

# A NEW ALGORITHM FOR SOLVING VARIATIONAL INEQUALITIES WITH APPLICATION TO THE TRAFFIC ASSIGNMENT PROBLEM

Patrice MARCOTTE

*Centre de Recherche sur les Transports, Université de Montréal, Canada*

Received 28 November 1983

Revised manuscript received 5 March 1985

The variational inequality problem in Euclidian space is formulated as a nonconvex, nondifferentiable optimization problem. We show that any stationary point is optimal, and we propose a solution algorithm that decreases the nondifferential objective monotonically. Application to the asymmetric traffic assignment problem is considered.

*Key words:* Variational Inequality, Nondifferentiable Optimization, Nonconvex Programming, Network Optimization.

## 1. Introduction

Recently, much attention has been focused on the variational inequality formulation of the generalized traffic assignment problem involving user-optimized behavior (Bertsekas and Gafni [2], Dafermos [4], Nguyen and Dupuis [17], Smith [21] for instance). Most solution methods have only been proposed in a transportation framework, although the algorithms are of more general application, and could be applied as well to solve  $N$ -person games or economic equilibrium problems. What singles out the multicommodity traffic assignment problem from the general variational inequality problem is the computational ease with which the linear approximation of the problem (replacing the cost mapping by a constant vector) can be solved, namely by performing shortest path computations.

In this paper the variational inequality problem is formulated as a nondifferentiable, nonconvex mathematical program. To evaluate the nondifferentiable objective, a linear approximation of the problem has to be solved; consequently this linear problem must possess some structure amenable to efficient solution procedures; the traffic assignment problem falls into such a class of problems.

After formulating the problem, we present the algorithm and give a proof of global convergence. Finally, we show how the algorithm can be implemented in the realm of large-scale traffic equilibrium models.

**2. Problem formulation. Basic definitions. Solution algorithms**

Consider a continuously differentiable mapping  $F: \Phi \rightarrow \mathbb{R}^n$ , where  $\Phi$  is a non-empty, convex and compact subset of  $\mathbb{R}^n$ . The variational inequality problem (VIP in short) associated with  $F$  and  $\Phi$  consists in finding a vector  $x^*$  in  $\Phi$  such that:

$$(x^* - y)^T F(x^*) \leq 0 \quad \text{for all } y \text{ in } \Phi \tag{1}$$

If we define a point-to-set mapping  $\Gamma$  as follows:

$$\Gamma(x) \triangleq \arg \min_{y \in \Phi} y^T F(x) \tag{2}$$

then a vector  $x^*$  solves VIP if and only if it is a fixed point of the uppersemicontinuous mapping  $\Gamma$ . Since  $\Phi$  is compact, Kakutani's fixed point theorem [12] ensures the existence of at least one solution to VIP.

**Definition 1.** A mapping  $F$  is:

1. Monotone on  $\Phi$  if:

$$\begin{aligned} (x - y)^T (F(x) - F(y)) &\geq 0 \quad \text{for all } x, y \text{ in } \Phi, \\ [(x - y)^T F'(x)(x - y) &\geq 0]. \end{aligned}$$

2. Strictly monotone on  $\Phi$  if:

$$(x - y)^T (F(x) - F(y)) > 0 \quad \text{for all } x, y \text{ in } \Phi (x \neq y).$$

3. Strongly monotone on  $\Phi$  if:

$$\begin{aligned} (x - y)^T (F(x) - F(y)) &\geq \eta \|x - y\|^2 \quad \text{for all } x, y \text{ in } \Phi, \\ [(x - y)^T F'(x)(x - y) &\geq \eta \|x - y\|^2], \end{aligned}$$

where  $\eta$  is some positive constant.

The relations inside the square brackets apply whenever  $F$  is continuously differentiable on  $\Phi$  with Jacobian  $F'$ .<sup>1</sup> See Auslender [1] for further details.

It is easily verified that the solution set of VIP is compact, convex if  $F$  is monotone, and a singleton if  $F$  is strictly monotone.

In the remainder of the paper we will assume that  $F$  is monotone. In this particular case, many solution algorithms have parallels in (convex) optimization: projection with predetermined stepsize (Glowinski et al. [8]), Newton and quasi-Newton methods (Josephy [10, 11]), Pang and Chan [18]), penalty and proximal point

<sup>1</sup> All vectors are column vectors; <sup>T</sup> denotes transposition. Thus:

$$F(x) = \begin{pmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{pmatrix} \quad \text{and} \quad F'(x) = (\nabla F_1(x), \dots, \nabla F_n(x)).$$

If  $F(x) = Ax$ , then  $F'(x) = A^T$ .

(‘augmented Lagrangian’) algorithms (Auslender [1], Rockafellar [20]), cutting-plane methods (Zuhovickii et al. [22], Nguyen and Dupuis [17]), etc. However, in contrast, with descent algorithms of convex optimization, the convergence of these algorithms is not related to the decrease of an objective function. Furthermore, some methods require that  $F$  be strongly monotone, or at least invertible in a neighborhood of a solution, which may fail to hold in a variety of applications.

We now introduce a merit function for VIP. An algorithm producing a sequence of iterates minimizing this merit function in a monotone fashion will be presented in Section 4.

**Definition 2.** The GAP FUNCTION (see Hearn [9]) of a VIP is defined as:

$$g(x) \triangleq \max_{y \in \Phi} (x - y)^T F(x) = (x - \tilde{y})^T F(x), \quad \text{for any } \tilde{y} \text{ in } \Gamma(x). \tag{3}$$

It is clear that  $g(x) = 0$  if and only if  $x$  is a solution to VIP. In economic equilibrium theory, the gap function corresponds to the excess demand function, and is a better measure of proximity to an equilibrium than the quasi-welfare function  $\int_0^x F(t) dt$  (when the latter is unambiguously defined), since it is directly related to the perception of the market structure by the economic agents.

In a traffic equilibrium framework, the gap function measures the difference between actual (perceived) travel costs and minimal (shortest path) costs.

### 3. Properties of the gap function

Let us first note that  $g$  is continuous but in general not differentiable (when  $\Gamma(x)$  is not single-valued) nor quasiconvex.

Let

$$\Phi = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + x_2 \leq 1, x_1, x_2 \geq 0 \right\} \quad \text{and} \quad F \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 + 4x_1x_2 \\ x_2^2 + 4x_1x_2 \end{pmatrix}. \tag{4}$$

$F$  is monotone on  $\Phi$  since

$$F' = \begin{pmatrix} 2x_1 + 4x_2 & 4x_1 \\ 4x_2 & 2x_2 + 4x_1 \end{pmatrix}$$

is positive semidefinite on  $\Phi$ . Also:

$$g(x) = (x - 0)^T F(x) \quad \text{and} \quad g \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{5}{4} > \max \left[ g \begin{pmatrix} 1 \\ 0 \end{pmatrix}, g \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = 1,$$

implying that  $g$  is NOT quasiconvex.

Useful differentiability properties of  $g$  are summarized in the next proposition.

**Proposition 1.** 1.  $g$  is Lipschitzian on  $\Phi$ .

2.  $\partial g(x) = \text{Co}_{y \in \Gamma(x)} [F(x) + F'(x)(x - y)]$

3.  $g$  is differentiable at  $x$  if  $\Gamma(x)$  is a singleton.

**Proof.** See Clarke [3] or Rockafellar [19].  $\square$

**Proposition 2.** If  $g(x)$  is positive and  $\Gamma(x) = \{y\}$  then  $d \triangleq y - x$  is a feasible descent direction for  $g$  at  $x$ .

**Proof.** We have, by Danskin's rule:

$$\begin{aligned} g'(x; d) &= \sup_{y \in \Gamma(x)} (y - x)^T (F(x) - F'(x)(y - x)) \\ &\leq -g(x) \quad \text{by monotonicity of } F \\ &< 0. \quad \square \end{aligned}$$

We now show that every stationary point of  $g$  is a zero of  $g$ .

**Proposition 3.** If  $g(x) > 0$  then there exists a point  $y^*$  in  $\Gamma(x)$  that satisfies:

$$g'(x; y^* - x) \leq -g(x). \tag{5}$$

**Proof.** Let  $y^*$  be solution to the linear VIP:

$$(y^* - y)^T F'^T(x)(y^* - x) \leq 0 \quad \text{for all } y \text{ in } \Gamma(x).$$

We have:

$$\begin{aligned} g'(x; y^* - x) &= \max_{h \in \partial g(x)} (y^* - x)^T h \\ &= \max_{y \in \Gamma(x)} (y^* - x)^T (F(x) + F'(x)(x - y)) \\ &= -g(x) + \max_{y \in \Gamma(x)} (y^* - x)^T F'(x)(x - y) \\ &\leq -g(x) - (y^* - x)^T F'(x)(y^* - x) \quad \text{by construction of } y^* \\ &\leq -g(x) \quad \text{by monotonicity of } F \\ &< 0. \quad \square \end{aligned}$$

When  $\Phi$  is a convex compact polyhedron, so is  $\Gamma(x)$ . Denote by  $\{u^1, \dots, u^m\}$  the extreme points of  $\Gamma(x)$ , and let  $z \triangleq \sum_{i=1}^m \lambda_i u^i$ . The directional derivative  $g'(x; z - x)$  is given as before, by the expression:

$$\begin{aligned} g'(x; z - x) &= -g(x) + \max_{u \in \Gamma(x)} (z - x)^T F'(x)(x - u) \\ &= -g(x) + \max_{i=1, \dots, m} \sum_{j=1}^m \lambda_j (u^j - x)^T F'(x)(x - u^i) \end{aligned}$$

since the maximum of the linear term  $(z - x)^T F'(x)(x - u)$  must be attained at some extreme point  $u^i$  of  $\Gamma(X)$ . Therefore:

$$g'(x; z - x) = -g(x) - \min_{i=1, \dots, m} (C\lambda)_i$$

where  $C$  is a positive semidefinite matrix with elements:

$$C_{ij} = (u^i - x)^T F'^T(x)(u^j - x).$$

The derivative  $g'(x; z - x)$  will achieve its minimum value at a point  $z^*$  which is solution to the mathematical program

$$\max_{z \in \Gamma(x)} g'(x; z - x)$$

which is equivalent to:

$$\max_{\lambda \geq 0} \min_{i=1, \dots, m} (C\lambda)_i,$$

or

$$\max_{\lambda, \gamma} \gamma$$

$$\text{s.t. } \sum_{i=1}^m \lambda_i = 1,$$

$$\lambda \geq 0,$$

(6)

$$C\lambda \geq \gamma \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Let  $(\lambda^*)$  be optimal for (6),  $z^* = \sum_{i=1}^m \lambda_i^* u^i$  and  $y^*$  as defined in the proof of Proposition 3. Then we have:  $g'(x; z^* - x) = -g(x) - \gamma^* \leq g'(x; y^* - x)$  where  $\gamma^*$  is nonnegative as a consequence of Proposition 3.<sup>2</sup> Notwithstanding direction norms, the descent direction  $z^*$  is therefore, in some sense, preferable to the descent direction  $y^* - x$ .

As an illustration, consider the VIP with:

$$\Phi = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ in } \mathbb{R}^2 \mid x_2 - 2x_1 \leq 2, 2x_1 + x_2 \leq 2, x_1, x_2 \geq 0 \right\} \text{ and } F \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(see Figure 1).

The equilibrium point is the origin, and the extreme points of  $\Phi$  are:  $u^1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ ,  $u^2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $u^3 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ . At  $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  we get:  $\Gamma(x) = \{u^1, u^2\}$ ,  $g(x^0) = 4$ . The directions  $d^1 = u^1 - x^0$  and  $d^2 = u^2 - x^0$  are descent directions for the gap function. Minimization along  $d^1$  yields  $x^1 = \begin{pmatrix} -7/10 \\ 6/10 \end{pmatrix}$ ,  $g(x^1) = \frac{31}{20}$  and  $\Gamma(x^1) = \{u^2\}$ . At  $x^1$ , minimization

<sup>2</sup> From the positive semidefiniteness of  $C$ , one can also directly deduce, using Tucker's theorem of the alternative (see [13]), that the system  $C^T \lambda \geq 0$  ( $\lambda \neq 0$ ) has a nonnegative solution.

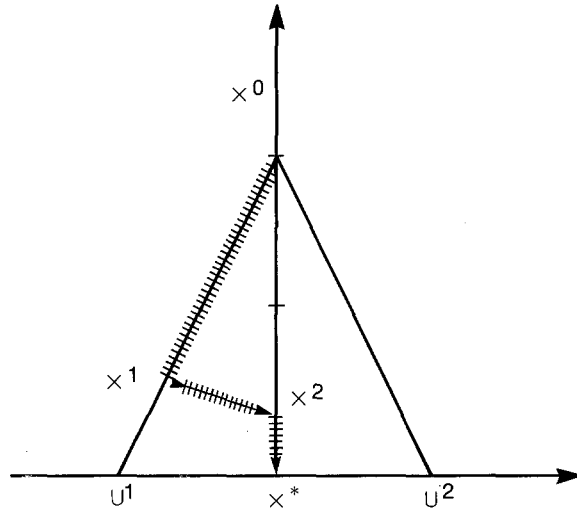


Fig. 1.

in the direction  $u^2 - x^1$  yields  $x^2 = (\frac{0}{6/17})$ ,  $g(x^2) = \frac{36}{289}$  and  $\Gamma(x^2) = \{u^1, u^2\}$ . At this point neither  $u^1 - x^2$  or  $u^2 - x^2$  are descent directions for  $g$ . The solution to (6) gives  $\lambda^* = (\frac{1/2}{0})$  and the descent direction  $(\frac{0}{0}) - x^2$ , which leads directly to the solution.

**4. A descent algorithm for the gap function**

In this section we propose a solution algorithm for the VIP, based on a descent direction obtained by solving LP programs of the form (6). The extremal directions  $u^i - x$  will play a role similar to the role of subgradients in convex nondifferentiable optimization. Since  $\Gamma(x)$  cannot in general be known explicitly, the algorithm will progressively refine a polyhedral approximation  $\bar{\Gamma}(x)$  to  $\Gamma(x)$ , in a spirit very similar to the  $\epsilon$ -subgradient algorithm described in [14], until a suitable descent direction for the gap function is found by solving (5) with  $\Gamma(x)$  replaced by  $\bar{\Gamma}(x)$ .

**Algorithm A.** Let  $x \in \Phi$  and  $\epsilon$  a small positive constant.

1. If  $g(x) < \epsilon$  STOP.  
 $m \leftarrow 1$ .  
 $C \leftarrow \emptyset$ .  
 Let  $y^1 \in \Gamma(x)$ .
2.  $C_{im} \leftarrow (y^i - x)^T F'^T(x)(y^m - x)$ ,  $i = 1, \dots, m$ .  
 $C_{mi} \leftarrow (y^m - x)^T F'^T(x)(y^i - x)$ ,  $i = 1, \dots, m$ .

$C \leftarrow$	$C$	$C_{1m}$
		$\vdots$
	$C_{m-1,m}$	$C_{mm}$
	$C_{m1} \quad \dots \quad C_{m,m-1}$	

3. Solve:

Max  $\gamma$

$$(C\lambda)_i \geq \gamma, \quad i = 1, \dots, m,$$

$$\sum_{i=1}^m \lambda_i = 1,$$

$$\lambda \geq 0,$$

whose solution is  $\bar{\lambda}, \bar{\gamma}$ .

$$\text{Let } \bar{y} = \sum_{i=1}^m \bar{\lambda}_i y^i.$$

4.  $\bar{x} \leftarrow x + \alpha g(x)(\bar{y} - x)$ .

$$m \leftarrow m + 1.$$

Let  $y^m \in \Gamma(\bar{x})$ .

5. If  $g(\bar{x}) > (1 - (\alpha/12)g(x))g(x)$  then go to 2

else let  $\theta^* \in \operatorname{argmin}_{0 \leq \theta \leq 1} g(x + \theta(\bar{y} - x))$ .

$$x \leftarrow x + \theta^*(\bar{y} - x).$$

go to 1.

Convergence of algorithm A depends on the existence of a positive constant  $\alpha$ , determined in such a way that the 'then' branch at step 5 cannot be taken infinitely often, at a given iteration.

In the following, we will assume that  $F'$  is Lipschitzian on  $\Phi$ , with Lipschitz constant  $M$ . Also, let  $L$  be the Lipschitz constant of  $F$ , and  $D$  be the diameter of  $\Phi$ , i.e.:

$$\|x - y\| \leq D \quad \text{for all } x, y \text{ in } \Phi, \tag{7a}$$

$$\|F(x) - F(y)\| \leq L\|x - y\| \quad \text{for all } x, y \text{ in } \Phi, \tag{7b}$$

$$\|F'^T(x) - F'^T(y)\| \leq M\|x - y\| \quad \text{for all } x, y \text{ in } \Phi. \tag{7c}$$

We also define:

$$L' \triangleq \sup_{x \in \Phi} \|F(x)\| < \infty, \tag{8a}$$

$$M' \triangleq \sup_{x \in \Phi} \|F'(x)\| < \infty, \tag{8b}$$

$$S \triangleq \sup_{x \in \Phi} g(x) < \infty. \tag{8c}$$

**Lemma 1.** Let  $x, \bar{x} \in \Phi, y \in \Gamma(x)$  and  $\bar{y} \in \Gamma(\bar{x})$ . Then there exists a positive constant  $\varepsilon_1$ , such that:

$$\|\bar{x} - x\| \leq \varepsilon_1 g(x) \Rightarrow (\bar{y} - x)^T F(x) \leq -\frac{1}{2}g(x) \quad \text{or} \quad g(\bar{x}) \leq \frac{2}{3}g(x). \tag{9}$$

**Proof.** Set  $\varepsilon_1 = 1/6(DL + L')$ . We have

$$\begin{aligned} (\bar{y} - x)^T F(x) &\leq (\bar{y} - \bar{x})^T F(x) + \|\bar{x} - x\| \cdot \|F(x)\| \\ &\leq (\bar{y} - \bar{x})^T F(\bar{x}) + \|\bar{y} - \bar{x}\| \cdot \|F(x) - F(\bar{x})\| + \|\bar{x} - x\| \cdot \|F(x)\| \\ &\leq -g(\bar{x}) + DL\varepsilon_1 g(x) + \varepsilon_1 g(x)L'. \end{aligned}$$

If  $g(\bar{x}) \geq \frac{2}{3}g(x)$  then

$$(\bar{y} - x)^T F(x) \leq -\frac{2}{3}g(x) + \frac{1}{6}g(x) = -\frac{1}{2}g(x). \quad \square$$

**Lemma 2.** *There exists a positive constant  $\varepsilon_2$  such that, whenever  $\alpha \leq \varepsilon_2$ , then ‘then’ branch at step 5 of algorithm A cannot be taken infinitely often, at a given iteration.*

**Proof.** Set

$$\varepsilon_2 = \text{Min} \left\{ \frac{4}{S}, \frac{1}{12D^2L}, \frac{1}{3D^2(2M + M' + DM)}, \frac{\varepsilon_1}{D} \right\}. \tag{10}$$

Suppose the result does not hold. Since  $\Phi$  is bounded, there must exist indices  $m$  and  $p$  ( $p \leq m$ ) such that:

$$\|y^m - y^p\| \leq \frac{\alpha DMg(x)}{M'}. \tag{11}$$

Let  $x, \bar{y}, \bar{x} = x + \alpha g(x)d$ ,  $d = \bar{y} - x = \sum_{i=1}^{m-1} \bar{\lambda}_i y^i - x$  be as in algorithm A (step 5),  $\tilde{x}$  be any point on the line segment from  $x$  to  $\bar{x}$  ( $\tilde{x} \neq x$ ) and  $\tilde{y}$  be any point in  $\Gamma(\tilde{x})$ . We have:

$$\tilde{x} = x + \tau \alpha g(x)d \quad \text{for some } \tau \in (0, 1],$$

and

$$\bar{x} = \tilde{x} + (1 - \tau) \alpha g(x)d.$$

We can write:

$$\begin{aligned} 0 &\geq (y^m - \tilde{y})^T F(\bar{x}) \quad \text{by definition of } y^m \\ &= (y^m - \tilde{y})^T [F(\tilde{x}) + F'^T(\tilde{x} + t(\bar{x} - \tilde{x}))(\bar{x} - \tilde{x})] \end{aligned}$$

for some  $t \in [0, 1]$  by the mean value theorem.

This implies:

$$\begin{aligned} 0 &\geq (y^m - \tilde{y})^T [F(\tilde{x}) + F'^T(x)(\bar{x} - \tilde{x})] - \|F'^T(\tilde{x} + t(\bar{x} - \tilde{x})) - F'^T(x)\| \cdot \|\bar{x} - \tilde{x}\| \\ &\geq (y^m - \tilde{y})^T [F(\tilde{x}) + F'^T(x)(\bar{x} - \tilde{x})] - M(\|\tilde{x} - x\| + t\|\bar{x} - \tilde{x}\|)(1 - \tau)\alpha g(x)D \\ &\geq (y^m - \tilde{y})^T [F(\tilde{x}) + F'^T(x)(\bar{x} - \tilde{x})] - M(1 - \tau)(\alpha g(x))^2 D^2 \end{aligned}$$



implying

$$\begin{aligned} (1-\tau)D^2M(\alpha g(x))^2 &\geq (y^m - \tilde{y})^T[F(\tilde{x}) + F'^T(x)(\bar{x} - \tilde{x})] \\ &\geq (y^m - \tilde{y})^T F'^T(x)(\bar{x} - \tilde{x}) \end{aligned}$$

since  $(y^m - \tilde{y})^T F(\tilde{x}) \geq 0$  by definition of  $\tilde{y}$ . Thus

$$(1-\tau)D^2M(\alpha g(x))^2 \geq (1-\tau)\alpha g(x)(y^m - \tilde{y})^T F'^T(x)d$$

and

$$\begin{aligned} D^2M\alpha g(x) &\geq (y^m - \tilde{y})^T F'^T(x)d = [(y^m - y^p) + (y^p - x) + (x - \tilde{y})]^T F'^T(x)d \\ &\geq -\frac{\alpha DMg(x)}{M'}M'D + (x - \tilde{y})^T F'^T(x)d \end{aligned}$$

from (11) and the fact that  $(y^p - x)^T F'^T(x)d \geq 0$  by construction of the direction  $d$ . Therefore

$$(x - \tilde{y})^T F'^T(x)d \leq 2\alpha D^2Mg(x). \tag{12}$$

Now

$$\begin{aligned} (\tilde{x} - \tilde{y})^T F'^T(\tilde{x})d &= (x - \tilde{y})^T F'^T(x)d + (\tilde{x} - x)^T F'^T(x)d \\ &\quad + (\tilde{x} - \tilde{y})^T (F'^T(\tilde{x}) - F'^T(x))d \\ &\leq 2\alpha D^2Mg(x) + D^2\alpha g(x)M' + MD^3\alpha g(x) \quad \text{by (12)} \\ &= (2D^2M + D^2M' + D^3M)\alpha g(x), \end{aligned}$$

implying

$$(\tilde{x} - \tilde{y})^T F'^T(\tilde{x})d \leq g(x)/3 \quad \text{since } \alpha \leq \epsilon_2.^3 \tag{13}$$

Also:

$$\begin{aligned} g'(\tilde{x}; d) &= \max_{y \in \Gamma(\tilde{x})} d^T[F(\tilde{x}) + F'(\tilde{x})(\tilde{x} - y)] \\ &\leq \frac{g(x)}{3} + (\bar{y} - x)^T F(\tilde{x}) \quad \text{by (13)} \\ &\leq \frac{g(x)}{3} + (\bar{y} - x)^T F(x) + \alpha D^2Lg(x) \\ &\leq \frac{g(x)}{3} + (\bar{y} - x)^T F(x) + \frac{g(x)}{12} \quad \text{by (10)}. \end{aligned}$$

Now, if  $g(\bar{x}) \leq \frac{2}{3}g(x)$ , we get:

$$g(\bar{x}) \leq \frac{2}{3}g(x) \leq \left(1 - \frac{1}{3} \cdot \frac{\alpha}{4/S}\right)g(x) \leq \left(1 - \frac{\alpha g(x)}{12}\right)g(x)$$

<sup>3</sup> This relation holds for any  $\tilde{y}$  in  $\Gamma(\tilde{x})$ .

which is the required contradiction. Otherwise, from Lemma 1,

$$(\bar{y} - x)^T F(x) \leq -\frac{1}{2}g(x)$$

since  $\|\bar{x} - x\| = \alpha g(x) \|d\| < \varepsilon_1 / D \cdot D = \varepsilon_1$ .

Therefore

$$g'(\tilde{x}; d) \leq \frac{5g(x)}{12} - \frac{1}{2}g(x) = -\frac{g(x)}{12}$$

Thus

$$\begin{aligned} g(\bar{x}) - g(x) &= \alpha g(x) g'(\tilde{x}; d) \quad \text{for some } \tilde{x} \in [x, \bar{x}], \\ &\quad \text{by the mean value theorem} \\ &\leq -\alpha \frac{(g(x))^2}{12} \end{aligned}$$

or:

$$g(\bar{x}) \leq \left(1 - \frac{\alpha g(x)}{12}\right) g(x)$$

which contradicts the assumption at the beginning of the proof.  $\square$

**Proposition 4.** Let  $\{x^k\}_k$  be a sequence generated by algorithm A. Then

$$\lim_{k \rightarrow \infty} g(x^k) = 0.$$

**Proof.** From Lemma 2 we get

$$g(x^{k+1}) \leq \left(1 - \frac{\alpha}{12} g(x^k)\right) g(x^k).$$

Taking limits on both sides, we have

$$g^\infty \triangleq \lim_{k \rightarrow \infty} g(x^k) \leq \left(1 - \alpha \frac{g^\infty}{12}\right) g^\infty \quad \text{and} \quad g^\infty = 0. \quad \square$$

**Remark.** When  $\Phi$  is polyhedral, and if the points  $y^i$  are always taken as extreme points of  $\Phi$ , then the maximal number of times step 5 has to be repeated at a given iteration of algorithm A can be uniformly bounded by the number of extreme points of  $\Phi$ .

## 5. Application to the traffic assignment problem

It is well known [4, 21] that the traffic assignment problem (TAP in short) can be formulated as a VIP, with  $\Phi$  representing the convex polyhedron of multicom-

modify link-flow vectors

$$x^{(k)}, k = 1, \dots, K^4 \quad \text{and} \quad F_a(x) \equiv F_a\left(\sum_{k=1}^K x^{(k)}\right)$$

is the unit cost along link  $a$ , possibly dependent on link flows over other links of the network.

In Marcotte [15] it has been observed that it occurs frequently than *any* direction of the form  $d = y - x$  with  $y \in \Gamma(x)$  is a descent direction for  $g$  at  $x$ , and the iteration<sup>5</sup>

$$x \leftarrow x + \lambda d \quad \text{with } \lambda = 1 \quad \text{if } (y - x)^T F(y) \leq 0,$$

$$(x + \lambda d - y)^T F(x + \lambda d) = 0 \quad \text{otherwise}$$

was gap decreasing, at least in the first iterations. Such a scheme can be exploited in the first iterations. When one gets closer to a solution, it is likely that many extreme points  $y^i$  generated previously by algorithm  $A$  will satisfy the relation

$$(y^i - x)^T F(x) \leq -\beta g(x)$$

for some positive constant  $\beta$ . This suggests that these points should be taken into account when looking for a descent direction, yielding the following modified algorithm.

**Algorithm B.** Let  $x \in \Phi$ ,  $B \leftarrow \emptyset$ .

1. If  $g(x) \leq \varepsilon$  STOP.  
 $m \leftarrow 1$ .  
 Let  $y^1 \in \Gamma(x)$ .
2.  $C_{im} \leftarrow (y^i - x)^T F'^T(x)(y^m - x)$ ,  $i = 1, \dots, m$ .  
 $C_{mi} \leftarrow (y^m - x)^T F'^T(x)(y^i - x)$ ,  $i = 1, \dots, m$ .

$C \leftarrow$	$C$	$C_{1m}$ $\vdots$ $C_{m-1,m}$
	$C_{m1} \quad \dots \quad C_{m,m-1}$	$C_{mm}$

3. Solve:

$$\text{Max } \gamma$$

$$(C\lambda)_i \geq \gamma, \quad i = 1, \dots, m,$$

$$\sum_{i=1}^m \lambda_i = 1,$$

$$\lambda \geq 0,$$

whose solution is  $\bar{\lambda}, \bar{\gamma}$ .

$$\text{Let } \bar{y} = \sum_{i=1}^m \bar{\lambda}_i y^i.$$

<sup>4</sup>  $\Phi$  can be assumed compact without loss of generality. Indeed if  $F_a(x)$ ,  $a = 1, \dots, n$ , is positive for every positive  $x$ , it can be shown that only loopless paths need be considered (see Marcotte [15]).

<sup>5</sup> This scheme actually consists in solving the VIP restricted to be the line segment  $[x, y = x + d]$ .

4.  $\bar{x} \leftarrow x + \alpha' g(x)(\bar{y} - x)$ .  
 $m \leftarrow m + 1$ .  
 Let  $y^m \in \Gamma(\bar{x})$ .
5. **If**  $g(\bar{x}) > (1 - (\alpha'/c)g(x))g(x)$  **then go to 2**  
**else** let  $\theta^* \in \arg \min_{\theta \in [0,1]} g(x + \theta(\bar{y} - x))$ .  
 $x \leftarrow x + \theta^*(\bar{y} - x)$ .  
 Let  $\hat{y} \in \Gamma(x)$ .  
**If**  $g(x) \leq \varepsilon$  **STOP**.

Let  $I = \{l_1, \dots, l_r\}$  the set of indices such that:

$$(y^i - x)^T F(x) \leq -\beta g(x) \tag{14}$$

Set

$$C_{ij} \leftarrow C_{i_l, j_l}, \quad i = 1, \dots, r, \quad j = 1, \dots, r.$$

$$C_{ij} \leftarrow 0, \quad i \geq r + 1, \quad j \geq r + 1.$$

$$m \leftarrow r + 1.$$

Let  $y^m = \hat{y}$  and return to 2.  $\square$

Only at step 5 do algorithms A and B differ. In algorithm B, all extreme points previously generated that satisfy (14) are retained in order to compute a descent direction. Convergence can be established along the lines of lemmas 1 and 2 and proposition 3, possibly with constants  $\alpha'$  and  $c$  different, respectively from the constants  $\alpha$  and 12 of algorithm A.

A somewhat similar idea has been proposed by Fukushima [7] to solve the TAP with separable cost function. At each step of the algorithm, one computes the directions  $d^1 = y^k - x^k$  with  $y^k \in \Gamma(x^k)$  and

$$d^2 = \sum_{i=1}^k \mu_i^k y^i - x^k$$

where  $y^i \in \Gamma(x^i)$ ,  $i = 1, \dots, k$  and  $\{\mu_i^k\}_i$  is an arbitrary set of positive weights summing to one, and satisfying the condition  $\mu_k^k \geq \dots \geq \mu_1^k$ .

At iteration  $k$ ,  $d^2$  is chosen in preference to the Frank-Wolfe direction  $d^1$  whenever:

$$\frac{d^{2T}GF(x^k)}{\|d^2\|} \leq \frac{d^{1T}F(x^k)}{\|d^1\|}.$$

In algorithm B, we propose a systematic approach for choosing the weights  $\{\mu_i^k\}_i$  in an optimal rather than heuristic manner.

Finally, we remark that when  $F$  is the gradient of some function  $f$  which is convex on  $\Phi$ , then the directions obtained from algorithms A and B are descent direction for the gap function  $g$  as well as for the 'natural' objective  $f$ .

## Acknowledgments

I wish to thank Jean-Louis Goffin and the referee for positive criticism of the first version of this paper.

## References

- [1] A. Auslender, *Optimisation: Méthodes numériques* (Masson, Paris, 1976).
- [2] D. Bertsekas and E.M. Gafni, "Projection methods for variational inequalities with application to the traffic assignment problem", *Mathematical Programming Study* 17 (1982) 139-159.
- [3] F.H. Clarke, "Generalized gradients and applications", *Transactions of the American Mathematical Society* 205 (1975) 247-262.
- [4] S.C. Dafermos, "Traffic equilibrium and variational inequalities", *Transportation Science* 14 (1980) 42-54.
- [5] S.C. Dafermos, "An iterative scheme for variational inequalities", *Mathematical Programming* 26 (1983) 40-47.
- [6] J.M. Danskin, "The Theory of max-min, with applications", *SIAM Journal of Applied Mathematics* 14 (1966) 641-664.
- [7] M. Fukushima, "A modified Frank-Wolfe algorithm for solving the traffic assignment problem", *Transportation Research B* 18B (1984) 169-177.
- [8] R. Glowinski, J.L. Lions and R. Trémolières, *Analyse numérique des inéquations variationnelles* (Dunod, Paris, 1979).
- [9] D.W. Hearn, "The gap function of a convex program", *Operations Research Letters* 1 (1981) 67-71.
- [10] N.H. Josephy, "Newton's method for generalized equations", Technical Report #1965, Mathematics Research Center, University of Wisconsin-Madison, Madison, WI, 1979.
- [11] N.H. Josephy, "Quasi-Newton methods for generalized equations", Technical Report #1966, Mathematics Research Center, University of Wisconsin-Madison, Madison, WI, 1979.
- [12] S. Kakutani, "A generalization of Brouwer's fixed point theorem", *Duke Mathematics Journal* 8 (1941) 457-459.
- [13] H.W. Kuhn and A.W. Tucker, "Linear inequalities and related systems", *Annals of Mathematics Studies* 38 (1956) 3-18.
- [14] C. Lemaréchal, "Nonsmooth optimization and descent methods", Research Report RR-78-4, IIASA (Laxenburg, Austria, 1978).
- [15] P. Marcotte, "Design optimal d'un réseau de transport en présence d'effets de congestion", Ph.D. Thesis, Département d'informatique et de recherche opérationnelle, Université de Montréal (Montréal, Canada, 1981).
- [16] R. Mifflin, "A modification and an extension of Lemaréchal's algorithm for nonsmooth minimization", *Mathematical Programming Study* 17 (1982) 77-90.
- [17] S. Nguyen and C. Dupuis, "An efficient method for computing traffic equilibria in networks with asymmetric transportation costs", *Transportation Science* 18 (1984) 185-202.
- [18] J.S. Pang and D. Chan, "Iterative methods for variational and complementarity problems", *Mathematical Programming* 24 (1982) 284-313.
- [19] R.T. Rockafellar, *La théorie des sous-gradients et ses applications à l'optimisation* (Presses de l'Université de Montréal, Montreal, Canada, 1981).
- [20] R.T. Rockafellar, "Monotone operators and augmented Lagrangian methods in nonlinear programming", in: O. Mangasarian, R.R. Meyer, and S.M. Robinson, eds., *Nonlinear Programming* 3 (Academic Press, New York, 1978).
- [21] M.J. Smith, "The existence, uniqueness and stability of traffic equilibrium", *Transportation Research B* 13B (1979) 295-304.
- [22] S. Zuhovickii, R.A. Polyak and M.E. Primak, "Two methods of search of equilibrium of  $N$ -person concave games", *Soviet Mathematics-Doklady* 10 (1969) 279-282.