Mathematical Programming 73 (1996) 251–268

Generalized equations and the generalized Newton method

Livinus U. Uko [1]

International Centre For Theoretical Physics, Trieste, Italy

Received 26 May 1995; revised manuscript received 30 June 1995

Abstract

We give some convergence results on the generalized Newton method (referred to by some authors as Newton's method) and the chord method when applied to generalized equations. The main results of the paper extend the classical Kantorovich results on Newton's method to (nonsmooth) generalized equations. Our results also extend earlier results on nonsmooth equations due to Eaves, Robinson, Josephy, Pang and Chan.

We also propose inner-iterative schemes for the computation of the generalized Newton iterates. These schemes generalize popular iterative methods (Richardson's method, Jacobi's method and the Gauss–Seidel method) for the solution of linear equations and linear complementarity problems and are shown to be convergent under natural generalizations of classical convergence criteria.

Our results are applicable to equations involving single-valued functions and also to a class of generalized equations which includes variational inequalities, nonlinear complementarity problems and some nonsmooth convex minimization problems.

Keywords: Generalized equations; Generalized Newton method; Variational inequalities; Nonlinear complementarity problem; Kantorovich theorem

1. Introduction

Let H be a Hilbert space equiped with a scalar product $(\cdot, \cdot)$, let $f: H \mapsto H$ be a Fréchet-differentiable function and let g be a nonempty subset of $H \times H$. In the sequel, we will regard the statements $[x, y] \in g$, $g(x) \ni y$, $-y + g(x) \ni 0$ and $y \in g(x)$ as synonymous.

[1] Permanent address: Mathematics Department, University of Ibadan, Ibadan, Nigeria.

SSDI 0025-5610(95)00055-0

We are interested in the (numerical) solution of the problem

$$f(u)+g(u)\ni 0. \tag{1.1}$$

Such problems have been studied by Robinson [31–33] who coined the term "generalized equations" for them. If $H=\mathbb{R}^n$ and $g(x_1,\dots,x_n)\equiv g_1(x_1)\times\cdots\times g_n(x_n)$ where $g_i\in\mathbb{R}\times\mathbb{R}$ for $i=1,\dots,n$. then (1.1) will be said to be separable. In this case, if we set

$$f(x_1,\dots,x_n)\equiv(f_1(x_1,\dots,x_n),\dots,f_n(x_1,\dots,x_n)),$$

then we can express (1.1) in the form

$$f_i(u_1,\dots,u_n)+g_i(u_i)\ni 0,\quad i=1,\dots,n. \tag{1.2}$$

An important class of problems of this type is obtained by taking

$$g_i=\{0\}\times(-\infty,0]\cup(0,\infty)\times\{0\}=\{[s,t]\in\mathbb{R}\times\mathbb{R}: s\geqslant 0,\ t\leqslant 0,\ st=0\} \tag{1.3}$$

for all i. In this case (1.2) becomes the nonlinear complementarity problem

$$f_i(u_1,\dots,u_n)\geqslant 0,\quad u_i\geqslant 0,\ i=1,\dots,n,\ \sum_{k=1}^{n}u_kf_k(u_1,\dots,u_n)=0. \tag{1.4}$$

Such problems have been studied extensively in the literature from the point of view of existence of solutions and approximation of solutions (cf. [8,13,19]).

On the other hand, if $\phi: H\mapsto(-\infty,\infty]$ is a proper lower semicontinuous convex function and

$$g(x)=\partial\phi(x)\equiv\{v\in H:\phi(x)-\phi(y)\leqslant(v,x-y),\ \forall y\in H\}$$

(called the subgradient of ϕ at x), then (1.1) becomes the variational inequality

$$f(u)+\partial\phi(u)\ni 0. \tag{1.5}$$

Such problems were introduced in the early sixties by Stampacchia [34] and have found important applications in the physical and engineering sciences and in many other fields [1–3,6,11].

If we take f as the gradient $\nabla\psi$ of a differentiable convex function $\psi: H\mapsto\mathbb{R}$, then (1.5) reduces (cf. [1, Theorem 3.3]) to the search for the minimum of the nonsmooth convex function $\psi+\phi$.

A basic class of variational inequalities is obtained by letting ϕ be the indicator function of a nonempty closed convex subset C of H, defined as

$$\phi(x)=\begin{cases}0, & \text{if } x\in C,\\ \infty, & \text{otherwise.}\end{cases} \tag{1.6}$$

In this case problem (1.5) reduces to the search for $u\in C$ satisfying

$$(f(u),u-v)\leqslant 0,\quad \forall v\in C. \tag{1.7}$$

This problem will be designated in the sequel as VI(f,C,H).

It is well known that every variational inequality of the form (1.5) can be associated with a variational inequality of the form VI(F,C,V). This is done by setting $C=$

$\{[x, \lambda] \in H \times \mathbb{R} : \phi(x) \leqslant \lambda\}$ (called the epigraph of ϕ), $F[x, \lambda] = [f(x), 1]$ for all $[x, \lambda] \in H \times \mathbb{R}$, and $V = H \times \mathbb{R}$, equiped with the scalar product $\langle [x, \lambda], [y, \tau] \rangle = (x, y) + \lambda\tau$. Then it is not difficult to see that if u solves (1.5) then $[u, \phi(u)]$ solves VI(F, C, V), and that if $[u, \lambda]$ solves VI(F, C, V) then u solves (1.5).

The generalized Newton method for the iterative solution of problem (1.7) is given by the scheme

$$\begin{aligned}&(f'(u_m)u_{m+1}, u_{m+1} - v) \leqslant (f'(u_m)u_m - f(u_m), u_{m+1} - v),\\ &\forall v \in C, \ m = 0, 1, \ldots,\end{aligned} \tag{1.8}$$

Early studies on the convergence of the scheme were carried out by Eaves [12], Robinson [32], Josephy [15,16], and by Pang and Chan [28]. These works were reviewed recently by Harker and Pang [13].

A version of this method that is directly applicable to the variational inequality (1.5) is the scheme

$$f'(u_m)u_{m+1} + \partial\phi(u_{m+1}) \ni f'(u_m)u_m - f(u_m), \quad m = 0, 1, \ldots, \tag{1.9}$$

which was studied in [37,39] without being aware of the previous work done on (1.8). When ϕ is of the form (1.6) this scheme reduces to (1.8).

If we use (1.8) to solve the epigraph formulation VI(F, C, V) of (1.5), what we obtain is precisely the epigraph formulation of the numerical scheme (1.9). However, a drawback of this approach to the numerical solution of (1.5) is the increase in the number of variables which could have an adverse effect on the numerical scheme. Another drawback arises from the fact that the derivative of the function F occurring in the epigraph formulation of (1.5) is given by $F'[x, \lambda][y, \tau] = [f'(x)y, 0]$ for all $[x, \lambda], [y, \tau] \in V$, and fails to satisfy the coercivity (positive definiteness) condition frequently required (cf. [28,37]) for the convergence of the scheme (1.8). However, some results on the convergence of (1.9) can be obtained – via its epigraph formulation – from known results on (1.8). The results of Josephy [15,16] and Robinson [31,33] are based on Robinson's notion of regular solution [32] instead of coercivity and can be applied to (1.9) in this manner. Eaves [12] follows a different approach and proves a local convergence result for (1.8) which when applied to VI(F, C, V), imposes a coercivity condition on $f'(x)$ instead of $F'[x, \lambda]$. This result could be used to obtain Theorem 1 of Uko [39].

The numerical scheme (1.9) is a special case of the generalized Newton scheme

$$f'(u_m)u_{m+1} + g(u_{m+1}) \ni f'(u_m)u_m - f(u_m), \quad m = 0, 1, \ldots, \tag{1.10}$$

which can be used for the iterative solution of the generalized equation (1.1). A related scheme is the generalized chord method which is given by

$$f'(u_0)u_{m+1} + g(u_{m+1}) \ni f'(u_0)u_m - f(u_m), \quad m = 0, 1, \ldots. \tag{1.11}$$

Unlike (1.10), the generalized chord scheme uses only one derivative evaluation for all iterations and is usually employed in situations in which the computation of derivatives is costly.

The scheme (1.8) is usually referred to (cf. [12,13,15,28]) as "Newton's method". In fact, if $C = H$, it reduces to Newton's method for the equation $f(u) = 0$. However, when g is single-valued, then (1.10) differs from Newton's method and could be used to solve the equation

$$f(u) + g(u) = 0 \tag{1.12}$$

in situations in which Newton's method cannot be used, due to the lack of differentiability of g. For this reason, we refer to (1.10) – and, by implication, to (1.8) – as the generalized Newton method.

The main advantage of the generalized Newton scheme is the fact that it is applicable to problems involving nondifferentiable and possibly multivalued functions, while retaining the quadratic termination property which is usually associated with Newton's method. In fact, when both methods are applicable, the generalized method may require less outer iterations than Newton's method to achieve a specified accuracy. To illustrate this, let $H = \mathbb{R}$, let f be a twice continuously differentiable function, and let g be single-valued and differentiable at the solution u of (1.12). Then the asymptotic error constant for the generalized Newton method for problem (1.1) is given [39] by

$$e^* \equiv \lim_{m \to \infty} \frac{|u_{m+1} - u|}{|u_m - u|^2} = \frac{|f''(u)|}{|f'(u) + g'(u)|}.$$

The asymptotic error constant for the classical Newton method is

$$e = \frac{|f''(u) + g''(u)|}{|f'(u) + g'(u)|}.$$

Therefore, if $g''(u)f''(u) > 0$, then $e^* < e$, which means that the generalized method will converge in fewer iteration steps than Newton's method.

The major drawback of the generalized Newton method is the fact that its' iterates – which are defined by (1.10) – actually need to be computed by means of some further inner-iterative method.

Another approach to the solution of generalized (nonsmooth) equations – especially those of the form (1.4) – uses generalized derivatives that are applicable to functions that are not differentiable in the traditional sense of Fréchet and Gâteaux. This idea (which is not pursued in the present paper) has lead to the development of Newton-like methods that are applicable to nonsmooth equations. Such methods include the B-derivative-based Newton method proposed by Pang [27], the generalized Jacobian-based Newton method proposed by Qi [30], the Gauss–Newton method proposed by Dennis and Schnabel [10], and the nonsmooth quasi-Newton methods studied by Ip and Kyparisis [14] and Kojima and Shindo [19]. The survey paper [29] by Pang and Qi introduces and motivates these schemes and also contains proofs of the superlinear convergence of each of them.

Many results on the convergence of the generalized Newton method and the generalized method of chords can be found in [15,16,31,33,12,13,28,37,39]. In the present paper we obtain further results not contained in the previous papers. We are particularly interested in results which extend to generalized equations the Kantorovich technique [17,25,35] for the solution – via Newton's method and the method of chords – of equations involving differentiable functions.

The main results of the present paper are contained in Section 2. The first results in this section extend previous results on the generalized Newton method obtained by Eaves [12], Pang and Chan [28], and Uko [37,39]. The last two theorems of the section are of the Kantorovich type and extend the Kantorovich technique to problem (1.1) approximated with the generalized Newton method and the generalized method of chords. Similar results have been obtained by Robinson [33, Theorem 5.1]. However, his hypotheses differ from ours, and his results are not applicable to the general problem (1.1).

In Section 3 we study methods for the numerical computation of the generalized Newton iterates. The methods studied in this section are generalizations of classical methods (Richardson's method, Jacobi's method, and the Gauss–Seidel method – cf. [9,21,24]) for solving linear equations and linear complementarity problems.

2. The main results

Let H be a Hilbert space with scalar product $(\cdot, \cdot)$ and norm $\|\cdot\|$. Let D_0 be the interior of a closed convex subset D of H, and for any $u_0 \in D_0$ and $r > 0$, let $B[u_0, r]$ designate the set $\{x \in H : \| x - u_0 \| \leqslant r\}$ while $B(u_0, r)$ designates the interior of $B[u_0, r]$.

All through this section we will assume that $f: D \mapsto H$ is a continuous function that is Fréchet differentiable at each point of D_0 and satisfies the condition

$$\| f'(x) - f'(y) \| \leqslant M \| x - y \| \quad \forall x, y \in D_0. \tag{2.1}$$

It is well known (cf. [25, p. 70]) that (2.1) implies that for all $z \in D$ and $y \in D_0$, we have

$$\| f(x) - f(y) - f'(y)(x - y) \| \leqslant \tfrac{1}{2} M \| x - y \|^2, \tag{2.2}$$

$$\| f(x) - f(z) - f'(y)(x - z) \| \leqslant M \max\{ \| x - y \|, \| z - x \| \} \| x - z \|. \tag{2.3}$$

We will make repeated use of these inequalities in the sequel.

We will also assume that g is a (multivalued) maximal monotone function from H to H. This means that g is nonempty subset of $H \times H$ which is monotone in the sense that there exists $\alpha \geqslant 0$ (its monotonicity modulus) such that

$$[x_1, y_1] \in g \text{ and } [x_2, y_2] \in g \Rightarrow (y_2 - y_1, x_2 - x_1) \geqslant \alpha \| x_1 - x_2 \|^2. \tag{2.4}$$

and which is not contained in any larger monotone subset of $H \times H$. It is well known (cf. [5,23]) that g is closed in the sense that

$$[x_m, y_m] \in g, \quad \lim_{m \to \infty} y_m = y \text{ and } \lim_{m \to \infty} x_m = x \Rightarrow [x, y] \in g, \tag{2.5}$$

and that, given any fixed positive real number μ, the resolvent operator $(1+\mu g)^{-1}$ exists as a single-valued function and satisfies

$$\|(1+\mu g)^{-1}(x)-(1+\mu g)^{-1}(y)\| \leqslant \frac{\|x-y\|}{1+\mu\alpha}, \quad \forall x, y \in H. \tag{2.6}$$

If $\alpha > 0$, g will be said to be strongly maximal monotone. A well known example [4,23] of a maximal monotone operator is the subgradient $\partial\phi$ of a proper lower semicontinuous convex function ϕ mapping H into $(-\infty, \infty]$.

One of the most important results in the theory of equations is the Kantorovich theorem [17, Ch. XVII, Theorem 6] which we state as follows.

Theorem 2.1. *Suppose that* (2.1) *holds. If* $f'(u_0)^{-1}$ *exists and is such that* $\|f'(u_0)^{-1}\| \leqslant b$, $\|f'(u_0)^{-1}f(u_0)\| \leqslant a$, $h \equiv Mab \leqslant \frac{1}{2}$, *and* $B[u_0, r_-] \subset D_0$ *where* $r_{\pm} = (a/h)(1 \pm \sqrt{1-2h})$, *then*

(a) *The equation* $f(u)=0$ *possesses a unique solution* u *in* $B(u_0, r_+) \cap D$.

(b) *The sequence defined inductively by Newton's method*

$$u_{m+1} = u_m - f'(u_m)^{-1} f(u_m), \quad m = 0, 1, \ldots,$$

converges to u *at the rate* $\|u-u_m\| \leqslant (a/h)(2h)^{2^m}2^{-m}$.

(c) *If* $h < \frac{1}{2}$, *then the sequence defined inductively by the chord method*

$$u_{m+1} = u_m - f'(u_0)^{-1} f(u_m), \quad m = 0, 1, \ldots,$$

converges to u *at the rate* $\|u-u_m\| \leqslant (a/h)[1-\sqrt{1-2h}]^{m+1}$.

Our main aim in this section is to extend this result to the generalized equation (1.1), approximated with the iterative schemes (1.10) and (1.11). However, before giving this extension we first give some general convergence results for these approximating iterative schemes. The first result gives sufficient conditions for each of these approximating generalized equations to have a unique solution.

Lemma 2.2. *Let* g *be a maximal monotone operator satisfying* (2.4) *and let* A *be a bounded linear operator mapping* H *into* H. *If there exists* $c \in \mathbb{R}$ *such that* $c > -\alpha$ *and*

$$(Ax, x) \geqslant c\|x\|^2, \quad \forall x \in H. \tag{2.7}$$

Then, for any $b \in H$, *there exists a unique* $z \in H$ *satisfying the generalized equation*

$$Az + g(z) \ni b. \tag{2.8}$$

Proof. It is easy to see that problem (2.8) is equivalent to the search for the fixed points of the operator $T_\mu x = (1+\mu g)^{-1}(x+\mu b-\mu Ax)$ for any fixed positive parameter μ.

Using (2.6) and (2.7), we see that

$$\begin{aligned}\|T_\mu x - T_\mu y\| &= \|(1+\mu g)^{-1}(x+\mu b-\mu Ax) \\ &\quad -(1+\mu g)^{-1}(y+\mu b-\mu Ay)\| \\ &\leqslant \frac{1}{1+\mu\alpha}\|x-y-\mu Ax+\mu Ay\| \\ &\leqslant \frac{\sqrt{1-2\mu c+\mu^2\|A\|^2}}{1+\mu\alpha}\|x-y\|, \quad \forall x, y \in D.\end{aligned}$$

If $\| A \| > \alpha$, we choose $0 < \mu < 2(c+\alpha)/(\| A \|^2 - \alpha^2)$; otherwise, we let μ be any arbitrary positive real number. Then we have $\sqrt{1 - 2\mu c + \mu^2 \| A \|^2} < 1 + \mu\alpha$, which shows that T_μ is a strict contraction mapping from the convex set D to itself. The existence of a unique fixed point for T_μ therefore follows from Banach's contraction mapping principle. □

Remark 2.3. If g is the subgradient of a convex function, and if (2.4) holds with $\alpha = 0$, then Lemma 2.2 becomes the classical Lions–Stampacchia theorem [20, Theorem 2.1] on the solvability of variational inequalities.

Remark 2.4. If (2.7) holds, we will say that A is weakly coercive. This concept of coercivity is weaker than the usual one (cf. [20]) since we do not require that c be positive.

Lemma 2.2 assures us that if

$$\exists c > -\alpha \quad \text{such that } (f'(z)x,\ x) \geqslant c \| x \|^2, \quad \forall x \in H,\ \forall z \in D_0, \tag{2.9}$$

then all the generalized Newton iterates u_{m+1} in (1.10) exist. The next theorem contains a convergence result that uses this hypothesis.

Theorem 2.5. *Let g be a maximal monotone operator satisfying* (2.4). *Suppose that* (2.1) *and* (2.9) *hold and that the generalized equation* (1.1) *has a unique solution u in D_0. If the initial vector in* (1.10) *satisfies the condition $d \equiv (M/2(c+\alpha)) \| u - u_0 \| < 1$, then the generalized Newton iterates u_m defined inductively by* (1.10) *converge to u at the rate*

$$\| u - u_m \| \leqslant \frac{2(c+\alpha)}{M} d^{2^m}.$$

Proof. The existence of the solutions to (1.10) follows from (2.9) and Lemma 2.2.

For $m = 0, 1, \ldots,$ if we use (2.4), (1.1) and (1.10), we obtain

$$\alpha \| u_{m+1} - u \|^2 \leqslant (f(u) - f(u_m) - f'(u_m)(u_{m+1} - u_m),\ u_{m+1} - u).$$

Rewriting this in the form

$$c \| u_{m+1} - u \|^2 + (f'(u_m)(u_{m+1} - u),\ u_{m+1} - u)$$
$$\leqslant (f(u) - f(u_m) - f'(u_m)(u - u_m),\ u_{m+1} - u)$$

and making use of (2.9) and (2.2), we obtain $\| u_{m+1} - u \| \leqslant (M/2(c+\alpha)) \| u_m - u \|^2$. An induction argument now shows that $\| u_m - u \| \leqslant (2(c+\alpha)/M)(M/2(c+\alpha) \| u - u_0 \|)^{2^m} = (2(c+\alpha)/M) d^{2^m}$, which completes the proof of the theorem. □

Remark 2.6. The conclusions of the theorem hold true — with the same proof — if instead of (2.1) we employ the weaker condition $\| f'(u) - f'(x) \| \leqslant M \| u - x \|$, $\forall x \in D_0$.

Remark 2.7. In the special case in which g is the subgradient of a proper lower-semicontinuous convex function and (2.4) holds with $\alpha = 0$, then Theorem 2.5 reduces to Corollary 26 of Pang and Chan [28]. A similar results was obtained by Josephy [15] using Robinson's notion [32] of regular solution instead of the coercivity condition (2.9).

Theorem 2.5 is of limited practical utility because its hypothesis and convergence rate are expressed in terms of the exact solution of (1.1) which is not usually known apriori. The next result is more useful in this regard since it provides a convergence rate which we could compute apriori if we knew the value of the constants M, α and c in (2.1), (2.4) and (2.9).

Theorem 2.8. *Let g be a maximal monotone operator satisfying* (2.4), *and suppose that* (2.1) *and* (2.9) *hold. Let $\{u_m : m = 0, 1, \ldots\}$ be the generalized Newton iterates from* (1.10), *and suppose that $d \equiv (M/2(c+\alpha)) \| u_1 - u_0 \| < 1$ and that $B[u_0, r] \subset D_0$, where $r = (2(c+\alpha)/M)\sum_{k=0}^{\infty} d^{2^k}$. Then all the u_m lie in $B[u_0, r]$ and converge to a solution u of* (1.1) *at the rate*

$$\| u - u_m \| \leqslant \frac{2(c+\alpha)}{M} \sum_{k=m}^{\infty} d^{2^k}.$$

Proof. The existence of the solutions u_{m+1} to (1.10) follows from (2.9) and Lemma 2.2.

For $m = 0, 1, \ldots,$ if we use (1.10), the corresponding generalized equation for $m-1$, and (2.4), we obtain

$$\begin{aligned}&\alpha \| u_{m+1} - u_m \|^2 + (f'(u_m)(u_{m+1} - u_m), u_{m+1} - u_m) \\ &\quad\leqslant (f(u_m) - f(u_{m-1}) - f'(u_{m-1})(u_m - u_{m-1}), u_m - u_{m+1}).\end{aligned}$$

If we now make use of (2.9) and (2.2), we obtain $\| u_{m+1} - u_m \| \leqslant (M/2(c+\alpha)) \| u_m - u_{m-1} \|^2$. An induction argument now shows that $\| u_{m+1} - u_m \| \leqslant (2(c+\alpha)/M)((M/2(c+\alpha)) \| u_1 - u_0 \|)^{2^m} = (2(c+\alpha)/M) d^{2^m}$. Using this fact we see that $\| u_{m+1} - u_0 \| \leqslant \sum_{k=0}^{m} \| u_{k+1} - u_k \| \leqslant r$, which implies that $u_{m+1} \in B[u_0, r]$, $\forall m$.

Finally, since $\| u_{m+p} - u_m \| \leqslant \sum_{k=m}^{m+p-1} \| u_{k+1} - u_k \| \leqslant (2(c+\alpha)/M) \sum_{k=m}^{m+p-1} d^{2^k}$, it follows that u_m is a Cauchy sequence, converging to some $u \in D$. The fact that u solves (1.1) follows from (1.10) and (2.4), and the convergence rate is obtained by letting p tend to infinity. □

Remark 2.9. Let $g = \partial\phi$, where ϕ is a proper lower-semicontinuous convex function ϕ, and suppose that (2.4) holds with $\alpha = 0$. If we take f as the gradient $\nabla\psi$ of a real-valued function ψ defined on D, and replace the condition $d < 1$ in Theorem 2.8 with the more restrictive condition $d < \frac{1}{2}$, we recover Theorem 2 of [39].

The coercivity condition (2.9) employed in the last two theorems implies that $f'(z)^{-1}$ exists for all $z \in D_0$ and is rather strong. Since this condition will not be satisfied in most problems occurring in applications we are led to consider the less restrictive hypothesis:

$$\exists c_0 > -\alpha \quad \text{such that } (f'(u_0)x, x) \geqslant c_0 \| x \|^2, \quad \forall x \in H. \tag{2.10}$$

This turns out to be the natural hypothesis for the extension of the Kantorovich technique to the solution of (1.1). For simplicity we first give the relevant Kantorovich-type result for the generalized chord method before giving the analogous result for the generalized Newton method.

Theorem 2.10. *Let g be a maximal monotone operator satisfying* (2.4), *and suppose that* (2.1) *and* (2.10) *hold. Let* $u_0 \in D_0$ *and suppose that there exists* $v_0 \in H$ *such that* $g(u_0) \ni v_0$ *and* $\| f(u_0) + v_0 \| \leqslant b_0$ *for some* $b_0 > 0$. *Let* $a_0 = b_0/(c_0 + \alpha)$ *and* $h_0 = Ma_0/(c_0 + \alpha)$ *and suppose that*

$$h_0 < \tfrac{1}{2}. \tag{2.11}$$

If $r_0 = 2a_0/(1 + \sqrt{1 - 2h_0})$, $R_0 = 2a_0/(1 - \sqrt{1 - 2h_0})$, *and* $B[u_0, r_0] \subset D_0$, *then there exists a unique solution u of* (1.1) *in* $B(u_0, R_0) \cap D$. *Moreover, the generalized chord iterates* u_m (*with initial vector* u_0) *converge to u at the rate*

$$\| u - u_m \| \leqslant \frac{a_0}{h_0}\left(1 - \sqrt{1 - 2h_0}\right)^{m+1}. \tag{2.12}$$

Proof. For any $x \in B[u_0, r_0]$, Lemma 2.2 ensures that we can define $w(x)$ uniquely by means of the generalized equation

$$f'(u_0)w(x) + g((w(x)) \ni f'(u_0)x - f(x).$$

Using (2.4) and the definitions of $w(x)$ and v_0, we obtain

$$\alpha \| w(x) - u_0 \|^2 \leqslant (v_0 + f(x) - f'(u_0)(x - w(x)), u_0 - w(x)).$$

Rewriting this in the form

$$\alpha \| w(x) - u_0 \|^2 + (f'(u_0)(w(x) - u_0), w(x) - u_0)$$
$$\leqslant (f'(u_0)(x - u_0) - f(x) - v_0, w(x) - u_0)$$

and making use of (2.2), (2.11) and the hypothesis on v_0, we obtain

$$\begin{aligned}(c_0 + \alpha) \| w(x) - u_0 \| &\leqslant \| v_0 + f(x) - f'(u_0)(x - u_0) \| \\ &= \| v_0 + f(u_0) + f(x) - f(u_0) - f'(u_0)(x - u_0) \| \\ &\leqslant \| f(u_0) + v_0 \| + \| f(x) - f(u_0) + f'(u_0)(x - u_0) \| \\ &\leqslant b_0 + \tfrac{1}{2}M \| x - u_0 \|^2.\end{aligned}$$

We conclude that

$$(c_0 + \alpha) \| w(x) - u_0 \| \leqslant b_0 + \tfrac{1}{2}M \| x - u_0 \|^2, \tag{2.13}$$

which implies that $\| w(x) - u_0 \| \leqslant a_0 + Mr_0^2/2(c_0 + \alpha) = r_0$ for all $x \in B[u_0, r_0]$. Therefore w maps the set $B[u_0, r_0]$ into itself.

Also, for any $x, y \in B[u_0, r]$, it follows from (2.4) and the definition of w that

$$\alpha \| w(x) - w(y) \|^2 \leqslant (f'(u_0)(w(x) - w(y)) + f(x) - f(y) - f'(u_0)(x-y), w(y) - w(x)).$$

We now rewrite this in the form

$$\alpha \| w(x) - w(y) \|^2 + (f'(u_0)(w(x) - w(y)), w(x) - w(y)) \leqslant (f'(u_0)(x-y) - f(x) + f(y), w(x) - w(y))$$

and make use of (2.3) and (2.11). We obtain

$$\begin{aligned} \| w(x) - w(y) \| &\leqslant \frac{1}{c_0 + \alpha} \| f(x) - f(y) - f'(u_0)(x-y) \| \\ &\leqslant \frac{M}{c_0 + \alpha} \| x - y \| \max\{ \| x - u_0 \|, \| y - u_0 \| \} \\ &\leqslant \frac{M r_0}{c_0 + \alpha} \| x - y \| = \left(1 - \sqrt{1 - 2h_0}\right) \| x - y \|. \end{aligned}$$

Letting $q = 1 - \sqrt{1 - 2h_0}$, we have $\| w(x) - w(y) \| \leqslant q \| x - y \|$. It therefore follows that w is a strict contraction mapping from the set $B[u_0, r_0]$ to itself. Since (1.11) is the successive approximation scheme $u_{m+1} = w(u_m)$, it follows from Banach's contraction-mapping theorem that there exists a unique $u \in B[u_0, r_0]$ satisfying $w(u) = u$, which is equivalent to (1.1). The error estimate (2.12) is obtained by observing that for $m = 0, 1, \ldots,$ we have

$$\| u_m - u \| = \| w(u_{m-1}) - w(u) \| \leqslant q^m \| u_0 - u \| \leqslant q^m r_0 = \frac{a_0}{h_0} q^{m+1}.$$

If v is another solution to (1.1), then v is another fixed point for w. On using (2.13), we obtain the inequality

$$(c_0 + \alpha) \| v - u_0 \| \leqslant b_0 + \tfrac{1}{2} M \| v - u_0 \|^2.$$

By solving this inequality, we see that either $\| v - u_0 \| \geqslant R_0$ or $\| v - u_0 \| \leqslant r_0$. Thus, if $v \in B(u_0, R_0)$, then we must have $v \in B[u_0, r_0]$ which implies that $v = u$ by the uniqueness assertion of the previous paragraph. The proof of the theorem is complete. □

The next result proves the second part of our Kantorovich-type result for the generalized equation (1.1), approximated with the generalized Newton scheme (1.10). Robinson [33, Theorem 5.1] has proved a similar result for problem (1.7), approximated with the generalized Newton scheme (1.8). He used a concept of regularity developed in his earlier work [32] in place of a coercivity hypothesis. However, his result is not applicable to the general generalized equation (1.1), and our proof is much simpler than his.

Theorem 2.11. *Let all the hypothesis of Theorem* 2.10 *hold, with the sole exception that* (2.11) *is replaced with the weaker assumption*

$$h_0 \leqslant \tfrac{1}{2}. \tag{2.14}$$

Then there exists a solution u of (1.1) *in $B[u_0, r_0]$ to which the generalized Newton iterates u_m converge at the rate*

$$\| u_m - u \| \leqslant \frac{a_0}{h_0}(2h_0)^{2^m} 2^{-m}. \tag{2.15}$$

Furthermore, u is the only solution of (1.1) *in $B(u_0, R_0) \cap D$.*

Proof. We prove by induction that for $m = 0, 1, \ldots,$ we have

$$u_m \in B[u_0, r_0], \tag{2.16}$$

$$\exists v_m \in H \quad \text{such that } v_m \in g(u_m), \tag{2.17}$$

$$\exists b_m > 0 \quad \text{such that } \| f(u_m) + v_m \| \leqslant b_m, \tag{2.18}$$

$$\exists c_m > -\alpha \quad \text{such that } (f'(u_m)x, x) \geqslant c_m \| x \|^2, \quad \forall x \in H, \tag{2.19}$$

$$h_m \equiv \frac{Mb_m}{(c_m + \alpha)^2} \leqslant \tfrac{1}{2}. \tag{2.20}$$

The hypothesis of the theorem ensure that the induction hypothesis is true if $m = 0$. We assume that $m \geqslant 0$ and that the induction hypothesis holds for m. Then it follows from (2.19) and Lemma 2.2 that there exists a unique $u_{m+1} \in H$ satisfying (1.10).

Using (2.17), (1.10) and (2.4), we obtain

$$\alpha \| u_{m+1} - u_m \|^2 + (v_m + f(u_m) - f'(u_m)(u_m - u_{m+1}), u_{m+1} - u_m) \leqslant 0.$$

Rewriting this in the form

$$\alpha \| u_{m+1} - u_m \|^2 + (f'(u_m)(u_{m+1} - u_m), u_{m+1} - u_m)$$
$$\leqslant (-f(u_m) - v_m, u_{m+1} - u_m)$$

and making use of (2.18) and (2.19), we see that

$$\| u_{m+1} - u_m \| \leqslant \frac{b_m}{c_m + \alpha} \equiv a_m. \tag{2.21}$$

If we set $r_k = 2a_k/(1 + \sqrt{1 - 2h_k})$, for $k = 0, 1, \ldots, m+1$ then it is easy to check that $r_k - r_{k+1} = a_k$ for $k = 0, \ldots, m$. Therefore, $\| u_{m+1} - u_0 \| \leqslant \sum_{k=0}^{m} \| u_{k+1} - u_k \| \leqslant \sum_{k=0}^{m} a_k \leqslant \sum_{k=0}^{m} (r_k - r_{k+1}) = r_0 - r_{m+1} < r_0$ which implies that $u_{m+1} \in B[u_0, r_0] \subset D_0$, and hence that (2.16) holds when m is replaced by $m + 1$.

Next, using (2.1) and (2.21), we obtain $\| f'(u_{m+1}) - f'(u_m) \| \leqslant M \| u_{m+1} - u_m \| \leqslant Ma_m$. Therefore,

$$(f'(u_m)x - f'(u_{m+1})x, x) \leqslant \| f'(u_m) - f'(u_{m+1}) \| \, \| x \|^2 \leqslant Ma_m \| x \|^2,$$

for all $x \in H$, which implies, because of (2.19), that $(f'(u_{m+1})x, x) \geqslant c_{m+1} \| x \|^2$ for all $x \in H$, where $c_{m+1} = c_m - Ma_m$. The inequality $c_{m+1} > -\alpha$ follows from (2.20) and shows that (2.19) holds when m is replaced by $m+1$.

Let $v_{m+1} = -f(u_m) - f'(u_m)(u_{m+1} - u_m)$. Then (2.8) shows that (2.17) holds when m is replaced with m is replaced with $m+1$, and we have

$$\| f(u_{m+1}) + v_{m+1} \| = \| f(u_{m+1}) - f(u_m) - f'(u_m)(u_{m+1} - u_m) \|$$
$$\leqslant \tfrac{1}{2} M \| u_{m+1} - u_m \|^2 \leqslant \tfrac{1}{2} M a_m^2 \equiv b_{m+1}.$$

Finally, (2.20) shows that $h_{m+1} = h_m^2 / 2(1-h_m)^2 \leqslant \frac{1}{2}$, which proves that (2.20) holds when m is replaced with $m+1$. That completes the proof of the fact that (2.16)–(2.20) hold for all positive integers m.

Because of (2.20), we have $h_m \leqslant \frac{1}{2} h_{m-1}^2 / (1-h_{m-1})^2 \leqslant 2h_{m-1}^2 \leqslant \cdots \leqslant \frac{1}{2}[2h_0]^{2^m}$. This implies that

$$a_m = \frac{a_{m-1} h_{m-1}}{2(1-h_{m-1})} \leqslant a_{m-1} h_{m-1} \leqslant h_{m-1} h_{m-2} \ldots h_0 a_0$$
$$\leqslant 2^{-m} [2h_0]^{2^{m-1} + 2^{m-2} + \cdots + 1} a_0 = 2^{-m} [2h_0]^{2^m - 1} a_0.$$

Consequently, we have

$$\| u_{m+p} - u_m \| \leqslant \sum_{k=m}^{m+p-1} \| u_{k+1} - u_k \| \leqslant \sum_{k=m}^{m+p-1} a_k$$
$$\leqslant 2^{-m} [2h_0]^{2^m - 1} a_0 \sum_{k=0}^{p-1} 2^{-k} [2h_0]^{2^k - 1}.$$

It follows that u_m is a Cauchy sequence and so converges to some $u \in D$ satisfying

$$\| u - u_m \| \leqslant 2^{-m} [2h_0]^{2^m - 1} a_0 \sum_{k=0}^{\infty} 2^{-k} [2h_0]^{2^k - 1} \leqslant 2^{-m} [2h_0]^{2^m} \frac{a_0}{h_0}.$$

Also, since $u_{m+1} \to u$, $f'(u_m)(u_m - u_{m+1}) - f(u_m) \to -f(u)$, and $g(u_{m+1}) \ni f'(u_m)(u_m - u_{m+1}) - f(u_m)$, it follows from (2.5) that $g(u) \ni -f(u)$, so that u solves (1.1).

If $h_0 < \frac{1}{2}$ uniqueness in $B(u_0, R_0)$ follows from Theorem 2.10. Let $h_0 = \frac{1}{2}$ and let v be another solution of (1.1) belonging to $B(u_0, R_0) \cap D$. Then there exists $0 \leqslant \theta < 1$ such that $\| v - u_0 \| = \theta R_0 = 2a_0 \theta$. For $m = 0, 1, \ldots,$ it follows from (1.1), (1.10) and (2.4) that

$$\alpha \| u_{m+1} - v \|^2 + \left(f'(u_m)(u_{m+1} - v), u_{m+1} - v \right)$$
$$\leqslant \left(f(v) - f(u_m) - f'(u_m)(v - u_m), u_{m+1} - v \right).$$

Therefore, using (2.19) and (2.2), we obtain $\| u_{m+1} - v \| \leqslant (M/2(c_m + \alpha)) \| u_m - v \|^2$. Since $\frac{1}{2} \leqslant 1 - h_m$ and $c_{m+1} + \alpha = (1 - h_m)(c_m + \alpha)$, an induction argument shows that $\| u_m - v \| \leqslant ((c_m + \alpha)/M) \theta^{2^m} \leqslant ((c_0 + \alpha)/M) \theta^{2^m}$ and implies that $u = \lim_{m \to \infty} u_m = v$. That completes the proof of the theorem. □

3. Inner iteration methods

In some (rare, one-dimensional) situations (cf. [37]) it is possible to solve the generalized equations (1.10) and (1.11) in closed form. For instance, let $H=\mathbb{R}$ and let g_1, g_2, h_1, h_2 be real constants satisfying the inequalities $g_1<0<g_2$ and $h_1<h_2$. Let g be the subgradient of the convex function

$$\phi(t)=\begin{cases} g_1(t-h_1), & t<h_1,\\ 0, & h_1\leqslant t\leqslant h_2,\\ g_2(t-h_2), & t>h_2.\end{cases} \tag{3.1}$$

Then it is easy to verify that the generalized Newton iterates (1.10) are given by

$$u_{m+1}=\begin{cases} v_m-g_1/f'(u_m), & \text{if } v_m<h_1+g_1/f'(u_m),\\ v_m+(h_1-v_m)^+-(v_m-h_2)^+, & \text{if } h_1+g_1/f'(u_m)\leqslant v_m\\ & \quad\leqslant h_2+g_2/f'(u_m),\\ v_m-g_2/f'(u_m), & \text{if } v_m>h_2+g_2/f'(u_m),\end{cases}$$

where $v_m=u_m-f(u_m)/f'(u_m)$ and $t^+=\max\{t,0\}$ for all $t\in\mathbb{R}$. An analogous expression can be easily be obtained for the generalized chord scheme.

In general, the generalized Newton iterates and generalized chord iterates satisfy a generalized equation of the form (2.8). This generalized equation would not usually have a closed-form solution and would have to be solved using some inner iterative numerical method.

One possible numerical scheme for the solution of problem (2.8) is the generalized Richardson scheme

$$z^{(k+1)}=(1+\mu g)^{-1}(z^{(k)}+\mu b-\mu Az^{(k)}),\quad k=0,1,\ldots. \tag{3.2}$$

When $g=0$ the scheme becomes the well known method of Richardson [24] for the solution of linear equations, and when g is the subgradient of the proper convex function ϕ, it reduces to the numerical scheme referred to by Cohen [7] as the "auxiliary problem principle" for the solution of (2.7).

The scheme (3.2) is the successive approximation scheme $z^{(k+1)}=T_\mu z^{(k)}$ corresponding to the operator T_μ employed in the proof of Lemma 2.2. It follows that if the conditions of Lemma 2.2 hold, and if μ is small enough to ensure that T_μ is a contraction mapping, then the generalized Richardson scheme converges globally — that is, for every choice of the initial vector $z^{(0)}$ — to the unique solution of (2.8).

For separable problems, the generalized Newton iterates are obtained by solving generalized equations of the form

$$\sum_{j=1}^{n} a_{ij}z_j+g_i(z_i)\ni b_i,\quad i=1,\ldots,n, \tag{3.3}$$

where $A=(a_{ij})$ is an $n\times n$ matrix and $(b_1,\ldots,b_n)\in\mathbb{R}^n$. On setting $P_i=(1+(1/a_{ii})g_i)^{-1}$, we can write each of these problems in the form

$$z_i=P_i\left[\frac{1}{a_{ii}}\left(b_i-\sum_{j\neq i}a_{ij}z_j\right)\right],\quad i=1,\ldots,n.$$

This formulation suggests the following iterative method

$$z_i^{(k+1)} = P_i\left[\frac{1}{a_{ii}}\left(b_i - \sum_{j \neq i} a_{ij} z_j^{(k)}\right)\right], \quad i = 1, \ldots, n, \tag{3.4}$$

which we will refer to as the generalized Jacobi scheme. Another iterative approach is to update one component at a time, in which case we are led to the numerical scheme

$$z_i^{(k+1)} = P_i\left[\frac{1}{a_{ii}}\left(b_i - \sum_{j<i} a_{ij} z_j^{(k+1)} - \sum_{j>i} a_{ij} z_j^{(k)}\right)\right], \quad i = 1, \ldots, n, \tag{3.5}$$

which we will refer to as the generalized Gauss–Seidel scheme.

If we choose $g_1 = \cdots = g_n = 0$ in these schemes, we recover the Jacobi and the Gauss–Seidel methods for the solution of linear equations, and if the g_i are given by (1.3) we recover the projected Jacobi and Gauss–Seidel methods proposed by Cryer [9] and Mangasarian [21] for the solution of linear complementarity problems. If the g_i are subgradients of general convex functions we recover the numerical schemes proposed by Uko [36] for separable variational inequalities.

The theorem below gives sufficient conditions for the iterates generated from the generalized Jacobi method and the generalized Gauss–Seidel method to converge globally in the maximum norm $\| x \|_\infty = \max_{1 \leqslant i \leqslant n} | x_i |$.

Theorem 3.1. *Let $g_1, \ldots, g_n$ be maximal monotone operators in $\mathbb{R} \times \mathbb{R}$ with monotonicity moduli $\alpha_1, \ldots, \alpha_n$ (respectively). If*

$$a_{ii} > 0 \quad \textit{and} \quad \sum_{j \neq i} | a_{ij} | < a_{ii} + \alpha_i, \qquad i = 1, \ldots, n, \tag{3.6}$$

then (3.3) has a unique solution z, for any $b \in \mathbb{R}^n$. Moreover, the iterates obtained from the generalized Jacobi scheme and the generalized Gauss–Seidel scheme converge globally to z in the maximum norm.

Proof. It is easy to see that the solution $z = (z_1, \ldots, z_n)$ of (3.3) is a fixed point of the operator defined, for $x \in \mathbb{R}^n$, by $Jx = ((Jx)_1, \ldots, (Jx)_n)$ where $(Jx)_i = P_i[(1/a_{ii})(b_i - \sum_{j \neq i} a_{ij} x_j)]$, $i = 1, \ldots, n$.

On using (2.6) and the definition of the α, we obtain the inequalities

$$\| (Jx)_i - (Jy)_i \| \leqslant \left(\frac{\sum_{j \neq i} | a_{ij} |}{a_{ii} + \alpha_i} \right) \| x - y \|_\infty, \quad \forall x, y \in \mathbb{R}^n, \qquad i = 1, \ldots, n.$$

This implies that

$$\| Jx - Jy \|_\infty \leqslant \max_{1 \leqslant i \leqslant n} \left(\frac{\sum_{j = i} | a_{ij} |}{a_{ii} + \alpha_i} \right) \| x - y \|_\infty, \quad \forall x, y \in \mathbb{R}^n.$$

Therefore, (3.6) shows that J is a contraction mapping from H into H. Since the generalized Jacobi scheme is the successive approximation scheme associated with J, its convergence follows from Banach's contraction mapping principle.

The solution $z = (z_1, \dots, z_n)$ of (3.3) is also a fixed point of the operator defined, for $x \in \mathbb{R}^n$, by $Gx = ((Gx)_1, \dots, (Gx)_n)$ where

$$(Gx)_i = P_i\left[\frac{1}{a_{ii}}\left(b_i - \sum_{j<i} a_{ij}(Gx)_j - \sum_{j>i} a_{ij}x_j\right)\right], \quad i = 1, \dots, n.$$

We prove by induction that

$$|(Gx)_m - (Gy)_m| \leqslant \left(\frac{\sum_{j \neq m} |a_{mj}|}{a_{mm} + \alpha_m}\right) \| x - y \|_\infty, \quad \forall x, y \in \mathbb{R}^n. \tag{3.7}$$

holds for all m. It is obviously true if $m = 1$, because of the definition of G and of the α_i. Suppose $k > 1$ and that it holds for $m = 1, \dots, k-1$.

Then it follows from (3.6) that $|(Gx)_m - (Gy)_m| \leqslant \| x - y \|_\infty$ for all such m, so that

$$\begin{aligned} |(Gx)_k - (Gy)_k| &\leqslant \sum_{j<k} \frac{|a_{kj}|}{a_{kk} + \alpha_k} |(Gx)_j - (Gy)_j| + \sum_{j>k} \frac{|a_{kj}|}{a_{kk} + \alpha_k} |x_j - y_j| \\ &\leqslant \sum_{j<k} \frac{|a_{kj}|}{a_{kk} + \alpha_k} \| x - y \|_\infty + \sum_{j>k} \frac{|a_{kj}|}{a_{kk} + \alpha_k} \| x - y \|_\infty \\ &\leqslant \max_{1 \leqslant i \leqslant n} \left(\frac{\sum_{j \neq i} |a_{ij}|}{a_{ii} + \alpha_i}\right) \| x - y \|_\infty, \quad \forall x, y \in \mathbb{R}^n. \end{aligned}$$

This proves, by induction, that (3.7) holds for all iterates, and hence, that

$$\| Gx - Gy \|_\infty \leqslant \max_{1 \leqslant i \leqslant n} \left(\frac{\sum j \neq i |a_{ij}|}{a_{ii} + \alpha_i}\right) \| x - y \|_\infty \quad \forall x, y \in \mathbb{R}^n.$$

Since (3.5) is the successive approximation scheme associated with G, the convergence of the generalized Gauss–Seidel scheme also follows from Banach's contraction mapping principle. □

Remark 3.2. If $\alpha_1 = \cdots = \alpha_n = 0$, then (3.6) reduces to the diagonal dominance condition:

$$\sum_{j \neq i} |a_{ij}| < a_{ii}, \quad i = 1, \dots, n. \tag{3.8}$$

It is well known (cf. [24]) that if the matrix $A = (a_{ij})$ satisfies this condition then the equation $Az = b$ has a unique solution z, and the iterates from the Jacobi and Gauss–Seidel schemes converge globally to z. However, if at least one of the α_i is strictly positive (that is, if some g_i is strongly maximal monotone), then condition (3.6) – which suffices for the convergence of (3.4) and (3.5) – is weaker than the diagonal dominance condition (3.8).

In order to make the numerical schemes (3.2), (3.4) and (3.5) fully constructive, we require algorithms for calculating the resolvents $(1 + \mu g)^{-1}(x)$, for all $\mu > 0$. Fortunately, such algorithms exist for the most common maximal monotone operators g that occur in applications.

For instance, in the nonlinear complementarity problem we come across the operator

$$g(x_1,\ldots,x_n) = g_1(x_1)\times\cdots\times g_n(x_n),$$

with the g_i defined by (1.3). In this case the resolvent is given by the simple expression

$$(1+\mu g)^{-1}(x_1,\ldots,x_n) = (x_1^+,\ldots,x_n^+),\quad \forall\mu>0.$$

Some models of heat flow through thick walls [11] employ variational inequalities which — when discretized — reduce to generalized equations involving the operator

$$g(x_1,\ldots,x_n) = \partial\phi_1(x_1)\times\cdots\times\partial\phi_n(x_n),$$

in which each ϕ_i is of the form (3.1). An explicit formula for resolvent of such an operator is given in [36, Example 2b].

If $g(x) = x\|x\|^q$, where q is a real exponent greater than or equal to one, then the resolvent can be computed with an algorithm given in [36, Example 2a]. However, if the exponent q has value 3, 4 or 5, then explicit formulae for the resolvent can be found in [38].

If $g(x_1,\ldots,x_n) = (x_1|x_1|^{q_1},\ldots,x_n|x_n|^{q_n})$ $(q_1\geqslant 1,\ldots,q_n\geqslant 1)$, then letting $g_i(x) = x_i|x_i|^{q_i}$ for $i=1,\ldots,n$, we have

$$(1+\mu g)^{-1}(x_1,\ldots,x_n) = \big((1+\mu g_1)^{-1}(x_1),\ldots,(1+\mu g_n)^{-1}(x_n)\big).$$

Therefore we can compute each component of the resolvent by using the algorithms employed in the previous paragraph.

The general variational inequality problem employs the operator $g=\partial\phi$, where ϕ is an indicator convex function of the form (1.6). In this case $(1+\mu g)^{-1}$ is independent of μ and coincides – cf. [39, Lemma 4.1] – with the orthogonal projection onto a closed convex set. Explicit expressions for such projections can be given in many cases.

Acknowledgements

The author would like to thank the referees for their patience and their suggestions which improved the contents and presentation of the paper. He is also grateful to Prof. Abdus Salam, the International Atomic Energy Agency and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

References

[1] C. Baiocchi, "Disequazioni variazionali", *Bolletino dell' Unione Matematica Italiana* 18-A (1981) 173–187.

[2] C. Baiocchi and A. Capelo, *Variational and Quasivariational Inequalities: Applications to Free Boundary Problems* (Wiley, New York, 1984).

[3] H. Brézis, "Problemes unilateraux," *Journal de Mathematique Pures et Appliqueés* 5 (1972) 1–168.

[4] H. Brézis, *Opérateur Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert* (North-Holland, Amsterdam, 1973).

[5] F.E. Browder "Nonlinear maximal monotone operators in Banach spaces," *Mathematische Annalen* 176 (1968) 88–113.

[6] M. Chipot, *Variational Inequalities and Flow in Porous Media* (Springer, New York, 1984).
[7] G. Cohen, "Auxilliary problem principle and decomposition of optimization problems," *Journal of Optimization Theory and Applications* 32 (1980) 277–305.
[8] R.W. Cottle, F. Gianessi and J.L. Lions, *Variational Inequalities and Complementarity Problems: Theory and Applications* (Wiley, New York, 1980).
[9] C.W. Cryer, "The solution of a quadratic programming problem using systematic overrelaxation," *SIAM Journal on Control* 9 (1971) 385–392.
[10] J.E. Dennis and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1983).
[11] G. Duvuat and J.L. Lions, *Inequalities in Physics and Mechanics* (Springer, Berlin, 1976).
[12] B.C. Eaves, "A locally quadratic algorithm for computing stationary points," Technical Report, Department of Operations Research, Stanford University (Stanford, CA, 1978).
[13] P.T. Harker and J.S. Pang, "Finite dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications," *Mathematical Programming* 48 (1990) 161–220.
[14] C.M. Ip and J. Kyparisis, "Local convergence of quasi-Newton methods for B-differentiable equations," *Mathematical Programming* 56 (1992) 71–90.
[15] N.H. Josephy, "Newton's method for generalized equations," Technical Report No. 1965, Mathematics Research Center, University of Wisconsin (Madison, WI, 1979).
[16] N.H. Josephy, "A Newton method for the PIES energy model," Technical Report No. 1977, Mathematics Research Center, University of Wisconsin (Madison, WI, 1979).
[17] L.V. Kantorovich and G.P. Akilov, *Functional Analysis* (Pergamon, New York, 1982).
[18] S. Karamardian, "The complementarity problem," *Mathematical programming* 2 (1972) 107–129.
[19] M. Kojima and S. Shindo, "Extensions of Newton and quasi-Newton methods to systems of PC^1 equations," *Journal of the Operations Research Society of Japan* 29 (1986) 352–374.
[20] J.L. Lions and G. Stampacchia, "Variational inequalities," *Communications in Pure and Applied Mathematics* 20 (1967) 493–519.
[21] O.L. Mangasarian, "Solution of symmetric linear complementarity problems by iterative methods," *Journal of Optimization Theory and Applications* 22 (1977) 465–485.
[22] G.J. Minty, "On the monotonicity of the gradient of a convex function," *Pacific Journal of Mathematics* 14 (1964) 243–247.
[23] G.J. Minty, "Monotone (nonlinear) operators in Hilbert space," *Duke Mathematics Journal* 29 (1973) 341–346.
[24] J.M. Ortega, *Introduction to Parallel and Vector Solution of Linear Systems* (Plenum, New York, 1988).
[25] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).
[26] A.M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces* (Academic Press, 1973).
[27] J.S. Pang, "Newton's method for B-differentiable equations," *Mathematics of Operations Research* 15 (1990) 311–341.
[28] J.S. Pang and D. Chan, "Iterative methods for variational and complementarity problems," *Mathematical Programming* 24 (1982) 284–313.
[29] J.S. Pang and L. Qi, "Nonsmooth equations: motivations and algorithms," *SIAM Journal on Optimization* 3 (1993) 443–465.
[30] L. Qi, "Convergence analysis of some algorithms for solving non-smooth equations," *Mathematics of Operations Research* 18 (1993) 227–244.
[31] S.M. Robinson, "Generalized equations and their solutions, part I: basic theory," *Mathematical Programming Study* 10 (1979) 128–141.
[32] S.M. Robinson, "Strongly regular generalized equations," *Mathematics of Operations Research* 5 (1980) 43–62.
[33] S.M. Robinson, "Generalized equations," in: A. Bachem, M. Grëtschel and B. Korte, eds., *Mathematical Programming: the State of the Art* (Springer, Berlin, 1982) pp. 346–367.
[34] G. Stampacchia, "Formes bilineares coercitives sur les ensembles convexes," *Comptes Rendus del L'Academie des Science de Paris* 258 (1964) 4413–4416.
[35] R.A. Tapia, "The Kantorovich Theorem for Newton's method," *American Mathematical Monthly* 78 (1971) 389–392.

[36] L.U. Uko, "The solution of finite dimensional variational inequalities using systematic relaxation," *Journal of the Nigerian Mathematical Society* 8 (1989) 61–75.

[37] L.U. Uko, "The generalized Newton's method," *Journal of the Nigerian Mathematical Society* 10 (1991) 55–64.

[38] L.U. Uko, "On a class of general strongly nonlinear quasivariational inequalities," *Rivista di Matematica Pura ed Applicata* 11 (1992) 47–55.

[39] L.U. Uko, "Remarks on the generalized Newton method," *Mathematical Programming* 59 (1993) 405–412.