \ldots, n$. By the definition of $(\mathbf{GNM})_N$, $z_{n+1,N}$ is the solution of

$$\mathcal{F}_N(\Delta_N \tilde{z}) + \mathcal{F}'_N(\Delta_N \tilde{z})(z - \Delta_N \tilde{z}) + v_N \in \mathbf{T}_N(z), \tag{3.15}$$

where

$$v_N = \mathcal{F}_N(z_{n,N}) + \mathcal{F}'_N(z_{n,N})(z_{n+1,N} - z_{n,N}) - \mathcal{F}_N(\Delta_N \tilde{z}) - \mathcal{F}'_N(\Delta_N \tilde{z})(z_{n+1,N} - \Delta_N \tilde{z}).$$

By the definition of $(\mathbf{GNM})$, z_{n+1} is the solution of

$$\mathcal{F}(\tilde{z}) + \mathcal{F}'(\tilde{z})(z - \tilde{z}) + y \in \mathbf{T}(z),$$

where

$$y = \mathcal{F}(z_n) + \mathcal{F}'(z_n)(z_{n+1} - z_n) - \mathcal{F}(\tilde{z}) - \mathcal{F}'(\tilde{z})(z_{n+1} - \tilde{z}) = l(\tilde{z}; z_{n+1}, z_n) \in \mathbf{Y}_R.$$

By Assumption $(\mathbf{D}_6)$ there exist $w_N \in \mathcal{Z}_N$ and $y_N \in \mathbf{Y}_N$ such that

$$\|w_N - \Delta_N z_{n+1}\|_{\mathcal{Z}_N} \leq c_2 h_N^p, \quad \|y_N - \tilde{\Delta}_N y\|_{\mathbf{Y}_N} \leq c_3 h_N^p,$$

and w_N is the solution of

$$\mathcal{F}_N(\Delta_N \tilde{z}) + \mathcal{F}'_N(\Delta_N \tilde{z})(z - \Delta_N \tilde{z}) + y_N \in \mathbf{T}_N(z). \tag{3.16}$$

Thus we obtain

$$\begin{aligned} \|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N} &\leq \|z_{n+1,N} - w_N\|_{\mathcal{Z}_N} + \|w_N - \Delta_N z_{n+1}\|_{\mathcal{Z}_N} \\ &\leq \|z_{n+1,N} - w_N\|_{\mathcal{Z}_N} + c_2 h_N^p. \end{aligned}$$

Since $z_{n+1,N}$ is the solution of (3.15) and w_N is the solution of (3.16), it follows from Assumption $(\mathbf{D}_4)$ that

$$\begin{aligned} \|z_{n+1,N} - w_N\|_{\mathcal{Z}_N} &\leq l_L \|v_N - y_N\|_{\mathbf{Y}_N} \\ &\leq l_L \|v_N - \tilde{\Delta}_N y\|_{\mathbf{Y}_N} + l_L \|\tilde{\Delta}_N y - y_N\|_{\mathbf{Y}_N} \\ &\leq l_L \|v_N - \tilde{\Delta}_N y\|_{\mathbf{Y}_N} + l_L c_3 h_N^p. \end{aligned}$$

Therefore, we have

$$\|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N} \leq l_L \|v_N - \tilde{\Delta}_N y\|_{\mathbf{Y}_N} + (c_2 + l_L c_3) h_N^p. \tag{3.17}$$

A simple calculation shows that

$$v_N - \tilde{\Delta}_N y = E_1 + E_2 + E_3 + E_4. \tag{3.18}$$

where

$$E_1 = \mathcal{F}_N(z_{n,N}) + \mathcal{F}'_N(z_{n,N})(\Delta_N z_n - z_{n,N}) - \mathcal{F}_N(\Delta_N z_n),$$

$$\begin{aligned}
E_2 &= \mathcal{F}_N(\Delta_N z_n) - \tilde{\Delta}_N \mathcal{F}_N(z_n) - \mathcal{F}_N(\Delta_N \tilde{z}) \\
&\quad + \tilde{\Delta}_N \mathcal{F}_N(\tilde{z}) + \mathcal{F}'_N(\Delta_N z_n)(\Delta_N z_{n+1} - \Delta_N z_n) \\
&\quad - \tilde{\Delta}_N \mathcal{F}'(z_n)(z_{n+1} - z_n) - \mathcal{F}'_N(\Delta_N \tilde{z})(\Delta_N z_{n+1} - \Delta_N \tilde{z}) \\
&\quad + \tilde{\Delta}_N \mathcal{F}'(\tilde{z})(z_{n+1} - \tilde{z}), \\
E_3 &= \left[\mathcal{F}'_N(z_{n,N}) - \mathcal{F}'_N(\Delta_N z_n)\right](\Delta_N z_{n+1} - \Delta_N z_n), \\
E_4 &= \left[\mathcal{F}'_N(z_{n,N}) - \mathcal{F}'_N(\Delta_N \tilde{z})\right](z_{n+1,N} - \Delta_N z_{n+1}).
\end{aligned}$$

From this and Assumptions ($\mathbf{D}_1$), ($\mathbf{D}_2$) we obtain the estimate

$$\|E_1\|_{\mathbf{Y}_N} \le \frac{1}{2} l_{\mathcal{F}} \|\Delta_N z_n - z_{n,N}\|^2_{\mathcal{Z}_N}.$$

By the induction assumption and (3.14) it follows then that

$$\|E_1\|_{\mathbf{Y}_N} \le \frac{l_F}{2}(\tilde{c}_1 h_N^p)^2 \le \frac{1}{12 l_L} \tilde{c}_1 h_N^p.$$

By Assumption ($\mathbf{D}_5$) we have

$$\|E_2\|_{\mathbf{Y}_N} \le 2c_0 h_N^p + c_1 h_N^p \|z_{n+1} - z_n\|_{\mathcal{Z}} + c_1 h_N^p \|z_{n+1} - \tilde{z}\|_{\mathcal{Z}}.$$

Since by (2.30)

$$\|z_{n+1} - \tilde{z}\|_{\mathcal{Z}} \le \frac{1}{2}\|z_0 - \tilde{z}\|_{\mathcal{Z}} \le \frac{1}{2}\rho_2,$$

and

$$\|z_{n+1} - z_n\|_{\mathcal{Z}} \le \|z_{n+1} - \tilde{z}\|_{\mathcal{Z}} + \|z_n - \tilde{z}\|_{\mathcal{Z}} \le \frac{3}{2}\|z_0 - \tilde{z}\|_{\mathcal{Z}} \le \frac{3}{2}\rho_2 \tag{3.19}$$

we obtain

$$\|E_2\|_{\mathbf{Y}_N} \le 2(c_0 + c_1\rho_2) h_N^p.$$

By (3.13) this implies that

$$\|E_2\|_{\mathbf{Y}_N} \le \frac{1}{12 l_L} \tilde{c}_1 h_N^p.$$

By Assumptions ($\mathbf{D}_1$), ($\mathbf{D}_2$) we have

$$\|E_3\|_{\mathbf{Y}_N} \le l_F \|z_{n,N} - \Delta_N z_n\|_{\mathcal{Z}_N} \|\Delta_N z_{n+1} - \Delta_N z_n\|_{\mathcal{Z}_N}.$$

Using the induction assumption and ($\mathbf{D}_3$) we obtain

$$\|E_3\|_{\mathbf{Y}_N} \le l_F \tilde{c}_1 h_N^p l_B \|z_{n+1} - z_n\|_{\mathcal{Z}}.$$

By (3.19) and (3.12) this implies that

$$\|E_3\|_{\mathbf{Y}_N} \le \frac{3}{2} l_F \tilde{c}_1 h_N^p l_B \rho_2 \le \frac{1}{12 l_L} \tilde{c}_1 h_N^p.$$

By Assumptions ($\mathbf{D}_1$) , ($\mathbf{D}_2$) we have

$$\|E_4\|_{\mathbf{Y}_N} \le l_F \|z_{n,N} - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N}.$$

Using the induction assumption, $(\mathbf{D}_3)$ and (2.30) (compare (3.19)) we obtain

$$\begin{aligned}\|z_{n,N} - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} &\le \|z_{n,N} - \Delta_N z_n\|_{\mathcal{Z}_N} + \|\Delta_N z_n - \Delta_N \tilde{z}\|_{\mathcal{Z}_N} \\ &\le \tilde{c}_1 h_N^p + l_B \|z_n - \tilde{z}\|_{\mathcal{Z}} \le \tilde{c}_1 h_N^p + l_B \rho_2.\end{aligned} \tag{3.20}$$

Hence, E_4 can be estimated by

$$\|E_4\|_{\mathbf{Y}_N} \le l_F(\tilde{c}_1 h_N^p + l_B \rho_2)\|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N}.$$

By (3.12) and (3.14) this implies that

$$\begin{aligned}\|E_4\|_{\mathbf{Y}_N} &\le \left(\frac{1}{6l_L} + \frac{1}{18l_L}\right)\|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N} \\ &\le \frac{1}{2l_L}\|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N}.\end{aligned}$$

From (3.17), (3.18) and from the estimates for E_i, $i = 1, \dots, 4$, we finally obtain

$$\begin{aligned}\|z_{n+1,N} - \Delta_N z_{n+1}\|_{Z_N} &\le l_L(\|E_1\| + \|E_2\| + \|E_3\| + \|E_4\|) + \frac{\lambda}{2}(c_2 + l_L c_3)h_N^p \\ &\le \frac{1}{4}\tilde{c}_1 h_N^p + \frac{1}{2}\|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N} + \frac{\lambda}{2}(c_2 + l_L c_3)h_N^p,\end{aligned}$$

which implies

$$\|z_{n+1,N} - \Delta_N z_{n+1}\|_{\mathcal{Z}_N} \le \frac{1}{2}\tilde{c}_1 h_N^p + \lambda(c_2 + l_L c_3)h_N^p.$$

This complete the induction. □

In view of the mesh-independence principle we need some additional estimates which are easily obtained from Theorem 3.3.

Corollary 3.4. *Let the hypothesis of Theorem* 3.3 *be satisfied. Then there are constants* $\tilde{c}_2$, $\tilde{c}_3$ *such that*

$$\|\mathcal{F}_N(z_{k,N}) - \tilde{Delta}_N \mathcal{F}(z_k)\|_{\mathbf{Y}_N} \le \tilde{c}_2 h_N^p, \quad k = 0, 1, \dots, \tag{3.21}$$

and

$$\|z_{k,N} - \tilde{z}_N - \Delta_N(z_k - \tilde{z})\|_{\mathcal{Z}_N} \le \tilde{c}_3 h_N^p, \quad k = 0, 1, \dots, \tag{3.22}$$

for all $N \ge N_2$ *and for all* $z_0 \in \mathcal{Z}_R \cap U_{\mathcal{Z}}(\tilde{z}, \rho_2)$.

Proof. By Assumptions $(\mathbf{D}_2)$ there exists $\tilde{l}_{\mathcal{F}}$ such that

$$\|\mathcal{F}'(z)\|_{\mathcal{Z}_n \to \mathbf{Y}_N} \le \tilde{l}_F \quad \text{for all} \quad z \in U_{\mathcal{Z}_N}(\Delta_N \tilde{z}, r_1).$$

Together with $(\mathbf{D}_5)$ we obtain

$$\begin{aligned}&\|\mathcal{F}_N(z_k,N)-\tilde{\Delta}_N\mathcal{F}(z_k)\|_{\mathbf{Y}_N}\\&\le\|\mathcal{F}_N(z_k,N)-\mathcal{F}_N(\Delta_N z_k)\|_{\mathbf{Y}_N}+\|\mathcal{F}_N(\Delta_N z_k)-\tilde{\Delta}_N\mathcal{F}(z_k)\|_{\mathbf{Y}_N}\\&\le\tilde{l}_{\mathcal{F}}\|z_{k,N}-\Delta_N z_k\|_{\mathcal{Z}_N}+c_0h_N^p.\end{aligned}$$

By Theorem 3.3 this implies (3.21). Since

$$\|z_{k,N}-\tilde{z}_N-\Delta_N(z_k-\tilde{z})\|_{\mathcal{Z}_N}\le\|z_{k,N}-\Delta_N z_k\|_{\mathcal{Z}_N}+\|\tilde{z}_N-\Delta_N\tilde{z}\|_{\mathcal{Z}_N}$$

inequality (3.22) follows from Theorem 3.3 and Theorem 3.2. □

As a consequence of the preceding results we can now prove a mesh-independence principles, which states that for sufficiently large N there is at most a difference of one between the number of iteration steps required by the two processes (**GNM**) and $(\mathbf{GNM})_N$ to converge to within a given tolerance $\epsilon>0$. The proof is a slight modification of the proof in [1][Corollary 1] or [9][Corollary 1].

Theorem 3.5. *Suppose that the hypotheses of Theorem 3.3 hold and that there is a constant* $l_D>0$ *for which*

$$\lim_{N\ge N_1}\inf\|\Delta_N z\|\ge 2l_D\|z\|_{\mathcal{Z}}\quad\text{for each}\quad z\in\mathcal{Z}_R.\tag{3.23}$$

Then for some $\rho_3\in]0,\rho_2]$ *and for any fixed* $\epsilon>0$ *and* $z_0\in U_{\mathcal{Z}}(\tilde{z},\rho_3)$ *there is a* $N_3=N_3(z_0,\epsilon)$ *such that*

$$|\min\{k\ge 0,\|z_k-\tilde{z}\|_{\mathcal{Z}}<\epsilon\}-\min\{k\ge 0,\|z_{k,N}-\tilde{z}_N\|_{\_N}<\epsilon\}|\le 1\tag{3.24}$$

for all $N\ge N_3$.

Proof. Let i be the unique integer defined by

$$\|z_{i+1}-\tilde{z}\|_{\mathcal{Z}}<\epsilon\le\|z_i-\tilde{z}\|_{\mathcal{Z}}\tag{3.25}$$

(compare Remark 2.7). By (3.23) there exists $M\ge N_2$ such that

$$\|\Delta_N(z_i-\tilde{z})\|\ge l_D\|z_i-\tilde{z}\|_{\mathcal{Z}}\tag{3.26}$$

for $N\ge M$. We choose $N_3\ge M$ such that

$$\max\{\tilde{c}_2,2\tilde{c}_3\}h_N^p\le l_B\epsilon,\quad\max\left\{3l_{\mathcal{F}}l_Ll_B\tilde{c}_1,\frac{3l_{\mathcal{F}}l_L\tilde{c}_1}{2l_D}\right\}h_N^p\le\frac{1}{4}\tag{3.27}$$

for $n\ge N_3$ and $0<\rho_2\le\rho_2$

$$\max\left\{3l_Fl_Ll_B^2,\frac{3l_Fl_Ll_B}{2l_D}\right\}\rho_3\le\frac{1}{4}.\tag{3.28}$$

By (3.22), Assumption $(\mathbf{D}_3)$ and (3.27)

$$\|z_{i+1,N}-\tilde{z}_N\|\le\|\Delta_N(z_{i+1}-\tilde{z})\|_{\mathcal{Z}_N}+\tilde{c}_3h_N^p\le l_B\epsilon+\tilde{c}_3h_N^p\le 2l_B\epsilon.$$

By Theorem 2.6 and (3.20) it then follows that

$$\|z_{i+2,N} - \tilde{z}_N\| \leq \frac{3}{2} l_F l_L \|z_{i+1,N} - \tilde{z}_N\|^2_{\mathcal{Z}_N} \leq \frac{3}{2} l_F l_L (\tilde{c}_1 h_N^p + l_B \rho_2) 2 l_B \epsilon,$$

and from (3.27), (3.28) we obtain

$$\|z_{i+2,N} - \tilde{z}_N\|_{\mathcal{Z}_N} \leq \frac{1}{2}\epsilon < \epsilon. \tag{3.29}$$

Because of (3.25) and (3.22) we have

$$\epsilon \leq \|z_i - \tilde{z}\|_{\mathcal{Z}} \leq \frac{1}{l_D}\|\Delta_N(z_i - \tilde{z})\|_{\mathcal{Z}_N} \leq \frac{1}{l_D}\|z_{i,N} - \tilde{z}_N\|_{\mathcal{Z}_N} + \tilde{c}_3 h_N^p,$$

or using (3.27)

$$\|z_{i,N} - \tilde{z}_N\|_{\mathcal{Z}_N} \geq l_D \epsilon - \tilde{c}_3 h_N^p \geq l_D \epsilon - \frac{l_D}{2}\epsilon = \frac{l_D}{2}\epsilon. \tag{3.30}$$

If $\|z_{i-1,N} - \tilde{z}_N\|_{\mathcal{Z}_N} < \epsilon$ then by (3.27) and (3.28) we obtain analogously to (3.29)

$$\|z_{i,N} - \tilde{z}_N\|_{\mathcal{Z}_N} \leq \frac{l_D}{2}\epsilon,$$

which contradicts (3.20). Therefore, we must have

$$\|z_{i,N} - \tilde{z}_N\|_{\mathcal{Z}_N} \geq \epsilon, \tag{3.31}$$

and it is easily seen that (3.25), (3.29) and (3.31) imply (3.24).

As in the case of operator equations, condition (3.23) is an immediate consequence of the convergence condition

$$\lim_{N\to\infty} \|\Delta_N z\|_{\mathcal{Z}_N} = \|z\|_{\mathcal{Z}} \qquad \text{for each} \qquad z \in \mathcal{Z}_R.$$

Moreover, for some discretization we have

$$\lim_{N\to\infty} \|\Delta_N z\|_{\mathcal{Z}_N} = \|z\|_{\mathcal{Z}} \qquad \text{uniformly for} \qquad z \in \mathcal{Z}_R. \tag{3.32}$$

In such cases the following stronger formulation of the mesh-independence principle applies, where N_3 is independent of the starting point (compare Argyros [1]-[8] or Allgower et al. [9] and Corollary 2). □

Corollary 3.6. *Suppose that the hypothesis of Theorem 3.5 is satisfied and that* (3.32) *holds. Then there exists a constant* $\rho_3 \in]0, \rho_2]$ *and for any fixed* $\epsilon > 0$ *there exists some* $N_3 = N_3(\epsilon)$ *such that* (3.24) *holds for all* $N \geq N_3$ *and all starting points* $z_0 \in U_{\mathcal{Z}}(\tilde{z}, \rho_3)$.

Remark 3.7. The results of the Section 3 reduce to the corresponding ones in [14] for $\alpha = 1$. Otherwise (i.e. if $\alpha \neq 1$) they constitute an improvement (see also Remark 2.17). Numerical examples where assumptions ($\mathbf{A}_1$) - ($\mathbf{A}_3$) and ($\mathbf{D}_1$) - ($\mathbf{D}_5$) are verified can be found in Hager et al. [22] and Dontcher et al. [16]. Finally examples where (**A**), (**A′**) verified and $\alpha < 1$ can be found in [1]-[8].

References

[1] I.K. Argyros, *An Improved Mesh Independence Principle for Solving Equations and their Discretizations using Newton's Method*, The Australian Journal of Mathematical Analysis and Applications, 6, (2009), 1–11.

[2] I.K. Argyros, *On a new Newton-Mysovskii-type theorem with applications to inexact Newton-like methods and their discretizations*, The IMA Journal of Numerical Analysis, 18 (1), (1997), 37–56.

[3] I.K. Argyros, *On the convergence of Newton's method for a class of nonsmooth operators*, Journal of Computational and Applied Mathematics, 205(1), (2007), 584–593.

[4] I.K. Argyros, *On the Secant method for solving nonsmooth equations*, 322, (2006), 146–157.

[5] I.K. Argyros, *A mesh independence principle for inexact Newton-like methods and their discretizations under generalized Lipschitz conditions*, Applied Mathematics and Computation, 87, (1997), 15–48.

[6] I.K. Argyros, *A mesh independence principle for inexact Newton type methods and their discretizations analysis*, Annales Univ. Sci. Budapest., Sect. Comp., 20, (2001) 31–53.

[7] I.K. Argyros, *A finer mesh independence of Newton's method*, Nonlinear Funct. Anal. Appl. 13, 3, 357–365 (2008).

[8] I.K. Argyros, Y.J. Cho and S. Hilout, *Numerical methods for equations and its applications*, CRC Press/Taylor and Francis Group, New York, 2012.

[9] E.L. Allgower, K. Bohmer, F.A. Potra and W.C. Rheinboldt, *A Mesh-Independence Principle for Operator Equations and Their Discretizations*, SIAM Journal on Numerical Analysis, 23 (1) (1986), 160–169.

[10] W. Alt, *Local convergence of the Lagrange-Newton method with applications to optimal control*, Control and Cybernetics, 23(12), (1994) 87–105.

[11] W. Alt, *Semi-local convergence of the Lagrange-Newton method with ap- plication to optimal control*, In R. Durier and Ch. Michelot, editors, Recent Developments in Optimization, Volume 429 of Lecture Notes in Economics and Mathematical Systems, pages 1–16. Springer-Verlag, 1995.

[12] W. Alt and K. Malanowski. *The Lagrange-Newton method for nonlinear op- timal control problems*, Computational Optimization and Applications, 2, 77–100, 1993.

[13] W. Alt and K. Malanowski, *The Lagrange-Newton method for state constrained optimal control problems*, Computational Optimization and Applications Volume 4, Number 3 (1995), 217-239.

[14] W. Alt, *Discretization and mesh-independence of Newton's method for generalized equations*, Mathematical Programming with Data Perturbations, (1998), 1–30.

[15] J.F. Bonnans, *Local analysis of Newton-type methods for variational inequalities and nonlinear programming*, Applied Mathematics & Optimization, 29 (2), (1994), 161–186.

[16] A.L. Dontchev and W.W. Hager, *Lipschitz stability in nonlinear control and optimization*, SIAM J. Control and Optimization, 31 (1993), 569–603.

[17] A.L. Dontchev and W.W. Hager, *Implicit Functions, Lipschitz Maps, and Stability in Optimization*, Mathematics of Operations Research, 2001 26:105-118.

[18] A.L. Dontchev, W.W. Hager, A.B. Poore and Bing Yang, *Optimality, stability, and convergence in nonlinear control*, Applied Mathematics & Optimization Volume 31, Number 3 (1995), 297-326.

[19] J.C. Dunn, and T. Tian, *Variants of the Kuhn - Tucker Sufficient Conditions in Cones of Nonnegative Functions*, SIAM J. Control Optim., 30(6), 13611384.

[20] A.V. Fiacco, Introduction to sensitivity and stability analysis in nonlinear programming, Academic Press, New York, 1983.

[21] R. Fletcher, Practical Methods of Optimization, John Wiley & SOns, New York, second edition, 1987.

[22] W.W. Hager, *Multiplier Methods for Nonlinear Optimal Control*, SIAM J. Numer. Anal., 27(4), 1061–1080.

[23] A.D. Ioffe and V.M. Tihomirov, Theory of Extremal Problems, North Holland, Amsterdam, 1979.

[24] N.H. Josephy, *Newton's method for generalized equations.* Technical Summary Report no. 1965. Mathematics Research Center, University of Wisconsin, Madison, 1979.

[25] C.T. Kelley and E.W. Sachs, *Mesh Independence of the Gradient Projection Method for Optimal Control Problems*, SIAM J. Control Optim., 30(2), 477–493, (1992).

[26] E.S. Levitin and B.T. Polyak, *Constrained minimization methods*, USSR Computational Mathematics and Mathematical Physics, 1966, 6(5), 1–50.

[27] K.C.P Machielsen, *Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space*, CWI, Tract, 53, 1987.

[28] K. Malanowski, C. Büskens and H. Maurer, *Convergence of approximations to nonlinear optimal control problems*, In Mathematical programming with data perturbations, Vol. 195 (1998), pp. 253-284.

[29] Malanowski, K, *Stability and sensitivity analysis of solutions to nonlinear optimal control problems*, Appl. Math, and Optim., 32:111–141, 1995.

[30] H. Maurer and H. J. Pesch, *Solution Differentiability for Nonlinear Parametric Control Problems*, SIAM J. Control Optim., 32(6), 15421554.

[31] H. Maurer and H.J. Pesch, *Solution Differentiability for Parametric Nonlinear Control Problems with Control-State Constraints*, Journal of Optimization Theory and Applications, 86(2), 285–309, (1995).

[32] S.M. Robinson, *Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear-programming algorithms*, Mathematical Programming Volume 7, Number 1 (1974), 1-16.

[33] S.M. Robinson, *Strongly Regular Generalized Equations*, Mathematics of Operations Research, 5:43–62, 1980.

[34] S.M. Robinson, *Generalized equations*, In A. Bachem, M. Grötschel, and B. Korte, editors, Mathematical Programming: The State of the Art, pages 165–207, Berlin, 1983, Springer–Verlag.

[35] S.M. Robinson, *Newton's method for a class of nonsmooth functions*, Set-Valued Analysis Volume 2, Numbers 1-2 (1994), 291-305.

[36] J. Stoer, *Principles of Sequential Quadratic Programming Methods for Solving Nonlinear Programs*, In: Computational Mathematical Programming, K. Schittkowski, editor, 23/07 - 02/08, 1984.

[37] E. Zeidler, Nonlinear Functional Analysis and its Applications: I: Fixed-Point Theorems. Springer Verlag, New Yourk, 1991.