IC/95/328

INTERNATIONAL CENTRE FOR
THEORETICAL PHYSICS

A SMOOTH GENERALIZED NEWTON METHOD FOR A CLASS OF NON-SMOOTH EQUATIONS

Livinus U. Uko

INTERNATIONAL ATOMIC ENERGY AGENCY

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

VOL. 27 № 07

MIRAMARE-TRIESTE

IC/95/328

International Atomic Energy Agency
and
United Nations Educational Scientific and Cultural Organization
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

A SMOOTH GENERALIZED NEWTON METHOD FOR A CLASS OF NON-SMOOTH EQUATIONS[1]

Livinus U. Uko[2]
International Centre for Theoretical Physics, Trieste, Italy.

ABSTRACT

This paper presents a Newton-type iterative scheme for finding the zero of the sum of a differentiable function and a multivalued maximal monotone function. Local and semi-local convergence results are proved for the Newton scheme, and an analogue of the Kantorovich theorem is proved for the associated modified scheme that uses only one Jacobian evaluation for the entire iteration. Applications in variational inequalities are discussed, and an illustrative numerical example is given.

Abbreviated title: Generalized Newton method.

Mathematics Subject Classification (1991): 47H15, 65H10, 65K05, 65K10, 49J40, 47H19

Keywords: Generalized Equations, Newton method, nonsmooth minimization, Variational inequalities, Nonlinear complementarity problem, Kantorovich theorem.

MIRAMARE - TRIESTE

October 1995

[1]Submitted to Proceedings of the American Mathematical Society.

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

1. Introduction

Let H be a real Hilbert space with norm $\|\cdot\|$ and associated scalar product $(\cdot,\cdot)$, and let f be a Frechét-differentiable function mapping a subset H into H. Let g be a maximal monotone subset of $H \times H$, in the sense [2,10] that

$$(y_2 - y_1, x_2 - x_1) \geq 0, \quad \forall [x_1, y_1], [x_2, y_2] \in g$$

and the resolvent maps $Q_k = (1 + kg)^{-1}$ are non-expansive single-valued maps for all $k > 0$. In the sequel, we will regard the statements $[x, y] \in g$ and $g(x) \ni y$ as synonymous.

We are interested in the iterative solution of the generalized equation

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

This problem includes as a special cases many problems occurring in nonsmooth optimization [9,15,19], nonlinear complementarity theory [3,9], variational inequality theory [1,4,6].

It is easy to see that a vector v solves the generalized equation (1.1) if and only if it can be expressed in the form $v = Q_k(u)$ for some $k > 0$ where u solves the non-smooth equation

$$F_k(u) \equiv f(Q_k(u)) + \frac{1}{k}(u - Q_k(u)) = 0. \tag{1.2}$$

In fact, if v solves (1.1) then, for all $k > 0$, it obviously satisfies relation $v + kg(v) \ni v - kf(v)$. This relation can be rewritten in terms of the resolvent map Q_k as the equation $v = Q_k(v - kf(v))$. Therefore, if we set $u = v - kf(v)$, then $v = Q_k(u)$, and we obtain (1.2) by substituting for v in the previous equation. Conversely, if (1.2) holds, then it follows from the definition of the resolvent that $-f(Q_k(u)) \in g(Q_k(u))$, which shows that the vector $v = Q_k(u)$ solves (1.1).

In this paper we will approximate the solution of problem (1.2) with the solutions of the the following Newton-type iterative scheme

$$f'(Q_k(u_m))(u_{m+1} - u_m) = \frac{1}{k}(Q_k(u_m) - u_m) - f(Q_k(u_m)), \quad m = 0, 1, \ldots \tag{1.3}$$

If $g = H \times \{0\}$, this scheme reduces to the classical method of Newton, a good account of which can be found in [8,11,12,20]. A related method is the modified scheme

$$f'(Q_k(u_0))(u_{m+1} - u_m) = \frac{1}{k}(Q_k(u_m) - u_m) - f(Q_k(u_m)), \quad m = 0, 1, \ldots \tag{1.4}$$

using only one derivative evaluation for the entire iteration. It is also possible to use variable stepsizes in (1.3), in which case we obtain the scheme:

$$f'(Q_{k_m}(u_m))(u_{m+1} - u_m) = \frac{1}{k_m}(Q_{k_m}(u_m) - u_m) - f(Q_{k_m}(u_m)), \quad m = 0, 1, \ldots \tag{1.5}$$

Some other Newton-type schemes have been proposed in the Literature for the solution of (1.1) and (1.2). The Josephy-Newton method attempts to solve (1.1) by iteratively solving the partially linearized problems

$$f'(v_m)v_{m+1} + g(v_{m+1}) \ni f'(v_m)v_m - f(v_m), \quad m = 0, 1, \ldots \tag{1.6}$$

This scheme has nice convergence properties and has been studied by many authors, especially Josephy [7], Robinson [19], Eaves [5], Harker and Pang [6], Pang and Chan [14], Uko [22,23]. Its major drawback arises from the fact that successive iterates are defined implicitly and have to be computed with some unspecified inner-iterative method.

Robinson [18] recently proposed the Newton-type scheme

$$DF_k(u_m, u_{m+1}) = 0$$

for the solution of (1.2) where DF_k is a 'point-based approximation' (PBA) for F_k. This scheme suffers from the same drawbacks as the Josephy-Newton scheme, for successive iterates have be computed by solving nonlinear and nonsmooth equations.

If f has a Lipschitz continuous Fréchet derivative, then it is easy to see using the arguments employed in [18, Proposition 4] that the function $Df(x,y) = f(x) + f'(x)(y-x)$ is a PBA for f, and $DF_k(x,y) = f(Q_k(x)) + f'(Q_k(x))(Q_k(y) - Q_k(x)) + \frac{1}{k}(y - Q_k(y))$ is a PBA for F_k. The Robinson-Newton scheme corresponding to this PBA can therefore be expressed in the form

$$f'(Q_k(u_m))(Q_k(u_{m+1}) - Q_k(u_m)) + \frac{1}{k}(u_{m+1} - Q_k(u_{m+1})) = -f(Q_k(u_m)), \quad m = 0,1,\dots \quad (1.7)$$

This scheme is related to our Newton scheme (1.3) and can be viewed as its fully implicit version. It is even more closely related to the Josephy-Newton scheme, for if $u_0 = Q_k(v_0)$, the sequence of iterates generated by (1.6) and (1.7) are related in the form $u_m = Q_k(v_m)$.

Numerous other authors (cf. [13,15,16,17]) have also proposed Newton-type schemes which are applicable to special cases of problem (1.2).

The subproblems occurring at each stage of the iterations in our Newton schemes (1.3) – (1.5) are linear equations. This gives these schemes an edge over the Josephy-Newton, the Robinson-Newton scheme and some other Newton-like schemes that have been proposed in the Literature [14,16,17]. The major drawback of the Newton schemes (1.3) – (1.5) is the need to compute the resolvent operators Q_k. Although this is not feasible in general, we will see in Section 3 that that it possible for a reasonably large class of problems which occur in applications.

The rest of the paper is subdivided into two sections. Section 2 begins with some local and semi-local convergence results for the Newton scheme and concludes with and an analogue of the Kantorovich theorem for the modified scheme (1.4). In Section 3 we discuss applications in variational inequalities and give an illustrative numerical example.

2. Convergence results

Let D_0 be the interior of a closed convex subset D of H, and let $f\colon D \mapsto H$ be a continuous function that is Fréchet differentiable at all points of D_0. We assume that there exists $M > 0$ such that

$$\|f'(x) - f'(y)\| \le M\|x - y\|, \quad \forall x, y \in D_0. \quad (2.1)$$

It is well known (cf. [11, Page 70]) that this Lipschitz condition implies that

$$\|f(x) - f(y) - f'(y)(x-y)\| \le \frac{M}{2}\|x-y\|^2, \quad (2.2)$$

$$\|f(x) - f(z) - f'(y)(x-z)\| \le M\max\{\|x-y\|, \|z-y\|\}\|x-z\|, \quad (2.3)$$

forall $y \in D_0$ and $x, z \in D$. We will also impose one or more conditions of the form

$$\|f'(Q_k(u_0))^{-1} - kI\| \le b - k, \qquad \forall k < b \quad (2.4)$$

$$\|f'(x)^{-1} - kI\| \le b - k, \quad \forall x \in D_0, \quad \forall k < b \quad (2.5)$$

$$(f'(x)h, z) = (f'(x)z, h), \quad \forall x \in D_0, \quad \forall h, z \in H \quad (2.6)$$

$$(f'(x)z, z) \ge b^{-1}\|z\|^2, \quad \forall x \in D_0, \quad \forall z \in H. \quad (2.7)$$

The symmetry and coercivity conditions (2.6) and (2.7) occur in a natural way when f is the gradient $\nabla\psi$, where $\psi\colon \mathbf{R}^n \mapsto \mathbf{R}$ is a strongly convex C^2 function. However, the condition (2.5) is more general than conditions (2.6) and (2.7) taken together. In fact, if (2.6) and (2.7) hold, then $\|f'(x)^{-1}\| \le b$ and, for any $k < b$, $I - kf'(x)$ is symmetric and self-adjoint. Therefore

$$\|I - kf'(x)\| = \sup_{\|z\|\le 1}(z - kf'(x)z, z) \le 1 - kb^{-1}. \quad (2.8)$$

This implies

$$\|f'(x)^{-1} - kI\| \le \|f'(x)^{-1}\|\|I - kf'(x)\| \le b(1 - kb^{-1}) \le b - k,$$

which shows that (2.5) holds.

The basic properties of the resolvent operators required in this paper are stated in the following Lemma.

Lemma 2.1. *For all $x, y \in H$ and $k > 0$, we have*

$$\|Q_k(x) - Q_k(y)\| \leq \|x - y\|, \tag{2.9}$$

$$\|x - y - Q_k(x) + Q_k(y)\|^2 \leq \|x - y\|^2 - \|Q_k(x) - Q_k(y)\|^2. \tag{2.10}$$

Proof. The definition of the resolvent implies that $Q_k(x)+kg(Q_k(x)) \ni x$ and $Q_k(y)+kg(Q_k(y)) \ni y$. Therefore, using the monotonicity of g we obtain $\|Q_k(x) - Q_k(y)\|^2 \leq (x - y, Q_k(x) - Q_k(y))$. The inequality (2.9) now follows from the Schwartz inequality, and the inequality (2.10) follows from the estimate

$$\begin{aligned}\|x - y - Q_k(x) + Q_k(y)\|^2 &= \|x - y\|^2 - 2(x - y, Q_k(x) - Q_k(y)) + \|Q_k(x) - Q_k(y)\|^2 \\ &\leq \|x - y\|^2 - \|Q_k(x) - Q_k(y)\|^2.\end{aligned}$$

□

Further properties of resolvents can be found in [2,10].

We now analyse the convergence of the Newton schemes (1.3) – (1.5). The first Theorem is on the linear convergence of (1.3) when it is assumed that (1.2) has a solution and the initial vector is close enough to the true solution.

Theorem 2.2. *Let g be a maximal monotone operator and suppose that (1.2) has a solution u, and that (2.1) and (2.5) hold. Let*

$$Mb\|u - u_0\| < 2 \tag{2.11}$$

and suppose that the stepsize is chosen to satisfy

$$0 < \frac{b}{k} - 1 < 1 - \frac{Mb}{2}\|u - u_0\|. \tag{2.12}$$

Then the iterates from (1.3) converge linearly to u.

Proof. It follows from (1.2) and (1.3) that

$$\begin{aligned}u_{m+1} - u &= f'(Q_k(u_m))^{-1}[f(Q_k(u)) - f(Q_k(u_m)) - f'(Q_k(u_m))(Q_k(u) - Q_k(u_m))] \\ &\quad + (\frac{1}{k}f'(Q_k(u_m))^{-1} - I)(u - Q_k(u) - u_m + Q_k(u_m))\end{aligned}$$

Therefore, on using (2.2) and (2.5) we get

$$\begin{aligned}\|u_{m+1} - u\| &\leq b\|f(Q_k(u)) - f(Q_k(u_m)) - f'(Q_k(u_m))(Q_k(u) - Q_k(u_m))\| \\ &\quad + (\frac{b}{k} - 1)\|u - Q_k(u) - u_m + Q_k(u_m)\| \\ &\leq \frac{Mb}{2}\|Q_k(u_m) - Q_k(u)\|^2 + (\frac{b}{k} - 1)\|u - Q_k(u) - u_m + Q_k(u_m)\|\end{aligned} \tag{2.13}$$

This implies that

$$\|u_{m+1} - u\| \leq (\frac{Mb}{2}\|u_m - u\| + \frac{b}{k} - 1)\|u_m - u\|.$$

An induction argument using (2.12) and this inequality shows that $\|u_m - u\| \leq \|u_0 - u\|$ for all m. Therefore, letting $L = \frac{Mb}{2}\|u - u_0\| + \frac{b}{k} - 1$, we immediately obtain the estimate

$$\|u_{m+1} - u\| \leq L\|u_m - u\|.$$

It follows that $\|u_m - u\| \leq L^m\|u - u_0\| \to 0$ as $m \to \infty$, which completes the proof. □

Remark 2.3: The inequality (2.13) implies that

$$\begin{aligned}\|u_{m+1}-u\| &\le \frac{Mb}{2}\|u_m-u\|^2\\ &+(\frac{b}{k}-1-\frac{Mb}{2}\|u-Q_k(u)-u_m+Q_k(u_m)\|)\|u-Q_k(u)-u_m+Q_k(u_m)\|\\ &\le \begin{cases}\frac{Mb}{2}\|u_m-u\|^2 & \text{if } \frac{b}{k}-1\le\frac{Mb}{2}\|u-Q_k(u)-u_m+Q_k(u_m)\|\\ (\frac{Mb}{2}\|u_m-u\|+\frac{b}{k}-1)\|u_m-u\| & \text{otherwise.}\end{cases}\end{aligned}$$

Now, the term $u-Q_k(u)$ determines the degree of nonsmoothness in (1.2) and can be viewed as the 'constraint' in the problem, and a vector x could be said to satisfy this constraint if $x-Q_k(x)=u-Q_k(u)$. Therefore, the preceding inequality could interpreted to mean that the Newton scheme converge quadratically whenever iterates violate the constraint beyond a certain threshold, and only linearly otherwise. In fact, if stepsizes are chosen iteratively in (1.5) in such a way that

$$0<\frac{b}{k_m}-1\le\frac{Mb}{2}\|u-Q_{k_m}(u)-u_m+Q_{k_m}(u_m)\|, \tag{2.14}$$

then the quadratic convergence criterion will prevail all through and successive iterates will satisfy the estimate $\|u_{m+1}-u\|\le\frac{Mb}{2}\|u_m-u\|^2$ for all m. Therefore, letting $d=\frac{Mb}{2}\|u-u_0\|$ and using an induction argument, we see that iterates converge quadratically at the rate

$$\|u_m-u\|\le\frac{2}{Mb}d^{2^m}.$$

However, in practice, the solution to (1.2) will not be known in advance, so that the stepsize rule (2.14) will not be feasible.

It would be interesting to see whether it is possible to vary the stepsizes iteratively in such a way that linear convergence occurs at the beginning of iterations while quadratic convergence occurs towards the end. This would be a cost-effective way of globalizing the convergence of our Newton schemes.

Theorem 2.2 is an "attraction" theorem which states that the iterates from the Newton scheme (1.3) converges whenever the initial vector is close enough to the unknown solution. In the next result is of the semi-local type and states that if the first iterate u_1 is close enough to the initial vector u_0, then (1.2) has a solution, and subsequent iterates converge to that solution.

Theorem 2.4. *Let g be a maximal monotone operator and suppose that (2.1), (2.6) and (2.7) hold. If*

$$Mb\|u_1-u_0\|<2, \tag{2.15}$$

$$0<\frac{b}{k}-1<1-\frac{Mb}{2}\|u_1-u_0\|, \tag{2.16}$$

where u_1 is the first iterate in (1.3), then the iterates from (1.3) converge linearly to a solution of (1.2).

Proof. Making use of (1.3) and the analoguous equation with m replaced with $m-1$, we obtain the equation

$$\begin{aligned}f'(Q_k(u_m))(u_{m+1}-u_m) &= f(Q_k(u_m))-f(Q_k(u_{m-1}))-f'(Q_k(u_{m-1}))(Q_k(u_m)-Q_k(u_{m-1}))\\ &+(\frac{1}{k}-f'(Q_k(u_{m-1})))(u_m-Q_k(u_m)-u_{m-1}+Q_k(u_{m-1})).\end{aligned}$$

Since (2.6) and (2.7) hold, (2.5) and (2.8) also hold. Therefore, on using (2.2), (2.5) and (2.8) we get

$$\begin{aligned}\|u_{m+1}-u_m\| &\le b\|f(Q_k(u_m))-f(Q_k(u_{m-1}))-f'(Q_k(u_{m-1}))(Q_k(u_m)-Q_k(u_{m-1}))\|\\ &+b\|(\frac{1}{k}-f'(Q_k(u_{m-1})))(u_{m-1}-Q_k(u_{m-1})-u_m+Q_k(u_m))\|\\ &\le\frac{Mb}{2}\|Q_k(u_m)-Q_k(u_{m-1})\|^2+b(\frac{1}{k}-\frac{1}{b})\|u_{m-1}-Q_k(u_{m-1})-u_m+Q_k(u_m)\|\\ &\le(\frac{Mb}{2}\|u_m-u_{m-1}\|+\frac{b}{k}-1)\|u_m-u_{m-1}\|\end{aligned}$$

A straightforward induction argument using (2.16) and this inequality establishes that $\|u_m - u_{m-1}\| \leq \|u_1 - u_0\|$ for all $m \geq 1$. Therefore, if $L = \frac{Mb}{2}\|u_1 - u_0\| + \frac{b}{k} - 1$, then the inequality $\|u_{m+1} - u_m\| \leq L\|u_m - u_{m-1}\|$ holds for all m. This implies that

$$\|u_{p+m+1} - u_{m+1}\| \leq [1 + L + \cdots + L^{p-1}]\|u_{m+1} - u_m\| \leq \frac{L^m(1-L^p)}{1-L}\|u_1 - u_0\|$$

for all $p, m \geq 1$, and shows that u_m is a Cauchy sequence. We conclude, by completeness, that u_m converges to some u in H. Letting m tend to infinity in (1.3) we see that u satisfies (1.2), and letting p tend to infinity in the last inequality above, we obtain the linear error estimate

$$\|u - u_{m+1}\| \leq \frac{L^m}{1-L}\|u_1 - u_0\|.$$

That completes the proof. □

Remark 2.5: Under the hypothesis of Theorem 2.4, it is easy to see that the analogue of Remark 2.3 holds, in the sense that

$$\frac{2}{Mb}(\frac{b}{k} - 1) \leq \|u_{m-1} - Q_k(u_{m-1}) - u_m + Q_k(u_m)\| \implies \|u_{m+1} - u_m\| \leq \frac{Mb}{2}\|u_m - u_{m-1}\|^2.$$

In the sequel, given any $u_0 \in D_0$ and $r > 0$, $B[u_0, r]$ will designate the set $\{x \in H\colon \|x - u_0\| \leq r\}$. The last result in this section on the convergence of the the classical modified Newton scheme.

Theorem 2.6. *Let g be a maximal monotone operator and suppose that (2.1) and (2.4) hold. Let*

$$r = \frac{1}{Mb}\left[1 - \sqrt{1 - 2Mab - 4(bk^{-1} - 1)^2}\,\right]$$

and suppose that

$$\|f'(Q_k(u_0))^{-1}[f(Q_k(u_0)) + \frac{1}{k}(u_0 - Q_k(u_0))]\| \leq a \tag{2.17}$$

$$2Mab < 1 \tag{2.18}$$

$$0 < \frac{b}{k} - 1 < \sqrt{\frac{1}{5}(1 - 2Mab)}. \tag{2.19}$$

Then (1.2) has a unique solution u in $B[u_0, r]$, and the iterates from the modified Newton scheme (1.4) converge to u at the rate

$$\|u - u_m\| \leq \frac{1}{Mb}\left[bk^{-1} - \sqrt{1 - 2Mab - 4(bk^{-1} - 1)^2}\,\right]^m \tag{2.20}$$

Proof. For any $x \in B[u_0, r]$, let

$$w(x) = x - f'(Q_k(u_0))^{-1}f(Q_k(x)) + \frac{1}{k}f'(Q_k(u_0))^{-1}(Q_k(x) - x).$$

Then it is easy to see that u solves (1.2) if and only if it is a fixed point for w. Moreover, the modified Newton scheme (1.4) is the merely successive approximation scheme $u_{m+1} = w(u_m)$.

We have

$$\begin{aligned} w(x) - u_0 &= f'(Q_k(u_0))^{-1}[f(Q_k(x)) - f(Q_k(u_0)) - f'(Q_k(u_0))(Q_k(x) - Q_k(u_0))] \\ &+ [\frac{1}{k}f'(Q_k(u_0))^{-1} - I](u_0 - x - Q_k(u_0) + Q_k(x)) \\ &+ f'(Q_k(u_0))^{-1}[\frac{1}{k}(Q_k(u_0) - u_0) - f(Q_k(u_0))]. \end{aligned}$$

Therefore, making use of (2.4), (2.17) and (2.2), we obtain

$$\|w(x) - u_0\| \le a + \frac{Mb}{2}\|Q_k(x) - Q_k(u_0)\|^2 + (bk^{-1} - 1)\|u_0 - x - Q_k(u_0) + Q_k(x)\|.$$

The right hand side in this inequality is bounded above by

$$a + \frac{Mb}{2}\|x - u_0\|^2 + (bk^{-1} - 1 - \frac{Mb}{2}\|u_0 - x - Q_k(u_0) + Q_k(x)\|)\|u_0 - x - Q_k(u_0) + Q_k(x)\|.$$

If $\frac{b}{k} - 1 \le \frac{Mb}{2}\|u_0 - x - Q_k(u_0) + Q_k(x)\|$, this latter term is bounded above by $a + \frac{Mb}{2}\|x - u_0\|^2$; otherwise, it is bounded above by $a + \frac{Mb}{2}\|Q_k(x) - Q_k(u_0)\|^2 + \frac{2}{Mb}(\frac{b}{k} - 1)^2$. In either case, we have

$$\begin{aligned} \|w(x) - u_0\| &\le a + \frac{Mb}{2}\|x - u_0\|^2 + \frac{2}{Mb}(\frac{b}{k} - 1)^2 \\ &\le a + \frac{Mbr^2}{2} + \frac{2}{Mb}(\frac{b}{k} - 1)^2 = r. \end{aligned}$$

It follows that w maps $B[u_0, r]$ into itself.

Given $x, y \in B[u_0, r]$, we have

$$\begin{aligned} w(x) - w(y) &= f'(Q_k(u_0))^{-1}[f(Q_k(y)) - f(Q_k(x)) - f'(Q_k(u_0))(Q_k(y) - Q_k(x))] \\ &\quad + [\frac{1}{k}f'(Q_k(u_0))^{-1} - I](y - x - Q_k(y) + Q_k(x)). \end{aligned}$$

Therefore, using (2.5) and (2.3), we obtain

$$\begin{aligned} \|w(x) - w(y)\| &\le Mb\max\{\|y - u_0\|, \|x - u_0\|\}\|x - y\| + (bk^{-1} - 1)\|y - x - Q_k(y) + Q_k(x)\| \\ &\le (Mbr + bk^{-1} - 1)\|x - y\|. \end{aligned}$$

Therefore, on setting $L = Mbr + bk^{-1} - 1 = bk^{-1} - \sqrt{1 - 2Mab - 4(bk^{-1} - 1)^2}$ we see that

$$\|w(x) - w(y)\| \le L\|x - y\|, \quad \forall x, y \in B[u_0, r].$$

The condition (2.18) implies that w is a strict contraction map on the set $B[u_0, r]$. Therefore it follows from Banach's contraction mapping principle that w has a unique fixed point in this set. The error estimate (2.20) follows immediately since

$$\|u_m - u\| = \|w(u_{m-1}) - w(u)\| \le L^{m-1}\|u_0 - u\| \le L^{m-1}r = \frac{L^m}{Mb}.$$

That completes the proof. □

If $g = H \times \{0\}$, we can take $\frac{b}{k} - 1$ to be an arbitrarily small positive number, in which case we recover the Kantorovich theorem [8, Chapter 18.1, Theorem 6] on the classical modified Newton scheme.

3. Implementation and applications in variational inequalities

In order to solve concrete problems with the Newton schemes (1.3) – (1.5) we require an algorithm for computing the values of the resolvent operators $Q_k(x) = (1 + kg)^{-1}$ for any given $x \in H$. No such algorithm is available in the general case. However, there exist some interesting problems of the form (1.1) which occur in applications and for which the pertinent resolvents are easily computed. This class includes many variational inequality problems, obtained on setting $g = \partial\phi$ in (1.1), where $\phi\colon H \mapsto (-\infty, \infty]$ is a proper lower semicontinuous convex function and

$$\partial\phi(x) \equiv \{v \in H : \phi(x) - \phi(y) \le (v, x - y) \quad \forall y \in H\}.$$

In this case (1.1) reduces to the variational inequality

$$f(v) + \partial\phi(v) \ni 0. \tag{3.1}$$

Such problems have important applications in the physical and engineering sciences and in many other fields [1,4,6].

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

$$\phi(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{otherwise.} \end{cases} \tag{3.2}$$

In this case problem (3.1) reduces to the search for $u \in C$ satisfying

$$(f(u), u - v) \leq 0, \quad \forall v \in C. \tag{3.3}$$

This problem will be designated in the sequel as VI(f,C,H). It is well known (cf. [24, Lemma 4.1]) that resolvent operator $Q_k = (1 + k\partial\phi)^{-1}$ corresponding to this problem is independent of k and coincides with the orthogonal projection onto C, which is computable in some cases (for instance, whenever C is a polyhedral set).

In principle, every variational inequality of the form (3.1) can be expressed in the form VI(F,C,V), where $V = H \times \mathbf{R}$, equipped with the scalar product $\langle [x,\lambda],[y,\tau]\rangle = (x,y) + \lambda\tau$, $C = \{[x,\lambda] \in V : \phi(x) \leq \lambda\}$, and $F[x,\lambda] = [f(x),1]$ for all $[x,\lambda] \in V$. It is not difficult to verify that if v solves (3.1) then $[v,\phi(v)]$ solves VI(F,C,V), and if $[v,\lambda]$ solves VI(F,C,V) then v solves (3.1). However, the derivative of F is given by $F'[x,\lambda][y,\tau] = [f'(x)y, 0]$ for all $[x,\lambda],[y,\tau] \in V$, and is not invertible at any point, so that the Newton-type schemes (1.3) – (1.5) are not applicable to VI(F,C,V). Newton-type iterative schemes should be applied directly to (3.1) and not to its equivalent problem of the the type (3.3).

The resolvent $Q_k(x)$ corresponding to the problem (3.1) is the unique value of v that minimizes the non-smooth function

$$\psi(v) = \frac{1}{2}\|x - v\|^2 + k\phi(v). \tag{3.4}$$

Therefore, resolvent computations for variational inequalities could in principle be done using non-smooth minimization techniques. This not always possible, for the associated nonsmooth minimum problems are not easily intractable. However, there are interesting special cases in which resolvents are relatively easy to compute.

The nonlinear complementarity problem

$$f_i(v_1,\ldots,v_n) \geq 0, \quad v_i \geq 0, \quad i = 1,\ldots,n, \quad \sum_{j=1}^{n} v_j f_j(v_1,\ldots,v_n) = 0 \tag{3.5}$$

is one such case. This problem corresponds to the variational inequality (3.1), with $H = \mathbf{R}^n$ and $C = \{x \in \mathbf{R}^n : x_i \geq 0,\ i = 1,\ldots,n\}$. It is easy to verify that the relevant resolvent formula for this problem is given by

$$Q_k(x_1,\ldots,x_n) = (x_1^+,\ldots,x_n^+),$$

where $t^+ = \max\{t,0\}$.

More generally, given scalar constants $g_{i1} < 0 < g_{i2}$ and $h_{i1} \leq h_{i2}$ ($i = 1, \ldots, n$), let ϕ be the convex function $\phi(x) = \sum_{i=1}^{n} \phi_i(x_i)$, where

$$\phi_i(t) = \begin{cases} g_{i1}(t - h_{i1}), & t < h_{i1} \\ 0, & h_{i1} \leq t \leq h_{i2} \\ g_{i2}(t - h_{i2}), & t > h_{i2}. \end{cases} \tag{3.6}$$

It not difficult to see that resolvent this convex function is given by the expression [21]

$$Q_k(x_1,\ldots,x_n) = (Q_{1,k}(x_1),\ldots,Q_{n,k}(x_n))$$

where

$$Q_{i,k}(t) = \begin{cases} t - kg_{i1} & \text{if } t < h_{i1} + kg_{i1} \\ t + (h_{i1} - t)^+ - (t - h_{i2})^+ & \text{if } h_{i1} + kg_{i1} \leq t \leq h_{i2} + kg_{i2} \\ t - kg_{i2} & \text{if } t > h_{i2} + kg_{i2}. \end{cases} \tag{3.7}$$

If we choose $g_{i1} = -\infty$ and $g_{i2} = \infty$ for all i then ϕ becomes an indicator function of the type (3.2), with $C = [h_{11}, h_{12}] \times \cdots \times [h_{n1}, h_{n2}]$. In this case $Q_{i,k}$ reduces to the simpler truncation form $Q_{i,k}(t) = t + (h_{i1} - t)^+ - (t - h_{i2})^+$. If we now let $h_{11} = \cdots = h_{n1} = 0$ and $h_{12} = \cdots = h_{n2} = \infty$ we recover the nonlinear complementarity problem (3.5). The convex functions (3.6) also occur in variational inequalities modelling heat flow through thick walls [4].

Another interesting class of convex functions are the functions $\phi(x) = \frac{1}{p}\|x\|^p$ with exponents $p \geq 1$. If $p > 1$, it is easy to see that the resolvent is given by the expression $Q_k(x) = \frac{x}{1+kq^{p-2}}$ where q is obtained by solving the scalar equation $q + kq^{p-1} = \|x\|$. If the value of the exponent p is either 3, 4 or 5, a formula for q can be obtained from Cardan's formulae, thereby giving $Q_k(x)$ in closed form [21]. In the general case we solve for q iteratively (and efficiently) with Newton's method. The case $p = 1$ is more tricky, and is handled by the following Lemma.

Lemma 3.1. *The resolvent $Q_k = (1 + k\partial\phi)^{-1}$ of the convex function $\phi(x) = \|x\|$ is given by the formula*

$$Q_k(x) = \begin{cases} x - k\frac{x}{\|x\|} & \text{if } \|x\| > k \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If $\|x\| > k$, and we let $v_k = x - k\frac{x}{\|x\|}$, then it is easy to verify that $v_k + kv_k/\|v_k\| = x$. Therefore, in this case, we have $Q_k(x) = v_k$.

Let $\|x\| \leq k$, and let ψ be the convex function defined in (3.4). Then, for any $v \in H$, we have

$$\begin{aligned}\psi(v) &= \frac{1}{2}\|v\|^2 - (v, x) + \frac{1}{2}\|x\|^2 + k\|v\| \\ &\geq \frac{1}{2}\|v\|^2 + (k - \|x\|)\|v\| + \frac{1}{2}\|x\|^2 \\ &\geq \frac{1}{2}\|x\|^2 = \psi(0).\end{aligned}$$

Therefore, ψ assumes its minimum value at 0, which shows that $Q_k(x) = 0$. That completes the proof. □

Further examples on resolvent computations can be found in [21,24].

In [23] the performance of the Josephy Newton scheme (1.6) was illustrated with a problem of the form (3.1) with $f(x,y) = (f_1(x,y), f_2(x,y))$ where $f_1(x,y) = x - 0.71\sin x - 0.473\cos y$, $f_2(x,y) = y - 0.71\cos x + 0.473\sin y$, and $\phi(x,y) = \phi_1(x) + \phi_2(y)$ where ϕ_1 and ϕ_2 are defined by (3.6) with $g_{11} = -1$, $g_{12} = 2$, $g_{21} = -2$, $g_{22} = 1$ and $h_{11} = -2$, $h_{12} = 2$, $h_{21} = 2$ and $h_{22} = 4$. We illustrate the performance of the Newton scheme (1.3) by solving this problem with the same initial vector $(x_0, y_0) = (5, 3)$ used in [23]. The results corresponding to the constant stepsize $k = 1.5$ are tabulated below.

m	x_m	y_m	$f_1(x_m, y_m) + \partial\phi_1(x_m)$	$f_2(x_m, y_m) + \partial\phi_2(y_m)$
1	0.89017756	2.63585654	0.75216624	2.41821755
2	2.00000000	1.53361275	1.33681505	0.30175005
3	0.90255678	2.00000000	0.54210553	[−0.00982197, 1.99017803]
4	−0.83341284	2.00000000	−0.11101187	[−0.04727323, 1.95272677]
5	−0.48104342	2.00000000	0.04431418	[−0.19932627, 1.80067373]
6	−0.59411325	2.00000000	0.00016387	[−0.15824041, 1.84175959]
7	−0.65422497	2.00000000	−0.02532097	[−0.13330138, 1.86669862]
8	−0.59057026	2.00000000	0.00161989	[−0.15964484, 1.84035516]

These values were obtained by computing the iterates (u_m, v_m) of the Newton scheme (1.3) and then computing the approximate solutions $x_m = Q_k(u_m)$ and $y_m Q_k(v_m)$ with the resolvent formulae (3.7).

These results compare favourably with the computations reported in [23] in which a final solution of $(x, y) = (-0.59451124, 2.00000000)$ was obtained after 8 Josephy-Newton iterations. However, the computations in [23] were done in conjunction with a nonlinear Gauss-Seidel inner iterative scheme taken from [21] and required considerably more computation time.

ACKNOWLEDGMENTS

The author is grateful to the International Centre for Theoretical Physics, Trieste, for hospitality.

REFERENCES

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

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

[3] R.W. Cottle, F. Gianessi and J.L. Lions, Variational inequalities and complementarity problems: theory and applications (John Wiley, New York, 1980).

[4] G. Duvuat, J.L. Lions, Inequalities in physics and mechanics (Springer-Verlag, Berlin, 1976).

[5] B.C. Eaves,"A locally quadratic algorithm for computing stationary points", Technical Report, Department of Operations Research, Stanford University (Stanford, CA, May 1978).

[6] 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.

[7] N.H. Josephy, "Newton's method for generalized equations", Technical Report No. 1965, Mathematics Research Center, University of Wisconsin (Madision, WI, 1979).

[8] L.V. Kantorovich and G.P. Akilov, Functional Analysis (Pergamon Press, New York, 1982).

[9] S. Karamardian, "The complementarity problem", Mathematical programming 2(1972)107-129.

[10] G.J. Minty, "Monotone (nonlinear) operators in Hilbert space", Duke Mathematics Journal 29(1973)341-346.

[11] J.M. Ortega and W.C. Rheinboldt, Iterative solution of nonlinear equations in several variables (Academic Press, New York, 1970).

[12] A.M. Ostrowski, Solution of equations in Euclidean and Banach spaces (Academic Press, 1973).

[13] J.S. Pang, "Newton's method for B-differentiable equations", Mathematics of Operations Research 15(1990)311-341.

[14] J.S. Pang and D. Chan, "Iterative methods for variational and complementarity problems", Mathematical Programming 24(1982)284-313.

[15] J.S. Pang and L. Qi, "Nonsmooth equations: motivations and algorithms", SIAM Journal on Optimization 3(1993)443-465.

[16] L. Qi, "Convergence analysis of some algorithms for solving non-smooth equations", Mathematics of Operations Research 18(1993)227-244.

[17] L. Qi, J. Sun, "A nonsmooth version of Newton's method", Mathematical Programming 58(1993)353-367.

[18] S.M. Robinson, "Newton's method for a class of nonsmooth functions", Set-Valued Analysis 2(1994)291-305.

[19] S.M. Robinson, "Generalized equations", in: A. Bachem, M. Grötschel and B. Korte, eds., Mathematical programming: the state of the art (Springer, Berlin, 1982) pp. 346-367.

[20] R.A. Tapia, "The Kantorovich Theorem for Newton's method", American Mathematical Monthly 78(1971)389-392.

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

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

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

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