CORE DISCUSSION PAPER

2006/107

Solving strongly monotone variational and quasi-variational inequalities

Yu. Nesterov* and L. Scrimali †

November, 2006

Abstract

In this paper we develop a new and efficient method for variational inequality with Lipschitz continuous strongly monotone operator. Our analysis is based on a new *strongly convex* merit function. We apply a variant of the developed scheme for solving quasivariational inequality. As a result, we significantly improve the standard sufficient condition for existence and uniqueness of their solutions. Moreover, we get a new numerical scheme, which rate of convergence is much higher than that of the straightforward gradient method.

Keywords: variational inequality, quasivariational inequality, monotone operators, complexity analysis, efficiency estimate, optimal methods.

*Center for Operations Research and Econometrics (CORE), Catholic University of Louvain (UCL), 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium; e-mail: nesterov@core.ucl.ac.be.

The research results presented in this paper have been supported by a grant "Action de recherche concertè ARC 04/09-315" from the "Direction de la recherche scientifique - Communautè française de Belgique". The scientific responsibility rests with its author(s).

†The research presented in the paper was done at CORE. Current address: Dipartimento di Matematica e Informatica Universit di Catania Viale Andrea Doria 6 95125 Catania Italy; e-mail: scrimali@dmi.unict.it.

1 Introduction

1.1 Motivation

A notable generalization of the variational inequality problem is the quasi-variational inequality problem, introduced by Bensoussan et al. [2] in the context of impulse control problems. A thorough study of these problems can be found in [1, 5, 8]. Many important and useful applications of these mathematical tools are known, which range from Nash games to transportation network equilibria. For instance, we may refer to Bensoussan [3] and Harker [6], who recognized the connection between generalized Nash games and quasi-variational inequalities. Recently, Pang and Fukushima [18] applied this result in order to formulate the noncooperative multi-leader-follower game in terms of generalized Nash games, which in turn (under opportune assumptions), can be expressed as a quasi-variational inequality. Kocvara and Outrata [7] dealt with a class of quasi-variational inequalities with applications to engineering. Wei and Smeers [21] introduced a formulation of a spatial oligopolistic electricity model with Cournot generators and regulated transmission prices in terms of quasi-variational inequalities. Bliemer and Bovy [4] discussed a quasi-variational inequality formulation of the dynamic traffic assignment problem. Applications to some economic and financial models can be found in [20], so that a wide class of problems can be solved as a quasi-variational inequality.

From the point of view of solution methods, quasi-variational inequalities do not have an extensive literature, we may address the reader only to few papers [4, 7, 15, 17, 19]. For variational inequalities (VI), the situation is better. Recently, there were proposed two new methods for solving VI with Lipschitz continuous operator [9, 13]. These methods can be seen as an alternative to smoothing technique developed for accelerating the gradient-type methods on large-scale structured convex optimization problems (see [12]). However, up to now, there is almost no specialized methods for VI with strongly monotone operators. The main goal of this paper is to close this gap. We also present a interesting application of the developed schemes to quasivariational inequalities.

1.2 Contents

In Section 2 we consider the standard VI-problem with Lipschitz continuous strongly monotone operator. For justifying the rate of convergence of numerical methods, we introduce a new *strongly convex* merit function. For the value of this function, we update recursively a simple quadratic model of our problem. Our technique can be seen as a mixture of the approach from [13] and the elements of estimate functions (see Section 2.2 [11]). As a result, we obtain a simple method, which significantly outperforms the straightforward gradient scheme. In Section 3 we introduce the problem of quasivariational inequality (QVI) and recall the main known existence results [16]. For the sake of completeness, we present a complexity analysis of corresponding gradient scheme.

In Section 4 we show how to apply the results of Section 2 to QVI. For that, we introduce a *relaxation operator* and establish the sufficient conditions for it to be a contraction. These conditions allow to guarantee the existence of a unique solution to QVI for a much wider class of problem as compared to [16]. On the other hand, we show that the simple iteration scheme with approximately computed values of the relaxation operator is much more efficient than the usual gradient method. The required approximate values of the operator are computed by a variant of the method developed in Section 2.

1.3 Notation

In this paper we denote by E a finite-dimensional real vector space. Notation E^* is used for the dual space. The value of a linear function $s \in E^*$ at $x \in E$ (that is the *scalar product* of s and x) is denoted by $\langle s, x \rangle$. The operator $B : E \to E^*$ is positive definite if

$$\langle Bx, x \rangle \quad > \quad 0 \quad \forall x \in E \setminus \{0\}.$$

It is self-adjoint if

$$\langle Bx, y \rangle \quad = \quad \langle By, x \rangle \quad \forall x, y \in E.$$

Using such an operator, we define on E and E^* the Euclidean norms:

$$\begin{aligned} \|x\| &= \langle Bx, x \rangle^{1/2}, \quad x \in E, \\ \|s\|_* &= \max_{x \in E} \{ \langle s, x \rangle : \ \|x\| \leq 1 \} \\ &= \langle s, B^{-1} s \rangle^{1/2}, \quad s \in E^*. \end{aligned}$$

By $\pi_Q(x)$ we denote the Euclidean projection of point x onto the set Q. The necessary and sufficient characterizations of the projection are as follows:

$$\begin{aligned} \pi_Q(x) &\in Q, \\ \langle B(\pi_Q(x) - x), y - \pi_Q(x) \rangle &\geq 0 \quad \forall y \in Q. \end{aligned} \tag{1.1}$$

2 Solving strongly monotone VI

Let Q be a closed convex set. Consider a continuous operator $g(x) : \ Q \to E^*$, which is *strongly monotone*:

$$\langle g(x) - g(y), x - y \rangle \quad \geq \quad \mu \, \|x - y\|^2, \quad \forall x, y \in Q. \tag{2.1}$$

The constant $\mu \geq 0$ is called the *parameter of strong monotonicity* of operator g. If $\mu = 0$, then g is a monotone operator. In what follows, we always assume $\mu > 0$.

In this section, the problem of our interest is the following *variational inequality* (VI):

$$\text{Find } x^*(Q) \in Q \ : \quad \langle g(x^*(Q)), y - x^*(Q) \rangle \geq 0 \quad \forall y \in Q. \tag{2.2}$$

In the absence of ambiguity, we often use a shortcut $x^* \equiv x^*(Q)$. Since g is strongly monotone, the solution x^* of problem (2.2) satisfies inequality

$$\langle g(y), x^* - y \rangle + \tfrac{1}{2}\mu \|y - x^*\|^2 \overset{(2.1)}{\leq} \langle g(x^*), x^* - y \rangle - \tfrac{1}{2}\mu \|y - x^*\|^2 \overset{(2.2)}{\leq} 0, \tag{2.3}$$

which is valid for all $y \in Q$. Since the main subject of this paper are numerical schemes, we simply assume that the solution x^* does exist. Clearly, in this case it is unique.

In order to speak about quality of approximate solutions to (2.2), we need to introduce the following *merit function*:

$$f(x) = \sup_{y \in Q} \left\{ \langle g(y), x - y \rangle + \tfrac{1}{2}\mu \|y - x\|^2 \right\}. \tag{2.4}$$

Theorem 1 *Merit function $f(x)$ is well defined and strongly convex on Q with convexity parameter μ. Moreover, it is non-negative on Q and vanishes only at the unique solution of variational inequality (2.2).*

Proof:
Indeed, in view of strong monotonicity of operator g, at any $x \in Q$ we have

$$\begin{aligned} \langle g(y), x - y \rangle + \tfrac{1}{2}\mu \|y - x\|^2 &\overset{(2.1)}{\leq} \langle g(x), x - y \rangle - \tfrac{1}{2}\mu \|y - x\|^2 \\ &\leq \|g(x)\|_* \cdot \|x - y\| - \tfrac{1}{2}\mu \|y - x\|^2 \leq \tfrac{1}{2\mu} \|g(x)\|_*^2. \end{aligned}$$

Thus, the value $f(x)$ is well defined. Further, function $f(x)$ is strongly convex in x with convexity parameter μ as a maximum of parametric family of strongly convex (in x) quadratic functions:

$$f(x) = \max_{y \in Q} \phi_y(x), \quad \phi_y(x) = \langle g(y), x - y \rangle + \tfrac{1}{2}\mu \|y - x\|^2, \; y \in Q.$$

Finally, if $x \in Q$, then clearly $f(x) \geq 0$. Consider now x^*, the solution to (2.2). In view of (2.3), we have $f(x^*) = 0$. On the other hand, if $f(\hat{x}) = 0$ for some $\hat{x} \in Q$, then is a solution to the weak variational inequality

$$\langle g(y), y - \hat{x} \rangle \geq 0 \quad \forall y \in Q.$$

Since $g(y)$ is continuous, this implies that $\hat{x}$ is a solution to (2.2). $\Box$

Let us show how we can generate points with small values of the merit function $f(\cdot)$. Consider a sequence of arbitrary points $\{y_i\}_{i=0}^N \subset Q$ and a sequence of positive weights $\{\lambda_i\}_{i=0}^N$. Denote

$$\begin{aligned} S_N &= \sum_{i=0}^N \lambda_i, \qquad \tilde{y}_N = \frac{1}{S_N} \sum_{i=0}^N \lambda_i y_i, \\ \Delta_N &= \max_{x \in Q} \left\{ \sum_{i=0}^N \lambda_i \left[\langle g(y_i), y_i - x \rangle - \tfrac{1}{2}\mu \|x - y_i\|^2 \right] \right\}. \end{aligned} \tag{2.5}$$

Note that the computation of Δ_N can be reduced to finding a Euclidean projection on Q.

Lemma 1

$$f(\tilde{y}_N) \leq \frac{1}{S_N} \Delta_N. \tag{2.6}$$

Proof:
Indeed, since $g(x)$ is a strongly monotone operator, we have

$$\begin{aligned}
f(\tilde{y}_N) &= \sup_{x \in Q} \left\{ \langle g(x), \tilde{y}_N - x \rangle + \tfrac{1}{2}\mu \| x - \tilde{y}_N \|^2 \right\} \\
&= \tfrac{1}{S_N} \sup_{x \in Q} \left\{ \sum_{i=0}^{N} \lambda_i \langle g(x), y_i - x \rangle + \tfrac{1}{2}\mu S_N \| x - \tilde{y}_N \|^2 \right\} \\
&\le \tfrac{1}{S_N} \sup_{x \in Q} \left\{ \sum_{i=0}^{N} \lambda_i \langle g(x), y_i - x \rangle + \tfrac{1}{2}\mu \sum_{i=0}^{N} \lambda_i \| x - y_i \|^2 \right\} \\
&\overset{(2.1)}{\le} \tfrac{1}{S_N} \max_{x \in Q} \left\{ \sum_{i=0}^{N} \lambda_i \left[\langle g(y_i), y_i - x \rangle - \tfrac{1}{2}\mu \| x - y_i \|^2 \right] \right\} \equiv \tfrac{1}{S_N} \Delta_N .
\end{aligned}$$

$\square$

Thus, our goal is to control the growth of Δ_N as compared to the sum S_N. For $\beta > 0$, denote

$$\begin{aligned}
\psi_y^\beta(x) &= \langle g(y), y - x \rangle - \tfrac{1}{2}\beta \| x - y \|^2, \\
\Psi_k(x) &= \sum_{i=0}^{k} \lambda_i \psi_{y_i}^{\mu}(x).
\end{aligned}$$

Note that $\psi_y^\beta(x)$ is a strongly concave quadratic function with concavity parameter β. Function $\Psi_k(x)$ is strongly concave with parameter μS_k. Clearly, $\Delta_k = \max\limits_{x \in Q} \Psi_k(x)$.

Consider the following iteration:

$$\begin{aligned}
x_k &= \arg\max_{x \in Q} \Psi_k(x), \\
y_{k+1} &= \arg\max_{x \in Q} \psi_{x_k}^{\beta}(x).
\end{aligned} \tag{2.7}$$

Theorem 2 *If $\lambda_{k+1} \le \frac{\mu}{\beta} S_k$, then*

$$\Delta_{k+1} \le \Delta_k + \tfrac{1}{2}\lambda_{k+1} \left[\tfrac{1}{\mu+\beta} \| g(y_{k+1}) - g(x_k) \|_*^2 - \beta \| y_{k+1} - x_k \|^2 \right]. \tag{2.8}$$

Proof:
Note that $\Psi_{k+1}(x) = \Psi_k(x) + \lambda_{k+1} \psi_{y_{k+1}}^{\mu}(x)$. Hence,

$$\begin{aligned}
\Delta_{k+1} &= \max_{x \in Q} \left\{ \Psi_k(x) + \lambda_{k+1} \psi_{y_{k+1}}^{\mu}(x) \right\} \\
&\le \Delta_k + \max_{x \in Q} \left\{ \langle \nabla \Psi_k(x_k), x - x_k \rangle - \tfrac{1}{2}\mu S_k \| x - x_k \|^2 + \lambda_{k+1} \psi_{y_{k+1}}^{\mu}(x) \right\} \\
&\le \Delta_k + \max_{x \in Q} \left\{ -\tfrac{1}{2}\mu S_k \| x - x_k \|^2 + \lambda_{k+1} \left[\langle g(y_{k+1}), y_{k+1} - x \rangle - \tfrac{1}{2}\mu \| x - y_{k+1} \|^2 \right] \right\}.
\end{aligned}$$

In view of definition of y_{k+1}, we have

$$\langle -g(x_k) - \beta B(y_{k+1} - x_k), x - y_{k+1}\rangle \;\leq\; 0 \quad \forall x \in Q.$$

Therefore,

$$\begin{array}{l} \langle g(y_{k+1}), y_{k+1} - x\rangle - \frac{1}{2}\mu\|x - y_{k+1}\|^2 \\[2ex] = \; \langle g(y_{k+1}) - g(x_k), y_{k+1} - x\rangle - \frac{1}{2}\mu\|x - y_{k+1}\|^2 + \langle g(x_k), y_{k+1} - x\rangle \\[2ex] \leq \; \|g(y_{k+1}) - g(x_k)\|_* \cdot \|y_{k+1} - x\| - \frac{1}{2}\mu\|x - y_{k+1}\|^2 + \beta\langle B(y_{k+1} - x_k), x - y_{k+1}\rangle. \end{array}$$

Note that

$$2\langle B(y_{k+1} - x_k), x - y_{k+1}\rangle \;\;=\;\; \|x - x_k\|^2 - \|y_{k+1} - x_k\|^2 - \|x - y_{k+1}\|^2.$$

Hence,

$$\begin{array}{l} \langle g(y_{k+1}), y_{k+1} - x\rangle - \frac{1}{2}\mu\|x - y_{k+1}\|^2 \\[2ex] \leq \; \frac{1}{2(\mu+\beta)}\|g(y_{k+1}) - g(x_k)\|_*^2 + \frac{1}{2}\beta\|x - x_k\|^2 - \frac{1}{2}\beta\|y_{k+1} - x_k\|^2. \end{array}$$

Putting all inequalities together, and using the upper bound on λ_{k+1}, we obtain the inequality (2.8). $\Box$

Corollary 1 *Assume that $g(x)$ is Lipschitz continuous on Q:*

$$\|g(x) - g(y)\|_* \;\;\leq\;\; L\|x - y\|, \quad \forall x, y \in Q. \tag{2.9}$$

Then for the choice

$$\beta \;=\; L, \quad \lambda_{k+1} \;=\; \tfrac{\mu}{L}\, S_k,$$

we have $\Delta_{k+1} \leq \Delta_k$.

Now we are ready to write down an algorithmic scheme for solving the variational inequality (2.2) with strongly monotone Lipschitz continuous operator. This method can be seen as a combination of the dual extrapolation method [13] with the technique of estimate functions (see Section 2.2 in [11]).

For simplicity, we assume that the constants μ and L are known. Denote by $\gamma = \frac{L}{\mu} \geq 1$ the *condition number* of the operator g.

Method for strongly monotone VI

Input : Choose $\bar{x} \in Q$. Set $\lambda_0 = 1$, and $y_0 = \bar{x}$.

Iteration $(k \geq 0)$**:**

$$\begin{array}{rcl} x_k & = & \arg\max\limits_{x \in Q} \Psi_k(x), \\ y_{k+1} & = & \arg\max\limits_{x \in Q} \psi^L_{x_k}(x), \\ \lambda_{k+1} & = & \frac{1}{\gamma} \cdot S_k. \end{array} \tag{2.10}$$

Output : $\tilde{y}_k = \frac{1}{S_k} \sum\limits_{i=0}^{k} \lambda_i y_i$.

Theorem 3 *Under conditions (2.1) and (2.9), for any $k \geq 0$, we have*

$$\begin{array}{rcl} \frac{\mu}{2} \cdot \|\tilde{y}_k - x^*\|^2 \;\leq\; f(\tilde{y}_k) & \leq & \left[f(\bar{x}) + \frac{\mu \cdot (\gamma^2 - 1)}{2} \cdot \|\bar{x} - x^*\|^2 \right] \cdot \exp\left\{ -\frac{k}{\gamma+1} \right\} \\ & \leq & f(\bar{x}) \cdot \gamma^2 \cdot \exp\left\{ -\frac{k}{\gamma+1} \right\}. \end{array} \tag{2.11}$$

Proof:
The first and the last inequality in (2.11) follow from strong convexity of function $f(x)$ with parameter μ. Let us prove the middle one.

In view of Corollary 1 we have $\Delta_{k+1} \leq \Delta_k$ for any $k \geq 0$. Note that $S_0 = \lambda_0 = 1$, and

$$S_{k+1} = S_k + \lambda_{k+1} = \left(1 + \frac{1}{\gamma} \right) S_k.$$

Hence,

$$f(\tilde{y}_k) \;\stackrel{(2.6)}{\leq}\; \frac{\Delta_k}{S_k} \;\leq\; \Delta_0 \cdot \left(1 - \frac{1}{\gamma+1} \right)^k \;\leq\; \Delta_0 \cdot \exp\left\{ -\frac{k}{\gamma+1} \right\}.$$

It remains to estimate Δ_0. Note that

$$\begin{array}{rcl} \Delta_0 & = & \max\limits_{x \in Q} \Psi_0(x) \;=\; \max\limits_{x \in Q} \psi^{\mu}_{y_0}(x) \;=\; \max\limits_{x \in Q} \left\{ \langle g(\bar{x}), \bar{x} - x \rangle - \frac{\mu}{2} \cdot \|x - \bar{x}\|^2 \right\} \\ & = & \max\limits_{x \in Q} \left\{ \langle g(\bar{x}) - g(x^*), \bar{x} - x \rangle + \langle g(x^*), \bar{x} - x \rangle - \frac{\mu}{2} \cdot \|x - \bar{x}\|^2 \right\} \\ & \stackrel{(2.2)}{\leq} & \langle g(x^*), \bar{x} - x^* \rangle + \max\limits_{x \in Q} \left\{ \langle g(\bar{x}) - g(x^*), \bar{x} - x \rangle - \frac{\mu}{2} \cdot \|x - \bar{x}\|^2 \right\} \\ & \stackrel{(2.9)}{\leq} & \langle g(x^*), \bar{x} - x^* \rangle + \frac{L^2}{2\mu} \cdot \|\bar{x} - x^*\|^2. \end{array}$$

Since $\langle g(x^*), \bar{x} - x^* \rangle + \frac{\mu}{2} \cdot \|x^* - \bar{x}\|^2 \stackrel{(2.4)}{\leq} f(\bar{x})$, we obtain the middle part of (2.11). $\square$

Note that the problem (2.2) can be solved by a standard gradient-type method:

$$\begin{array}{rcl} x_0 & = & \bar{x} \in Q, \\ x_{k+1} & = & \pi_Q(x_k - \lambda B^{-1} g(x_k)), \quad k \geq 0. \end{array} \tag{2.12}$$

However, it is well known that this method converges very slowly. For the reader convenience, let us estimate its rate of convergence. Since $x^* \overset{(2.2)}{=} \pi_Q(x^* - \lambda B^{-1} g(x^*))$, choosing the optimal step $\lambda = \frac{\mu}{L^2}$, we obtain

$$\|x_{k+1} - x^*\|^2 \;\leq\; \|x_k - \lambda B^{-1} g(x_k) - (x^* - \lambda B^{-1} g(x^*))\|^2$$

$$= \; \|x_k - x^*\|^2 - 2\lambda \langle g(x_k) - g(x^*), x_k - x^* \rangle + \lambda^2 \|g(x_k) - g(x^*)\|_*^2$$

$$\overset{(2.1),(2.9)}{\leq} \; (1 - 2\lambda\mu + \lambda^2 L^2) \cdot \|x_k - x^*\|^2 \;\leq\; \|x_k - x^*\|^2 \cdot \exp\left\{ -\frac{k}{\gamma^2} \right\}.$$

For big values of the condition number, this estimate is much worse than (2.11). Note that the rate of convergence (2.11) cannot be improved by any black-box method as applied to the problem class (2.1), (2.9) (see [10]). At the same time, from the viewpoint of implementation, the method (2.10) is comparable with (2.12): at each iteration, it needs two projections and two computations of the operator instead of one in (2.12).

3 Quasi-variational inequalities

Let $\mathcal{Q} : E \to 2^E$ be a multifunction with nonempty closed and convex values. We are interested in the following *quasi-variational inequality problem* (QVI):

$$\text{Find } x_* \in \mathcal{Q}(x_*) : \quad \langle g(x_*), y - x_* \rangle \geq 0, \quad \forall y \in \mathcal{Q}(x_*). \tag{3.1}$$

For the reader convenience we recall the following existence result (see [16], Theorem 9).

Theorem 4 *Suppose that the following assumptions hold:*

- *(a) Operator g is Lipschitz continuous and strongly monotone on E with constants L and $\mu > 0$ respectively.*
- *(b) There exists $\alpha < \frac{1}{\gamma(\gamma + \sqrt{\gamma^2 - 1})}$ such that*

$$\|\pi_{\mathcal{Q}(x)}(z) - \pi_{\mathcal{Q}(y)}(z)\| \leq \alpha \|x - y\|, \quad \forall x, y, z \in E. \tag{3.2}$$

Then the problem (3.1) has a unique solution.

We will get this result as a corollary of Theorem 5 justifying the rate of convergence of the gradient method as applied to problem (3.1). Now we just mention that assumption (3.2) is a kind of strengthening of the contraction property for multifunction $\mathcal{Q}(x)$. Let us give an example of such a mapping.

Lemma 2 *Let function $c(x) : E \to E$ be Lipschitz continuous:*

$$\|c(x) - c(y)\| \le \alpha \|x - y\|, \quad x, y \in E.$$

And let $\bar{Q}$ be a closed convex set. Then

$$\mathcal{Q}(x) \stackrel{\mathrm{def}}{=} c(x) + \bar{Q}$$

satisfies (3.2) with the same value of α.

Proof:
Indeed, for arbitrary x and z from E we have

$$\pi_{c(x)+\bar{Q}}(z) = c(x) + \pi_{\bar{Q}}(z - c(x))$$

Denote $z_1 = z - c(x)$, $z_2 = z - c(y)$. Since we project in Euclidean norm, we have

$$\begin{aligned}
\|\pi_{c(x)+\bar{Q}}(z) - \pi_{c(y)+\bar{Q}}(z)\|^2 &= \|c(x) + \pi_{\bar{Q}}(z - c(x)) - c(y) - \pi_{\bar{Q}}(z - c(y))\|^2 \\
&= \|z_2 - \pi_{\bar{Q}}(z_2) - z_1 + \pi_{\bar{Q}}(z_1)\|^2 \\
&= \|z_2 - z_1\|^2 - 2\langle B(z_2 - z_1), \pi_{\bar{Q}}(z_2) - \pi_{\bar{Q}}(z_1)\rangle + \|\pi_{\bar{Q}}(z_2) - \pi_{\bar{Q}}(z_1)\|^2.
\end{aligned}$$

Note that

$$\begin{aligned}
\langle B(z_2 - z_1), \pi_{\bar{Q}}(z_2) - \pi_{\bar{Q}}(z_1)\rangle &= \langle B(z_2 - \pi_{\bar{Q}}(z_2) + \pi_{\bar{Q}}(z_2) - z_1), \pi_{\bar{Q}}(z_2) - \pi_{\bar{Q}}(z_1)\rangle \\
&\stackrel{(1.1)}{\ge} \langle B(\pi_{\bar{Q}}(z_2) - \pi_{\bar{Q}}(z_1) + \pi_{\bar{Q}}(z_1) - z_1), \pi_{\bar{Q}}(z_2) - \pi_{\bar{Q}}(z_1)\rangle \stackrel{(1.1)}{\ge} \|\pi_{\bar{Q}}(z_2) - \pi_{\bar{Q}}(z_1)\|^2.
\end{aligned}$$

Hence, $\|\pi_{c(x)+\bar{Q}}(z) - \pi_{c(y)+\bar{Q}}(z)\|^2 \le \|z_2 - z_1\|^2 \le \alpha^2 \|x - y\|^2$. □

Note that problem (3.1) can be solved by a standard gradient-type method:

$$x_{k+1} = \pi_{\mathcal{Q}(x_k)}(x_k - \lambda B^{-1} g(x_k)), \quad k \ge 0. \tag{3.3}$$

Let us estimate its rate of convergence.

$$\begin{aligned}
\|x_{k+1} - x_*\| &= \|\pi_{\mathcal{Q}(x_k)}(x_k - \lambda B^{-1} g(x_k)) - \pi_{\mathcal{Q}(x_*)}(x_* - \lambda B^{-1} g(x_*))\| \\
&= \|\pi_{\mathcal{Q}(x_k)}(x_k - \lambda B^{-1} g(x_k)) - \pi_{\mathcal{Q}(x_*)}(x_k - \lambda B^{-1} g(x_k)) \\
&\quad + \pi_{\mathcal{Q}(x_*)}(x_k - \lambda B^{-1} g(x_k)) - \pi_{\mathcal{Q}(x_*)}(x_* - \lambda B^{-1} g(x_*))\| \\
&\stackrel{(3.2)}{\le} \alpha \|x_k - x_*\| + \|\pi_{\mathcal{Q}(x_*)}(x_k - \lambda B^{-1} g(x_k)) - \pi_{\mathcal{Q}(x_*)}(x_* - \lambda B^{-1} g(x_*))\|.
\end{aligned}$$

Since g is strongly monotone and Lipschitz continuous, we have

$$\begin{aligned}
&\|\pi_{\mathcal{Q}(x_*)}(x_k - \lambda B^{-1} g(x_k)) - \pi_{\mathcal{Q}(x_*)}(x_* - \lambda B^{-1} g(x_*))\|^2 \\
\le\ & \|(x_k - \lambda B^{-1} g(x_k)) - (x_* - \lambda B^{-1} g(x^*))\|^2 \\
=\ & \|x_k - x_*\|^2 - 2\lambda \langle g(x_k) - g(x_*), x_k - x^*\rangle + \lambda^2 \|g(x_k) - g(x_*\|^2 \\
\le\ & (1 - 2\lambda\mu + \lambda^2 L^2)\, \|x_k - x_*\|^2.
\end{aligned}$$

Thus, we have proved the following theorem.

Theorem 5 *If operator g is strongly monotone and Lipschitz continuous with constants L and μ, and multifunction $\mathcal{Q}(x)$ satisfies condition (3.2) with $\alpha < \frac{1}{\gamma(\gamma+\sqrt{\gamma^2-1})}$, then the gradient method (3.3) with optimal stepsize $\lambda = \frac{\mu}{L^2}$ converges to the unique solution of problem (3.1) with the following rate:*

$$\|x_k - x_*\| \quad \leq \quad \exp\left\{-k \cdot \left(\frac{1}{\gamma(\gamma+\sqrt{\gamma^2-1})} - \alpha\right)\right\} \cdot \|x_0 - x_*\|. \tag{3.4}$$

Thus, we have seen that quasivariational inequality (3.1) is solvable by the gradient scheme (3.3) only if the variation rate of the feasible set $\mathcal{Q}(x)$ is very small as compared with inverse condition number of the operator g:

$$\alpha \quad < \quad \frac{1}{\gamma(\gamma+\sqrt{\gamma^2-1})} \;\approx\; \frac{1}{2\gamma^2}. \tag{3.5}$$

In the next section we will see that this limitation can be significantly weakened.

4 Relaxation operator for QVI

For problem (3.1), let us introduce the *relaxation operator* $T(x) \stackrel{\mathrm{def}}{=} x^*(\mathcal{Q}(x))$. In the case of strongly monotone operator g, operator $T(x)$ is fully defined by the following relations:

$$\begin{aligned} T(x) \quad &\in \quad \mathcal{Q}(x), \\ \langle g(T(x)), y - T(x)\rangle \quad &\geq \quad 0 \quad \forall y \in \mathcal{Q}(x). \end{aligned} \tag{4.1}$$

Clearly, the solution of problem (3.1) is a fixed point of this operator:

$$x_* = T(x_*).$$

It appears that under our standard conditions operator $T(x)$ is Lipschitz continuous.

Theorem 6 *Suppose that operator g is Lipschitz continuous and strongly monotone with constants L and $\mu > 0$ respectively. Assume that there exists some $\alpha \geq 0$ such that*

$$\|\pi_{\mathcal{Q}(x)}(z) - \pi_{\mathcal{Q}(y)}(z)\| \quad \leq \quad \alpha\|x - y\|, \quad \forall x, y, z \in E. \tag{4.2}$$

Then $T(x)$ is Lipschitz continuous with constant $\alpha\gamma$, where $\gamma = L/\mu$.

Proof:
Let us fix two arbitrary points $x_1, x_2 \in E$. Denote

$$T_i = T(x_i), \quad g_i = g(T_i), \quad Q_i = \mathcal{Q}(x_i), \quad i = 1, 2.$$

Let us fix some $\lambda > 0$. By definition,

$$T_1 \quad \stackrel{(1.1),(4.1)}{=} \quad \pi_{Q_1}(T_1 - \lambda B^{-1} g_1).$$

Denote $y_2 = \pi_{Q_2}(T_1 - \lambda B^{-1} g_1)$. By condition (4.2), we have

$$\|y_2 - T_1\| \;\; \leq \;\; \alpha \|x_1 - x_2\|. \tag{4.3}$$

On the other hand, $\langle B(y_2 - (T_1 - \lambda B^{-1} g_1)), T_2 - y_2 \rangle \overset{(1.1)}{\geq} 0$. Therefore,

$$\begin{aligned} \langle B(y_2 - T_1), T_2 - y_2 \rangle \;\; & \geq \;\; \lambda \langle g_1, y_2 - T_2 \rangle \; = \; \lambda \langle g_1, y_2 - T_1 \rangle + \lambda \langle g_1, T_1 - T_2 \rangle \\ & \overset{(2.1)}{\geq} \;\; \lambda \langle g_1, y_2 - T_1 \rangle + \lambda \langle g_2, T_1 - T_2 \rangle + \lambda \mu \|T_1 - T_2\|^2 \\ & \overset{(4.1)}{\geq} \;\; \lambda \langle g_1 - g_2, y_2 - T_1 \rangle + \lambda \mu \|T_1 - T_2\|^2. \end{aligned}$$

Thus,

$$\begin{aligned} \lambda \mu \|T_1 - T_2\|^2 \;\; & \leq \;\; \lambda \langle g_1 - g_2, T_1 - y_2 \rangle + \langle B(y_2 - T_1), T_2 - y_2 \rangle \\ & \leq \;\; \lambda \langle g_1 - g_2, T_1 - y_2 \rangle + \langle B(y_2 - T_1), T_2 - T_1 \rangle \\ & \overset{(2.9)}{\leq} \;\; (1 + \lambda L) \cdot \|T_1 - T_2\| \cdot \|y_2 - T_1\|. \end{aligned}$$

Since λ can be chosen arbitrarily large, from (4.3) we obtain $\|T_1 - T_2\| \leq \frac{\alpha L}{\mu} \|x_1 - x_2\|$. $\square$

As a byproduct of our considerations, we get the following existence result.

Corollary 2 *If $\alpha < \gamma^{-1}$, then there exists a unique solution to problem (3.1).*

Proof:
Indeed, under conditions of the corollary, $T(x)$ is a contraction. $\square$

Note that the latter statement significantly strengthen the statement of Theorem 4. Moreover, we can combine it with the technique presented in Section 2 and develop an efficient numerical scheme for problem (3.1). Let us start with one auxiliary statement. In what follows, we always assume that operator g is Lipschitz continuous and strongly monotone with constants L and μ respectively.

Lemma 3 *Let $x^*(Q)$ be a solution for VI (2.2). For $\hat{x} \in Q$ define $\bar{x} = \arg\max_{y \in Q} \psi^{\beta}_{\hat{x}}(y)$ with some $\beta > 0$. Then*

$$\mu \|\hat{x} - x^*(Q)\|^2 + \beta \|\hat{x} - \bar{x}\|^2 \;\; \leq \;\; (\beta + L) \cdot \|\hat{x} - x^*(Q)\| \cdot \|\hat{x} - \bar{x}\|. \tag{4.4}$$

Therefore

$$\tfrac{\mu}{\beta + L} \cdot \|\hat{x} - x^*(Q)\| \;\; \leq \;\; \|\hat{x} - \bar{x}\| \; \leq \; \tfrac{\beta + L}{\beta} \cdot \|\hat{x} - x^*(Q)\|. \tag{4.5}$$

Moreover,

$$f(\bar{x}) \;\; \leq \;\; \tfrac{\beta^2 + L^2}{2\mu} \|\hat{x} - \bar{x}\|^2 \; \leq \; \tfrac{(\beta^2 + L^2) \cdot (\beta + L)^2}{2\mu\beta^2} \|\hat{x} - x^*(Q)\|^2. \tag{4.6}$$

Proof:
By definition,

$$\langle g(\hat{x}) + \beta B(\bar{x} - \hat{x}), y - \bar{x} \rangle \;\; \geq \;\; 0 \quad \forall y \in Q. \tag{4.7}$$

Therefore,

$$\begin{array}{rcl} \beta\langle B(\bar{x} - \hat{x}), x^*(Q) - \bar{x}\rangle & \geq & \langle g(\hat{x}), \bar{x} - x^*(Q)\rangle \\ & = & \langle g(\hat{x}), \bar{x} - \hat{x}\rangle + \langle g(\hat{x}), \hat{x} - x^*(Q)\rangle \\ & \stackrel{(2.1)}{\geq} & \langle g(\hat{x}), \bar{x} - \hat{x}\rangle + \langle g(x^*(Q)), \hat{x} - x^*(Q)\rangle + \mu\|\hat{x} - x^*(Q)\|^2 \\ & \stackrel{(2.2)}{\geq} & \langle g(\hat{x}) - g(x^*(Q)), \bar{x} - \hat{x}\rangle + \mu\|\hat{x} - x^*(Q)\|^2 \end{array}$$

Hence,

$$\begin{array}{rcl} \mu\|\hat{x} - x^*(Q)\|^2 + \beta\|\hat{x} - \bar{x}\|^2 & \leq & \beta\langle B(\bar{x} - \hat{x}), x^*(Q) - \hat{x}\rangle + \langle g(\hat{x}) - g(x^*(Q)), \hat{x} - \bar{x}\rangle \\ & \stackrel{(2.9)}{\leq} & (\beta + L) \cdot \|\hat{x} - x^*(Q)\| \cdot \|\hat{x} - \bar{x}\|, \end{array}$$

and (4.5) follows.

Further,

$$\begin{array}{rcl} f(\bar{x}) & = & \max\limits_{y \in Q}\{\langle g(y), \bar{x} - y\rangle + \frac{1}{2}\mu\|y - \bar{x}\|^2\} \stackrel{(2.1)}{\leq} \max\limits_{y \in Q}\{\langle g(\bar{x}), \bar{x} - y\rangle - \frac{1}{2}\mu\|y - \bar{x}\|^2\} \\ & \stackrel{(4.7)}{\leq} & \max\limits_{y \in Q}\{\langle g(\bar{x}) - g(\hat{x}) - \beta B(\bar{x} - \hat{x}), \bar{x} - y\rangle - \frac{1}{2}\mu\|y - \bar{x}\|^2\}. \end{array}$$

Since g is monotone,

$$\begin{array}{rcl} \|g(\bar{x}) - g(\hat{x}) - \beta B(\bar{x} - \hat{x})\|^2 & = & \|g(\bar{x}) - g(\hat{x})\|^2 - 2\beta\langle g(\bar{x}) - g(\hat{x}), \bar{x} - \hat{x}\rangle + \|\bar{x} - \hat{x}\|^2 \\ & \stackrel{(2.9)}{\leq} & (L^2 + \beta^2)\|\bar{x} - \hat{x}\|^2. \end{array}$$

Hence, $f(\bar{x}) \leq \frac{\beta^2 + L^2}{2\mu}\|\hat{x} - \bar{x}\|^2$, and the remaining inequality follows from the second inequality in (4.5). □

Let us present a modified version of method (2.10). The main difference with the original consists in a preliminary *damping step* in the spirit of [14]. This allows to obtain the efficiency estimates of the new scheme entirely in terms of distances to the solution.

$$\boxed{\begin{array}{ll} \multicolumn{2}{c}{\textbf{Fixed-length method for strongly monotone VI}} \\ \hline \textbf{Input}: & \text{Fix the number of steps } N \geq 1 \text{ and choose } \hat{x} \in Q. \\ & \text{Set } \lambda_0 = 1, \text{ and } y_0 = \bar{x} \stackrel{\mathrm{def}}{=} \arg\max\limits_{x \in Q} \psi^L_{\hat{x}}(x). \\ \textbf{For } k = 0 & \textbf{to } N-1 \textbf{ do:} \\ & \begin{array}{rcl} x_k & = & \arg\max\limits_{x \in Q} \Psi_k(x), \\ y_{k+1} & = & \arg\max\limits_{x \in Q} \psi^L_{x_k}(x), \\ \lambda_{k+1} & = & \frac{1}{\gamma} \cdot S_k. \end{array} \\ \textbf{Output}: & \tilde{y}_N(Q, \hat{x}) \stackrel{\mathrm{def}}{=} \frac{1}{S_N} \sum\limits_{i=0}^{N} \lambda_i y_i. \end{array}} \tag{4.8}$$

Theorem 7 *Assume that the operator g satisfies conditions (2.1), (2.9). Then the method (4.8) as applied to VI-problem (2.2) converges as follows:*

$$\| \tilde{y}_N(Q, \hat{x}) - x^*(Q) \| \quad \leq \quad 3\gamma \exp\left\{ -\tfrac{N}{2(\gamma+1)} \right\} \cdot \| \hat{x} - x^*(Q) \|, \quad N \geq 1. \tag{4.9}$$

Proof:
Indeed,

$$\begin{array}{rcl} \| \tilde{y}_N(Q, \hat{x}) - x^*(Q) \|^2 & \stackrel{(2.11)}{\leq} & \frac{2}{\mu} \left[f(\bar{x}) + \frac{\mu \cdot (\gamma^2 - 1)}{2} \cdot \| \bar{x} - x^*(Q) \|^2 \right] \cdot \exp\left\{ -\frac{N}{\gamma+1} \right\} \\ & \stackrel{(4.6)}{\leq} & \frac{2}{\mu} \left[\frac{4L^2}{\mu} + \frac{\mu \cdot (\gamma^2 - 1)}{2} \right] \cdot \exp\left\{ -\frac{N}{\gamma+1} \right\} \cdot \| \hat{x} - x^*(Q) \|^2 \\ & \leq & 9\gamma^2 \exp\left\{ -\frac{N}{\gamma+1} \right\} \cdot \| \hat{x} - x^*(Q) \|^2. \qquad \Box \end{array}$$

Let us discuss an efficient numerical strategy for solving a slowly changing QVI. The complexity of this problem depends on the condition number $\gamma = \frac{L}{\mu}$ of the operator g and on *contraction gap*

$$\delta \quad \stackrel{\mathrm{def}}{=} \quad 1 - \alpha\gamma, \tag{4.10}$$

which we assume to be positive. If the rate of variation α of the sets $\mathcal{Q}(x)$ is very small, then the gap can be close to one. For example, for QVI solvable by the gradient method (3.3), this gap is at least $\frac{1}{2}$ (see (3.5)).

Taking into account the estimate (4.9), let us define the minimal number of steps $\hat{N} = N(\alpha, \gamma)$, satisfying condition

$$3\gamma \exp\left\{ -\tfrac{\hat{N}}{2(\gamma+1)} \right\} \leq \tfrac{\delta}{4} \quad \Rightarrow \quad \hat{N} = \lfloor 2(\gamma+1) \ln \tfrac{12\gamma}{1-\alpha\gamma} \rfloor + 1. \tag{4.11}$$

Consider the following two-level numerical scheme:

$$\boxed{\begin{array}{l} \text{Choose } u_0 \in E. \text{ For } k \geq 0 \text{ iterate:} \\[2ex] \hat{x}_k = \pi_{\mathcal{Q}(u_k)}(u_k), \quad u_{k+1} = \tilde{y}_{\hat{N}}(\mathcal{Q}(u_k), \hat{x}_k). \end{array}} \tag{4.12}$$

Theorem 8 *Assume that $\delta > 0$. Then there exists a unique solution x_* to QVI (3.1), and method (4.12) converges as follows:*

$$\|u_k - x_*\| \;\leq\; \tfrac{1}{\delta} \cdot \exp\left\{-\tfrac{\delta}{2}k\right\} \cdot \|u_0 - T(u_0)\|. \tag{4.13}$$

Proof:
Denote $r_k = \|u_k - T(u_k)\|$. Note that

$$\begin{array}{rcl} \|u_{k+1} - T(u_k)\| & \stackrel{(4.9)}{\leq} & \frac{\delta}{4} \cdot \|\hat{x}_k - T(u_k)\| \stackrel{(4.12)}{\leq} \frac{\delta}{4} \cdot \|u_k - T(u_k)\|, \\[2ex] \|T(u_{k+1}) - T(u_k)\| & \stackrel{\text{T.6}}{\leq} & \alpha\gamma \cdot \|u_{k+1} - u_k\|. \end{array}$$

Therefore

$$\begin{array}{rcl} r_{k+1} & \leq & \|u_{k+1} - T(u_k)\| + \|T(u_{k+1}) - T(u_k)\| \;\leq\; \frac{\delta}{4} \cdot r_k + \alpha\gamma \cdot \|u_{k+1} - u_k\| \\[2ex] & \leq & \left(\alpha\gamma + \frac{\delta}{4}\right) \cdot r_k + \alpha\gamma \cdot \|u_{k+1} - T(u_k)\| \;\leq\; \left(1 - \frac{\delta}{2}\right) \cdot r_k. \end{array}$$

Thus, $r_k \leq \exp\left\{-\frac{\delta}{2}k\right\} \cdot r_0$. On the other hand,

$$r_k \;=\; \|u_k - T(u_k)\| \;\geq\; \|u_k - x_*\| - \|T(x_*) - T(u_k)\| \stackrel{\text{T.6}}{\geq} \delta \cdot \|u_k - x_*\|. \quad \Box$$

Corollary 3 *In order to get a point u_k with $\|u_k - x_*\| \leq \epsilon$, method (4.12) needs at most*

$$\tfrac{4(\gamma+1)}{1-\alpha\gamma} \cdot \ln \tfrac{12\gamma}{1-\alpha\gamma} \cdot \ln \tfrac{\|u_0 - T(u_0)\|}{\epsilon \cdot (1-\alpha\gamma)} \tag{4.14}$$

computations of operator $g(x)$.

It is easy to see that this estimate is much better than the corresponding result for the gradient method (3.3).

References

[1] C. Baiocchi, A. Capelo. *Variational and quasivariational inequalities: applications to Free boundary problems*. Wiley, New York, 1984.

[2] A. Bensoussan, M. Goursat, J.-L- Lions. Contrôle impulsionnel et inéquations quasi-variationnelle. *Compte rendu de l'Académie des Sciences Paris*, Série A **276**, 1279–1284 (1973).

[3] A. Bensoussan. Points de Nash dans le cas de fontionnelles quadratiques et jeux differentiels linaires a N personnes. *SIAM Journal on Control*, **12**, 460-499 (1974).

[4] M. Bliemer, P. Bovy. Quasi-Variational inequality formulation of the multiclass dynamic traffic assignment problem. *Transportation Research B*, **37**, 501–519 (2003).

[5] D. Chan, J.-S. Pang. The generalized quasi-variational inequality problem. *Mathematics of Operations Research*, **7** (2), 211–222 (1982).

[6] P.T. Harker. Generalized Nash games and quasivariational inequalities. *European Journal of Operations Research*, **54**, 81–94 (1991).

[7] M. Kocvara M., J.V. Outrata. On a class of quasi-variational inequalities. *Optimization Methods and Software*, **5**, 275-295 (1995).

[8] U. Mosco. Implicit variational problems and quasi variational inequalities. In: Proc. Summer School (Bruxelles, 1975) 'Nonlinear operators and Calculus of variations', Lectures Notes Math, no. 543, Springer-Verlag, Berlin, 83–156 (1976).

[9] A. Nemirovski. Prox-method with rate of convergence O(1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization* **15**, 229-251 (2004)

[10] A. Nemirovsky and D. Yudin. *Problem Complexity and Method Efficiency in Optimization*. Wiley, 1983.

[11] Yu. Nesterov. *Introductory Lectures on Convex Optimization*. Kluwer, Boston, 2004.

[12] Yu. Nesterov. Smooth minimization of nonsmooth function. *Mathematical Programming*, **103**(1), 127-152 (2005)

[13] Yu.Nesterov. Dual extrapolation and its applications to solving variational inequalities and related problems. CORE DP 2003/68. Accepted by *Mathematical Programming.*

[14] Yu. Nesterov. Minimizing functions with bounded variation of the gradient. CORE DP 2005/79.

[15] M.A. Noor, Z.A. Menon. Algorithms for general mixed quasi variational inequalities. *Journal of Inequalities in Pure and Applied Mathematics*, **3** (4) Art. 59 (2002).

[16] M. A. Noor, W. Oettli, On general nonlinear complementarity problems and quasi-equilibria, *Le Matematiche* XLIX, 313-331 (1994)

[17] J. Outrata, J. Zowe. A numerical approach to optimization problems with variational inequality constraints. *Mathematical Programming*, **68**, 105–130 (1995).

[18] J.-S. Pang, M. Fukushima. Quasi-variational inequalities, generalized Nash equilibria and Multi-leader-follower games. *Computational Management Science*, **1**, 21–56 (2005).

[19] Salahuddin. Projection methods for quasi-variational inequalities. *Mathematical and Computational Applications*, **9** (2), 125–131 (2004).

[20] J.C. Yao. The generalized quasi-variational inequality problem with applications. *Journal of Mathematical Analysis and Applications*, **158**, 139–160 (1991).

[21] J.Y. Wei, Y. Smeers. Spatial oligopolistic electricity models with Cournot generators and regulated transmission prices. *Operations Research*, **47**, 102-112 (1999).