# THE GENERAL TRAFFIC ASSIGNMENT PROBLEM: A PROXIMAL POINT METHOD FOR EQUILIBRIUM COMPUTATION WITH APPLICATIONS TO THE DEMAND ADJUSTMENT PROBLEM.

Lídia Montero, Esteve Codina
Statistics and Operations Research Department
Polytechnic University of Catalonia, Spain
C/ Pau Gargallo, 5
Barcelona 08028, Spain

Lídia Montero
C/ Pau Gargallo, 5
Barcelona 08028, Spain
Tel. +34 93 401 1038
Fax. +34 93 401 5855
lidia.montero@upc.es

April 30, 2003

ABSTRACT

An adaptation of the proximal algorithm for the traffic assignment problem under a user equilibrium formulation for a general asymmetric traffic network is presented in this paper. It follows the recently published results of Pennanen regarding convergence under non monotonicity. As it is well known the problem can be formulated as a variational inequality and the algorithmic solutions developed up to date guarantee convergence only under too restrictive conditions which are difficult to appear in practice. In this paper it is also discussed the possibility of including the algorithm on a demand adjustment problem formulated as a bilevel program with lower level traffic equilibrium constraints expressed as a variational inequality.



**Keywords:** Traffic assignment, variational inequality, proximal point algorithm, bilevel programming, demand adjustment.

**AMS Classification:** 90C31, 90C59, 90C99, 90B10, 90B30

# 1. Introduction

The traffic assignment problem is one of the core problems in the transportation planning process. Given a transportation network, and certain assumptions of the route choice behaviour of the tripmakers, the traffic assignment problem consists on assigning traffic onto the network, so as to fulfill demand for transportation and to minimize some merit function, related to the behavioural assumption made. Several assumptions on the route choice behaviour have been made. However, the most natural assumption is that each driver chooses the shortest perceived route to his/her destination under prevailing traffic conditions. The result from a generalized decision like that made by all the travelers yields a situation in which no driver can reduce his/her journey time by choosing another route. This is the *user optimal criteria* for route choice. Wardrop [29] was the first to state this route choice criteria.

The asymmetric traffic assignment problem (i.e. when there are link flow interactions on the link travel costs) was first formulated as a variational inequality problem by Smith [28] and a number of algorithms emerged to solve the resulting formulation and basically the convergence of all them required at least the strong monotonicity of the link travel costs. The proximal point method of Martinet [20] and its application by Rockafellar in order to develop the method of proximal multipliers has been the subject of study of many researchers that apply it on a variety of problems. The recent work of Pennanen [24] has shown that these methods present local convergence under some weaker conditions that the ones stated on previous works.

On the other hand, updating an obsolete origin and destination matrix of trips within a transportation area, using available information, is a very common problem in traffic and transportation planning. For traffic networks, a source of information of a relatively moderate cost are traffic counts on a subset of links of the traffic network. The paper is organized as follows. Following these paragraphs, in Subsection 1.1 the basic formulations and notation used to described the user equilibrium in traffic networks is presented. In Section 2, the variational inequality formulation is developed, reproducing already known results mainly due to Smith [28] for existence and uniqueness of solutions. In Section 3 the convergence conditions of many algorithms for solving the variational inequality formulation of the traffic assignment problem are revisited (linearization methods, projection methods and diagonalization methods and simplicial decomposition methods). Section 4 describes recently developed convergence conditions for the proximal point algorithm for inclusions due to Pennanen [24]. His results basically state that the convergence can be achieved even if the application map of the inclusion is hypomonotone, while up to date conditions ensured convergence only under maximal monotonicity conditions. In Section 5, the application of the proximal point algorithm for the demand adjustment problem is shown; where the problem is formulated as a bilevel program in which the lower level (the traffic assignment problem) takes the form of a variational inequality in which the link travel costs are strongly monotone.

## 1.1 Formulations

The development of general cost functions was due to the unrealistic assumption that the travel time on a link is independent of the flow on other links. One just has to imagine traffic near a turning intersection or a narrow two-way street to realize this. One reason for the late development of models incorporating more general cost functions has been the difficulty of catching a realistic cost relationship.

Let us consider the Wardrop's user equilibrium, as it is discussed by authors as Smith and Heydecker in references [28] and Heydecker [18]:

> *A traffic distribution is a Wardrop equilibrium when no driver has a less costly alternative route.*

Assume a transportation network $\mathcal{G}$ with a single mode of transit and fixed demand $\mathcal{D}$. Consider a specific origin destination pair $(i, j)$ and let $\mathcal{R}_{ij}$ be the set of available paths joining OD pair $(i, j)$. The route flow vector $\mathbf{h}$ induces flows $v_a$ on each link $a \in \mathcal{A}$ given by the expression,

$$v_a = \sum_{(i,j)\in\mathcal{C}} \sum_{r\in\mathcal{R}_{ij}} \delta_{ar}\, h_{ijr} \tag{1}$$

where $\delta_{ar} = 1$ if link $a$ belongs to path r; and $\delta_{ar} = 0$ otherwise. In words, $v_a$ is the sum of flows $h_{ijr}$ on all paths $r$, over all OD pairs $(i, j)$.

Let $\mathbf{v}$ denote the vector of arc flows and let $\mathbf{\Delta}$ denote the link path incident matrix. Then in vector form, link flows and path flows are related by the following expression

$$\mathbf{v} = \mathbf{\Delta}\,\mathbf{h}$$

Let each unit of flow on link $a$ incur a travel cost $c_a(\mathbf{v})$ which depends upon the vector $\mathbf{v}$ of link flows in the network. In the separable traffic assignment problem, the cost on a link depends solely upon the flow $v_a$ on that link, in which case it usually increases with increased levels of the link flow.

If we assume that the cost on any path of the network, as a function of path flows, is the sum of travel costs on the links of that path, then

$$C_{ijr} = \sum_{a\in\mathcal{A}} \delta_{ar}\, c_a(v_a)$$

We refer to this form of route costs as an *additive model*. Stating this relationship more compactly in vector form, we obtain

$$\mathbf{C}(\mathbf{h}) = \mathbf{\Delta}^\top\,\mathbf{c}(\mathbf{v}),$$

where $^\top$ denotes transposition.

In this expression, $\mathbf{C}(\mathbf{h}) = (\ C_{ijr}(\mathbf{h})\ )$ is a vector-valued function specifying the travel costs on each path $r$ and $\mathbf{c}(\mathbf{v})$ is a vector valued function whose components specify the link travel costs. The route travel costs $C_{ijr}(\mathbf{h})$ on each route $r$ joining an OD pair $(i,j)$ defines the least travel cost $u_{ij}$ over all paths joining that OD pair. This is

$$u_{ij} \equiv \min_{r \in \mathcal{R}_{ij}} C_{ijr}(\mathbf{h})$$

These least travel costs certainly provide a reference point against which to measure any route's ability to attract trips. If different routes are used for a given OD pair $(i,j)$, the travel cost of the routes used is $u_{ij}$, and Wardrop's equilibrium may be rewritten as

$$\begin{aligned} h_{ijr} &> 0 \Rightarrow C_{ijr} = u_{ij} \\ h_{ijr} &= 0 \Rightarrow C_{ijr} \geq u_{ij} \end{aligned}$$

or equivalently as

$$h_{ijr} \cdot (C_{ijr} - u_{ij}) = 0 \quad \forall\, r \in \mathcal{R}_{ij} \ ; \forall\, (i,j) \in \mathcal{C} \tag{2}$$

$$C_{ijr} - u_{ij} \geq 0 \quad \forall\, r \in \mathcal{R}_{ij} \ ; \forall\, (i,j) \in \mathcal{C} \tag{3}$$

$$\sum_{r \in \mathcal{R}_{ij}} h_{ijr} = d_{ij} \quad \forall\, (i,j) \in \mathcal{C} \tag{4}$$

$$h_{ijr} \geq 0 \quad \forall\, r \in \mathcal{R}_{ij} \ ; \forall\, (i,j) \in \mathcal{C} \tag{5}$$

$$u_{ij} \geq 0 \quad \forall\, (i,j) \in \mathcal{C} \tag{6}$$

## 2. Variational Inequality Formulation [TAP-VI]

Wardrop conditions for the general traffic assignment model can be reformulated into three main types of equivalent problems: variational inequality formulations, nonlinear complementary formulations and fixed point formulations. The variational inequality formulations became very popular, due to the large literature devoted to **VIP** , compacteness for writing and flexibility in choosing the ground set $\mathbf{X}$.

**Theorem 1.1** *Wardrop conditions for user equilibrium are equivalent to the following variational inequality formulation, defining the general traffic assignment problem* [**TAP-VI**] *,*

$$\textit{Find}\ \boldsymbol{h} \in \boldsymbol{\mathcal{H}}\ s.\ t.\ \boldsymbol{C}(\boldsymbol{h})^{\top}(\boldsymbol{h} - \boldsymbol{h}) \geq 0\ \ \forall\, \boldsymbol{h} \in \boldsymbol{\mathcal{H}} \tag{7}$$

*and*

$$\mathcal{H} \overset{\text{def}}{=} \{\, \mathbf{h} \mid \sum_{r \in \mathcal{R}_{ij}} h_{ijr} = d_{ij},\ h_{ijr} \geq 0,\ \forall\, r \in \mathcal{R}_{ij}\ , \forall\, (i,j) \in \mathcal{C}\ \} \tag{8}$$

It should be noted that the variational inequality formulation **[TAP-VI]** is independent of the form of the cost functions $C_{ijr}(\mathbf{h})$, this is, the travel cost on a route needs not to be additive. The additivity property is assumed to hold when developing algorithmic approaches. There are two main reasons for this:

1. Under the non additive assumption there does not exist a transformation of the above arc route formulation of **[TAP-VI]** in terms of arc flows.
2. On working with large networks, as our case, the arc route formulations of traffic assignment models involves very large scale problems, and it has been preferred to remove the extra size due to the route flow model and concentrate efforts on solving the aggregated arc flow model.

We state the arc flow variational inequality formulation for the traffic assignment problem when travel costs on a route satisfy the additivity property. This formulation is studied below and serves as the basis for developing algorithms to compute equilibrium solutions to **[TAP-VI]**.

$$\text{Find } \mathbf{v}^* \in \mathcal{V} \; s.\,t.\; \mathbf{c}(\mathbf{v}^*)^\top \cdot (\mathbf{v} - \mathbf{v}^*) \;\geq\; 0 \;\; \forall\, \mathbf{v} \in \mathcal{V} \tag{9}$$

and

$$\mathcal{V} \stackrel{\text{def}}{=} \{\, \mathbf{v} \mid v_a \;=\; \sum_{(i,j)\in\mathcal{C}} \;\; \sum_{r\in\mathcal{R}_{ij}} \;\; \delta_{ar}\, h_{ijr} \;\; \forall\, a \in \mathcal{A}\;, \;\; \mathbf{h} \in \mathcal{H} \,\} \tag{10}$$

Smith [28] discusses the properties of equilibrium solutions and formulates the problem **[TAP-VI]** for both the arc node case and the arc route case for additive route costs. Dafermos develops an equivalent finite dimensional variational inequality formulation. Formulations for the elastic case have been proposed by Dafermos and Florian. Dafermos and Nagurney [11] developed an arc flow formulation, equivalent to the above model, and study the stability and sensitivity of equilibria, based on that formulation.

In this point we introduce the primal and dual gap functions $\mathbf{G_P}(\mathbf{x})$ and $\mathbf{G_D}(\mathbf{x})$ for **[TAP-VI]**.

$$\mathbf{G_P}(\mathbf{v}) \;=\; -\min_{w\in\mathcal{V}} \mathbf{c}(\mathbf{v})^\top \cdot (\mathbf{w} - \mathbf{v}) \tag{11}$$

$$\mathbf{G_D}(\mathbf{w}) \;=\; -\max_{v\in\mathcal{V}} \mathbf{c}(\mathbf{v})^\top \cdot (\mathbf{w} - \mathbf{v}) \tag{12}$$

The primal gap function may be interpreted in this context as the difference of total travel costs between the current flow $\mathbf{v}$ and the shortest route flow $\mathbf{z}$. Thus a positive gap function corresponds to a potential benefit for some travelers in adjusting their route choices. On the other hand, $\mathbf{G}(\mathbf{v}) = 0$ precisely when no traveler has an incentive to change route, that is, when the flow satisfies Wardrop conditions of equilibria. Viewing the gap function in terms of an error in Wardrop conditions, Hearn [16] suggests the use of this formulation even in the separable case, since the objective function of its mathematical formulation is rather artificial.

Algorithms based on primal gap function are given Hearn [16]; methods based on simplicial descomposition are given by Lawphongpanich and Hearn [19], and Pang and Yu [21]. Cutting plane methods are given by Nguyen and Dupuis.

Now, we present conditions for the existence and uniqueness of solutions for the variational inequality formulation **[TAP-VI]** . Although they are, in the traffic equilibrium context, essentially equivalent, the existence results presented differ. It seems that too strong assumptions have sometimes been put on the model when establishing the existence of a solution, mainly because these conditions serves to validate the proposed algorithms as well.

We recall that sufficient conditions for the existence of a solution to **VIP** , where based on:

- Boundedness of the feasible set.
- Sufficient monotonicity of the mapping defining the problem.

The feasible set in the traffic assignment problem is, in general, not bounded, due to the existence of cycles in the network. Anyway, we can ensure the existence of a solution to **[TAP]** , by restricting the cost functions $\mathbf{c}(\mathbf{v})$ or $C_{ijr}(\mathbf{h})$ to be strictly positive for all feasible flows. The implication of this is that an optimal solution cannot include cyclic flows and the variables can therefore be restricted as follows,

$$0 \leq v_a \leq \sum_{(i,j)\in\mathcal{C}} d_{ij}, \;\; \forall\, a \in \mathcal{A}$$

Lawphongpanich and Hearn [19] consider for **[TAP-VI]** how the original problem can be replaced by restricted problems over sets of simple or loop-free routes. In convergence results or existence and uniqueness results, boundness it is always assumed.

Existence and uniqueness results are extracted from Aashtiani and Magnanti [1] and Smith [27]. The graph $\mathcal{G}$ is assumed to be *strongly connected,* i.e., that there exists at least one route connecting each origin and destination.

The problem **[TAP-VI]** has an equilibrium solution $\mathbf{v}^{(*)}$ under one of the following conditions:

1. Under additive cost model and $\mathbf{c}(\mathbf{v})$ continuous, positive and $d_{ij} \geq 0$ (Aashtiani and Magnanti).
2. Under additive cost model and $\mathbf{c}(\mathbf{v})$ continuous, nonnegative and the feasible set closed and convex (Smith).

3. Under nonadditive cost model and $\mathbf{C}(\mathbf{h})$ continuous, positive and $d_{ij} \geq 0$ (Aashtiani and Magnanti).

Uniqueness results listed below proceed from the same sources mentioned before:

1. Under additive cost model and $\mathbf{c}(\mathbf{v})$ continuous, nonnegative, strictly monotone and the feasible set closed and convex, the equilibrium link flow solution $\mathbf{v}^{(*)}$ is unique (Smith.)

2. Under additive cost model and $\mathbf{c}(\mathbf{v})$ continuous, positive, strictly monotone and $d_{ij} \geq 0$, the link flow volumes $\mathbf{v}^{(*)}$ and the accessibility vector $\boldsymbol{u}$ are unique (Aashtiani and Magnanti.)

Observe that the results require that the vector $\mathbf{c}(\mathbf{v})$ of cost functions be strictly monotone in terms of link volumes $\mathbf{v}$, not in terms of path flows $\mathbf{h}$. Path flows need not to be unique, since two collections of path flows might correspond to the same link flows.

# 3. Convergence conditions for [TAP - VI] algorithms

The first methods applied to the general finite dimensional variational inequality problem were based on fixed point problem reformulations, but for large scale problems, these algorithms are impractical, due to the large memory requirements, and their failure to utilize problem structure. Efficient methods for **VIP** may be grouped into the following methodology classes:

- Linearization methods. They are based on iterative approximation of the mapping defining **VIP** by affine mappings at the current point.
- Diagonalization methods. Similar to the algorithms proposed for solving equations: Extensions of Jacobi, Gauss-Seidel and Newton methods.
- Simplicial Decomposition methods.
- Dual cutting plane methods. These methods apply to solving the maximization of the dual gap function, which it is equivalent to solving **[TAP-VI]**. Convergence results require strict monotonicity and the set of feasible arc flows to be a compact polyhedron.
- Gap descent Newton method. Essentially, the method generates a sequence such that each iterate is the solution to a variational inequality problem involving a linear approximation of the functional cost at the previous iterate, whose solution is used to define a descent direction to the gap function. Monotonicity and continuous differentiability of the functional cost are required for global convergence results.

## 3.1 Linearization Methods

Several variants are described in reports, but all of them are based on an iterative approximation of $\mathbf{f}(\mathbf{x})$ on the current iteration point $\mathbf{x}^{(t)}$, called $\mathbf{f}^{(t)}(\mathbf{x})$ and can be described as special cases of the following scheme:

> **Algorithm.** Given $\mathbf{x}^{(t)} \in X$, let $\mathbf{x}^{(t+1)}$ solve the variational inequality subproblem $\mathbf{VI}(\mathbf{f}^{(t)}, \mathbf{X})$, where $\mathbf{f}^{(t)}(\mathbf{x})$ is some mapping approximating the original $\mathbf{f}(\mathbf{x})$ at the point $\mathbf{x}^{(t)}$.

Pressumably, each subproblem $\mathbf{VI}(\mathbf{f}^{(t)}, \mathbf{X})$ is numerically easier to solve than the original problem $\mathbf{VI}(\mathbf{f}, \mathbf{X})$.

The method is classified as a *linear approximation method* if $\mathbf{f}^{(t)}(\mathbf{x})$ is of the form

$$\mathbf{f}(\mathbf{x}) \sim \mathbf{f}^{(t)}(\mathbf{x}) = \mathbf{c} + \mathbf{A}^{(t)}\mathbf{x}$$

where $\mathbf{A}^{(t)}$ is a constant matrix at iteration $t$ and $\mathbf{c} = \rho\mathbf{f}(\hat{\mathbf{x}}) - \mathbf{A}^{(\mathbf{t})}\hat{\mathbf{x}}$ for $\hat{x} \in X$ and $\rho > 0$.

Included in the family of linear approximation methods are:

**Newton method.** The mapping $\mathbf{f}(\mathbf{x})$ is assumed to be differentiable and $\mathbf{A}^{(t)} = \nabla\mathbf{f}(\mathbf{x}^{(t)})$. The linear approximation is in this case the first order Taylor expansion of $\mathbf{f}(\mathbf{x})$ around $\mathbf{x}^{(t)}$.
**Quasi-Newton methods.** $\mathbf{A}^{(t)}$ is taken to be an approximation to the Jacobian matrix $\nabla\mathbf{f}(\mathbf{x}^{(t)})$.
**Linear SOR methods.** Successive overrelaxation methods in which the matrix $\mathbf{A}^{(t)}$ is taken to be

$$\mathbf{A}^{(t)} = \begin{cases} \mathbf{L}^{(t)} + \mathbf{D}^{(t)}/\omega^* \\ \mathbf{U}^{(t)} + \mathbf{D}^{(t)}/\omega^* \end{cases} \quad 0 < \omega^* < 2$$

Assuming $\mathbf{L}^{(t)}$, $\mathbf{U}^{(t)}$ and $\mathbf{D}^{(t)}$ are respectively the strictly lower and upper triangular parts and diagonal of $\nabla\mathbf{f}(\mathbf{x}^{(t)})$. When $\omega^*$ is taken to be 1, the scheme becomes *the linearized Gauss-Seidel method*.
**Linear Jacobi method.** The matrix $\mathbf{A}^{(t)}$ is taken to be the diagonal of the Jacobian matrix at the current point, i.e., $\mathbf{A}^{(t)} = \mathrm{Diag}(\nabla\mathbf{f}(\mathbf{x}^{(t)}))$.
**Projection methods.** The matrix $\mathbf{A}^{(t)}$ is taken to be a fixed symmetric and positive definite matrix $\mathbf{G}$ equal for all iterations $t$.

Although a fast local convergence is ensured for Newton's method, the problem $\mathbf{VI}(\mathbf{f}^{(t)}, \mathbf{X})$ is not trivial, and the affine variational inequality defined is not equivalent to a mathematical program, since $\nabla\,\mathbf{f}(\mathbf{x})$ is, in general, assymetric. Quasi-Newton methods have been proposed with an approximation matrix chosen to be symmetric, to ensure that the variational inequality defined is equivalent to a mathematical programming program. SOR methods lead to decomposition of the original subproblem into independent problems.

The reason why methods in the last category are termed projection methods is due to the following geometrical interpretation of the iterates $\{\mathbf{x}^{(t)}\}$. Indeed, it is easy to show that if the set $\mathbf{X}$ is closed and convex and if $\mathbf{A}^{(t)} = \frac{1}{\gamma}\mathbf{G}$ is symmetric and positive definite, then the vector $\mathbf{x}^{(t+1)}$ that solves the subproblem $\mathbf{VI}(\mathbf{f}^{(t)}, \mathbf{X})$ is the projection of the point $\mathbf{x}^{(t)} - \gamma\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}^{(t)})$ onto the set $\mathbf{X}$ where the projection is defined with respect to the G-norm, i. e.,

$$\mathbf{x}^{(t+1)} = \mathbf{Pr}_{\mathbf{X}}^{\mathbf{G}} \left(\mathbf{x}^{(t)} - \gamma\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}^{(t)})\right)$$

where for a given vector $\mathbf{z}$, $\mathbf{Pr}_{\mathbf{X}}^{\mathbf{G}}(\mathbf{z})$ is the unique vector solving the mathematical programming program

$$\begin{gathered}\min \parallel \mathbf{z} - \mathbf{y} \parallel_{\mathbf{G}} \\ \text{s. t. } \mathbf{y} \in \mathbf{X}\end{gathered}$$

and

$$\parallel \mathbf{x} \parallel_{\mathbf{G}} = (\mathbf{x}^{\top}\mathbf{G}\,\mathbf{x})^{\frac{1}{2}}$$

is the G-norm of the vector $\mathbf{x}$.

## 3.2 Projection methods

Projection methods are easily converted to mathematical programming equivalents. A projection algorithm for solving **VIP** sets $\mathbf{A}^{(t)} = \frac{1}{\gamma}\mathbf{G}$, for all $t$, where $\mathbf{G}$ is an arbitrary symmetric and positive definite matrix. The projection algorithm is defined by the iterative formula

$$\mathbf{x}^{(t+1)} = \mathbf{Pr}_{\mathbf{X}}^{\mathbf{G}} \left(\mathbf{x}^{(t)} - \gamma\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}^{(t)})\right) \quad t = 1, 2, \cdots \tag{13}$$

Each iteration $t$ is equivalent to a quadratic mathematical programming program stated as,

$$\mathbf{x}^{(t+1)} = \min_{\mathbf{y}\in\mathbf{X}} \frac{1}{2} \parallel \mathbf{x}^{(t)} - \gamma\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}^{(t)}) - \mathbf{y} \parallel_{\mathbf{G}}^2 \tag{14}$$

defining the projection of the point $\mathbf{x}^{(t)} - \gamma\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}^{(t)})$ onto the feasible set $\mathbf{X}$ according to the metric $\parallel \cdot \parallel_{\mathbf{G}}$.

The quadratic program is a special case of affine variational inequalities. Let us define the function $H(\cdot)$ as

$$H(\mathbf{y}) = \frac{1}{2} \parallel \mathbf{x}^{(t)} - \gamma\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}^{(t)}) - \mathbf{y} \parallel_{\mathbf{G}}^2 \quad \forall \mathbf{y} \in \mathbf{X}$$

Then, according to the correspondence between a convex program (derived from the positive definiteness of matrix $\mathbf{G}$) and the optimality conditions, $\mathbf{x}^{(t+1)}$ solves the quadratic program if and only if $\mathbf{x}^{(t+1)}$ satisfies local minima condition (that are sufficient conditions for global minimum in convex programming)

$$\nabla H(\mathbf{x}^{(t+1)})^{\top}(\mathbf{y} - \mathbf{x}^{(t+1)}) \geq 0 \;\forall \mathbf{y} \in \mathbf{X}$$

or equivalently,

$$(\mathbf{f}(\mathbf{x}^{(t)}) + \frac{1}{\gamma}\mathbf{G}(\mathbf{x}^{(t+1)} - \mathbf{x}^{(t)}))^\top(\mathbf{y} - \mathbf{x}^{(t+1)}) \geq 0 \ \forall \mathbf{y} \in \mathbf{X}$$

or equivalently,

$$\mathbf{f}^{(t)}(\mathbf{x}^{(t+1)})^\top(\mathbf{y} - \mathbf{x}^{(t+1)}) \geq 0 \ \forall \mathbf{y} \in \mathbf{X}$$

At this point the equivalence between a mathematical program and an specific affine variational inequality has been shown. It is interesting to note, that any linearization method formerly described can be viewed as a projection algorithm with variable metric; the metric is defined by the approximation matrix $\mathbf{A}^{(t)}$ defined at each iteration $t$, provided the matrices are symmetric and positive definite. The projection algorithm with variable metric is equivalent to the following iterative statement:

$$\mathbf{x}^{(t+1)} = \mathbf{Pr}_{\mathbf{X}}^{\mathbf{A^{(t)}}} \left( \mathbf{x}^{(t)} - (\mathbf{A}^{(t)})^{-1}\mathbf{f}(\mathbf{x}^{(t)}) \right) \quad t = 1, 2, \cdots \tag{15}$$

In fact, it is not easy to solve the quadratic program equivalent to the affine variational inequality for **[TAP-VI]** and the convergence of a pure fixed projection method is very poor. Most linearization algorithms applied to **[TAP-VI]** have been proposed in the context of simplicial decomposition methods, because an affine variational inequalities defined over a convex hull of extreme flows might be converted into a quadratic program with simple constraints.

Dafermos was the first to apply general projection methods to **[TAP-VI]** . She also gave convergence properties in the context of traffic equilibrium. Convergence results for projection algorithms are complex and it seems that conditions are rather restrictive: it is required to $\mathbf{f}(\mathbf{x})$ be continuously differentiable and strongly monotone, provided $\rho > 0$ to be sufficiently small.

## 3.3 Diagonalization Methods

Diagonalization methods have in common that interactions among blocks of variables are removed. These algorithms, when applied to variational inequalities, extend the Jacobi and Gauss-Seidel methods for linear and nonlinear equations. They are also called *relaxation methods*, due to the relaxation of the cost interactions. Algorithms of this type were developed for in the 70's in the analysis of multiclass user networks.

Dafermos [10] presents a unified description of diagonalization methods. She introduces a smooth function

$$\mathbf{g}(\mathbf{x}, \mathbf{y}) : \mathbf{X} \times \mathbf{X} \mapsto \mathbb{R}^n$$

with the following properties

1. $\mathbf{g}(\mathbf{x}, \mathbf{x}) = \mathbf{f}(\mathbf{x})$, for all $\mathbf{x} \in \mathbf{X}$.

2. $\nabla_{\mathbf{x}}\mathbf{g}(\mathbf{x},\mathbf{y})$ is positive definite and symmetric.

The latter properties ensure a unique solution of the variational inequality

$$\mathbf{g}(\mathbf{x}^{(t+1)},\mathbf{x}^{(t)})^\top(\mathbf{x}-\mathbf{x}^{(t+1)}) \geq 0 \ \ \forall \mathbf{x}\in\mathbf{X}$$

that, by symmetry, reduces to the strictly convex program

$$\mathbf{x}^{(t+1)} = \min_{\mathbf{x}\in\mathbf{X}} H(\mathbf{x},\mathbf{x}^{(t)})$$

where $\nabla_{\mathbf{x}}H(\mathbf{x},\mathbf{y}) = \mathbf{g}(\mathbf{x},\mathbf{y})$ and $\mathbf{x}^{(t)}$ solves **VIP** if $\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)}$.

Some methods that fit into Dafermos general framework are

| | | | | |
|---|---|---|---|---|
| $\mathbf{g}(\mathbf{x},\mathbf{y})$ | $=$ | $\mathbf{f}(\mathbf{y}) + \mathbf{A}(\mathbf{y})\cdot(\mathbf{x}-\mathbf{y})$ | | Linearization |
| $(g_i(\mathbf{x},\mathbf{y}))$ | $=$ | $\mathbf{f}_i(y_1,\cdots,y_{i-1},x_i,y_{i+1},\cdots,y_n)$ | $i=1,\cdots,n$ | Nonlinear Jacobi |
| $(g_i(\mathbf{x},\mathbf{y}))$ | $=$ | $\mathbf{f}_i(x_1,\cdots,x_i,y_{i+1},\cdots,y_n)$ | $i=1,\cdots,n$ | Nonlinear Gauss-Seidel |

Dafermos' convergence conditions are tested practically to be too restrictive and much more complex, but less restrictive sufficient conditions are proposed by some authors, these conditions have an intuitive interpretation that the Jacobian of the cost function should have weak asymmetries and can be expressed as,

$$\left\|\nabla_{\mathbf{x}}\mathbf{g}^{-\frac{1}{2}}(\mathbf{x}^1,\mathbf{y}^1)\nabla_{\mathbf{y}}\mathbf{g}(\mathbf{x}^2,\mathbf{y}^2)\nabla_{\mathbf{x}}\mathbf{g}^{-\frac{1}{2}}(\mathbf{x}^3,\mathbf{y}^3)\right\|_2 < 1,$$

for $\forall \mathbf{x}^i, \mathbf{y}^i \in \mathbf{X}$ and $i = 1, 2, 3$.

## 3.4 Simplicial Decomposition Methods

Simplicial decomposition techniques have been applied to **[TAP-VI]** by Smith [28], Pang and Yu [21], Bertsekas and Gafni and Lawphongpanich and Hearn [19].

Bertsekas and Gafni utilize the arc route formulation of **[TAP-VI]** and define a master problem as the variational inequality over the current restricted set, i. e., $\mathbf{VI}(\mathbf{f}(\mathbf{x}),\mathbf{X}^{(t)})$, and apply a projection algorithm for solving the master problem. The algorithm is a linearized method, where new routes are added after each iteration of the linearization scheme. No column dropping is used. Algorithms with fixed and variable projection metric are proposed for solving the master problem. The global convergence of the algorithm is shown in the fixed projection metric case, under the following assumptions :

1. $\mathbf{f}(\mathbf{x})$ strongly monotone.

2. $\mathbf{f}(\mathbf{x})$ Lipschitz continuous.

and a linear convergence rate is obtained. For the variable projection method case, a safeguard on the selection of the variable projection matrix at each iteration has to be forced in order to show a similar convergence result.

Lawphongpanich and Hearn [19] propose an algorithm where dropping extremes is considered for extremes with zero weight in the current point if a sufficient gap function descent has been encountered in the current iteration. Two levels of aggregation are proposed:

1. Link flow extremes.
2. Commodity link flow extremes, where a commodity groups OD pairs with the same origin, we refer to them as *origin link flow* extremes.

The first option is shown to be the most efficient. The master problem is defined as the variational inequality over the current restricted set, i.e., $\mathbf{VI}(\,\mathbf{f}(\mathbf{x}), \mathbf{X}^{(t)})$ and the resolution of the master problem is proposed by means of a fixed metric projection method derived from that proposed by Bertsekas and Gafni.

Pang and Yu [21] propose a linearized simplicial decompositon method similar to that of Lawphongpanich and Hearn [19]. The aggregation level is set to link flow extremes. The master problem is defined as a linear approximation to the functional defining the variational inequality at the current point and restricted to the subset of link flow extremes, i.e.,$\mathbf{VI}(\,\mathbf{f}^{(t)}(\mathbf{x}), \mathbf{X}^{(t)})$. Extreme points with weight coefficients greater that zero in the current iteration are maintained to the next. The method for solving the master problem is a modified version of the Dantzig, van de Panne and Whinston algorithm for quadratic programming that operates on the simplicial representation of $\boldsymbol{X^{(t)}}$ in terms of its extreme points. The linear approximations proposed for defining the master problem are of the type:

1. For symmetric problems, the Jacobian matrix of $\mathbf{f}(\mathbf{x})$ at the current point.
2. For asymmetric problems, the diagonal part of the Jacobian of $\mathbf{f}(\mathbf{x})$ at the current point.

Some remarks related to the RSD algorithm are:

- A small positive tolerance $\delta$ for primal gap descent is required for ensuring convergence when extreme dropping is used.
- A convergent monotone sequence $\{\epsilon^{(t)}\}$ is required, i.e.,

$$\epsilon^{(t)} > \epsilon^{(t+1)} > 0, \text{ and } \epsilon^{(t)} \rightarrow 0 \text{ as } t \rightarrow \infty$$

- $\mathbf{c}(\mathbf{v})$ is assummed to be strictly monotone and nonnegative and the demand feasible set of total link flows $\mathcal{V}$ is assummed to be compact (i.e., closed and bounded) and convex.
- Any convergent method for solving variational inequalities may be applied to the master problem which needs to be solved approximately. The authors use a variant of Bertsekas and Gafni projection method.
- The gap function value is a natural byproduct of solving the subproblem and provides a stopping criteria, as well as monitoring the dropping extreme process. A deep study on gap functions may be found in Hearn [16].

- For large scale problem a *relative gap* is suggested as more suitable and defined as
$$\bar{G}(\mathbf{v}^{(t)}) = \frac{\mathbf{c}(\mathbf{v}^{(t)}) \cdot (\mathbf{v}^{(t)} - \mathbf{f}^{(t)})}{\mathbf{c}(\mathbf{v}^{(t)}) \cdot \mathbf{f}^{(t)}}$$
and a certain $\epsilon$-solution determined more practical for implementation purposes.

Lawphongpanich and Hearn [19] results demonstrate the potential of the simplicial decomposition algorithm with regard to computer time. They also show that the method obtains good solutions while retaining only a small number of extreme flow patterns.

# 4. Proximal point methods for VI's

The proximal point algorithm has been analyzed for inclusions of the form

$$0 \in \mathbf{T}(\mathbf{x}) \tag{16}$$

in Rockafellar [25] for the case of $\mathbf{T}$ being a monotone point-to-set map on a Hilbert space as a generalization of the algorithm of Martinet [20] for convex minimization problems. The results of Rockafellar [25] apply when the mapping is maximal monotone. In this case the proximal operator $\mathbf{P}_c(\mathbf{x}) \triangleq (\mathbf{I}+c\mathbf{T})^{-1}(\mathbf{x})$ is also maximal monotone and single valued (see for instance, Rockafellar and Wets [26], theorem 12.12). The iterations of the proximal point method for the problem (16) can be simply described as,

$$\text{Find } x_{k+1} \text{ so that } x_{x+1} \approx P_{c_k}(x_k) \tag{17}$$

where $c_k$ is a sequence so that $inf\, c_k > 0$ and the tolerances $\epsilon_k \to 0+$. If exactly $x_{x+1} = P_{c_k}(x_k)$, then this is equivalent to solve the problem

$$\text{Find } \mathbf{x}_{k+1} \text{ so that } 0 \in c_k^{-1}(\mathbf{x}_{k+1} - \mathbf{x}_k) + \mathbf{T}(\mathbf{x}_{k+1}) \tag{18}$$

or for a variational inequality on $\mathbb{R}^n$

$$\text{Find } \mathbf{x}^* \in \mathbf{X} \text{ so that } \mathbf{F}(\mathbf{x}^*)^\top(\mathbf{x} - \mathbf{x}^*) \geq \mathbf{0},\ \forall \mathbf{x} \in \mathbf{X} \tag{19}$$

iteration (17) expresses as:

$$\text{Find } \mathbf{x}_{k+1} \in \mathbf{X} \text{ so that } (\mathbf{F}(\mathbf{x}) + c_k^{-1}(\mathbf{x} - \mathbf{x}_k))^\top(\mathbf{x} - \mathbf{x}_k) \geq 0,\ \forall \mathbf{x} \in \mathbf{X} \tag{20}$$

Attaining to the general methods for VI's described in previous sections, iteration (20) seems to be well posed for its resolution provided that $\nabla\mathbf{F}(\cdot) + c_k^{-1}\mathbf{I}$ is positive definite.

Eckstein and Bertsekas [14] develope a relaxed algorithm that allows for the proximal iteration to be solved only approximately for a sequence of tolerances $\epsilon_k \to 0+$:

$$\text{Find } x_{k+1} \text{ so that } \|\, x_{x+1} - P_{c_k}(x_k)\, \| \leq \epsilon_k \tag{21}$$

If iteration (20) is carried out approximately, then, accordingly to (21) this is equivalent to choose a sequence of $\delta_k \to 0+$ and solve the proximal subproblems (20) approximately,

$$\text{Find } \mathbf{x}_{k+1} \in \mathbf{X},\ (\mathbf{F}(\mathbf{x}) + c_k^{-1}(\mathbf{x} - \mathbf{x}_k))^\top(\mathbf{x} - \mathbf{x_k}) \geq -\delta_k,\ \forall \mathbf{x} \in \mathbf{X} \tag{22}$$

suggesting clearly a stopping criterion based on the primal gap. The algorithm of Eckstein and Bertsekas [14] can be summarized as follows. At iteration $k$:

$$\mathbf{x}_{k+1} = \sigma_k \bar{\mathbf{x}}_k + (1 - \sigma_k)\mathbf{x}_k \tag{23}$$

being $\bar{\mathbf{x}}_k$ a solution of (21) (or (22) in the context of VI's), and the parameters $\sigma_k$, $\epsilon_k$ and $c_k$ taken so that, $inf\ \sigma_k > 0$, $sup\ \sigma_k < 2$, $inf\ c_k = \bar{c} > 0$ and $\sum_k \epsilon_k < +\infty$. They prove the convergence of the algorithm under the maximal monotonicity of the operator $\mathbf{T}$ and provided that a solution of the inclusion (16) exists (otherwise the algorithm may not converge).

The work of Pennanen [24] shows that, in fact the so stringent condition of maximal monotonicity required up to now for the convergence of the proximal point algorithm, can be substituted by the hypomonotonicity condition. (a mapping $\mathbf{T}$ is called hypomonotone if the $\mathbf{T} + \rho\mathbf{I}$ is monotone for some $\rho > 0$). He uses the Yosida regularization $\mathbf{T}_\rho$ of $\mathbf{T}$, $\mathbf{T}_\rho \triangleq (\mathbf{T}^{-1} + \rho\mathbf{I})^{-1}$ and shows that, in fact:

$$(\mathbf{I} + c\mathbf{T}_\rho)^{-1} = \frac{1}{c + \rho}\,(c\mathbf{P}_{c+\rho} + \rho\mathbf{I})$$

for $c \neq 0$. This is equivalent to say that the application of the relaxed proximal point of Eckstein-Bertsekas to the Yosida regularization $\mathbf{T}_\rho$ of $\mathbf{T}$ is equivalent to the application of the proximal point algorithm of Rockafellar to $\mathbf{T}$. This is stated if the following theorem due to Pennanen:

**Theorem 1.2** *(Pennanen (2002), theorem 9) Let $\mathbf{T} : X \mapsto Y$, with $0 \in Y$ be an hypomonoton point-to-set map for some $\rho > 0$ and let $\bar{\mathbf{x}} \in \mathbf{T}^{-1}(0) \cap X \neq \emptyset$ and $\mathbf{T}^{-1}(0) \cap X$ a closed set. Then, there exists $\epsilon > 0$ so that if the starting point $\mathbf{x}_0 \in B_\epsilon(0)$. Then, the iteration:*

$$\mathbf{x}_{k+1} = B_\epsilon(0) \cap (\mathbf{I} + c_k\mathbf{T})^{-1}(\mathbf{x}_k) \tag{24}$$

*converges weakly and Fejér monotonically to a point $\bar{\mathbf{x}} \in \mathbf{T}^{-1}(0) \cap X$ if $c_k \nearrow \bar{c} \leq +\infty$ and $inf\ c_k > 2\rho$. Further if the map $\mathbf{y} \mapsto \mathbf{T}_\rho^{-1}(\mathbf{y}) \cap X$ satisfies for some positive constant $L_T$ the calming condition at $\bar{\mathbf{x}}$, $\mathbf{x} \in \mathbf{T}_\rho(\mathbf{y}) \Rightarrow \|\, \mathbf{x} - \bar{\mathbf{x}}\, \| \leq L_T \|\, \mathbf{y}\, \|$, then the sequence of points generated by (24) converges to $\mathbf{T}^{-1}(0) \cap X$ linearly with rate*

$$\left(1 - \frac{\bar{c}}{\bar{c} - \rho}\left(2 - \frac{\bar{c}}{\bar{c} - \rho}\right)\frac{\bar{c}^2}{(L_T + \rho)^2 + \bar{c}^2}\right)^{\frac{1}{2}} < 1$$

It is proved that applying the proximal point iteration to $\mathbf{T}$ with the parameters stated in the above theorem is equivalent to the application of the over-relaxed proximal point of Eckstein-Bertsekas with parameters $c'_k = c_k - \rho$ and $\sigma_k = c_k/(c_k - \rho)$ and alternatively that the application of the over-relaxed iteration of Eckstein-Bertsekas on $\mathbf{T}$ with $\sigma_k = (c'_k - \rho)/c'_k$, with $inf\ c'_k > 2\rho$ is equivalent to the proximal point iteration $\mathbf{x}_{k+1} = B_\epsilon(0) \cap (\mathbf{I} + (c'_k - \rho)\mathbf{T})^{-1}(\mathbf{x}_k)$.

# 5. Applications to the demand adjustment problem

The matrix adjustment problem on a traffic network can be formulated as the following bilevel programming problem,

$$\begin{array}{lll} Min_{\,g\geq 0} & F(g) \;=\; z_1 F_1(v) + z_2 F_2(g) \\ s.t. & v \in \mathcal{V}^*(g) \end{array} \tag{25}$$

where by $v$ it is denoted the vector of total link flows and by $g$ it is denoted the origin-destination trip matrix so that, when assigned to the traffic network reproduces some traffic observed conditions.

The functions $F_1$ and $F_2$ can be viewed as distances between the response of the lower level problem and the data input that is used for the adjustment, i.e., the link flow observations and an obsolete or reference trip matrix. Usually $F_1$ is adopted as $F_1(v) = \frac{1}{2}(v_1 - \hat{v}_1)^\top U\,(v_1 - \hat{v}_1)$ and $F_2$ as $F_2(g) = \frac{1}{2}(g - \bar{g})^\top B\,(g - \bar{g})$ or an entropy function and the subscript 1 on the link flows $v_1$ stands for a subset of components where there are observations available, $\hat{v}_1$. $U$ and $W$ are weighting matrices. Usually $\bar{g}$ is referred to as the target matrix.

Also in (25), $\mathcal{V}(g)$ is the set of feasible link flows corresponding to the O-D matrix $g$ and $\mathcal{V}^*(g)$ is the set of solutions in terms of the total link flows of the variational inequality (26) that formulates the traffic assignment problem on an asymmetric traffic network.

$$\text{Find } v \in \mathcal{V}(g), \text{ so that } \; c(v)^\top (v' - v) \geq \; 0, \; \forall v' \in \mathcal{V}(g) \tag{26}$$

In general if the assignment map $\mathcal{V}^*(g)$ is a singleton, then the bilevel programming problem (25) for the demand adjustment problem can be reformulated as:

$$Min_{\,g\geq 0}\, F(g) \;=\; z_1 F_1(v^*(g)) + z_2 F_2(g) \tag{27}$$

It must be remarked that the upper level function $F(g)$ is non-differentiable.

The proximal point algorithm for minimizing a (continuously differentiable) function $F(g)$ on a convex set $X$ consists of solving a sequence of problems of the type:

$$Min_{\,g\in X}\;\; F(g) + \frac{\mu_\ell}{2}\,\|\,g - g^\ell\|_2^{\;2} \tag{28}$$

for a sequence $\{\mu_\ell\}$ of parameters, $\mu_\ell \leq \bar{\mu}$. It must be also noted that, if $\mu_\ell$ is high enough, a solution $g^*(\mu_\ell)$ of the above problem (28) provides a good approximation to an element of the generalized gradient of Clarke [2] $\partial F(g^*(\mu_\ell))$ for the function $F(g)$ at $g^*(\mu_\ell)$, because the Fermat rule applied to problem (28) results in:

$$-\mu_\ell(g^*(\mu_\ell) - g^\ell) \in \partial F(g^*(\mu_\ell)). \tag{29}$$

In order to apply the proximal point iteration for the demand adjustment problem, the previous problem (28) needs to be transformed in the following one:

$$\begin{aligned} Min_{(v,g)\in\mathbf{\Omega}} \quad & H(v,g) + \frac{\mu_\ell}{2} \, \|g - g^\ell\|_2^{\,2} \\ s.t. \quad & G(v,g) = 0 \end{aligned} \tag{30}$$

For a non negative O-D matrix $g$, $\mathbf{\Omega}$ denotes the feasible cone of flows-demands or the set of pairs $(v,g)$ so that $g \geq 0$ and $v \in \mathcal{V}(g)$ and $G(v,g)$ is a gap function for the variational inequality (26) parametrized by $g$. The primal and dual gap functions $G_P(v)$ and $G_D(v)$, have already been mentioned in section 2. As the methods to be used are primal feasible it is immaterial the value of $G(v,g)$ outside $\mathbf{\Omega}$.

In this paper we shall deal only with the primal gap function $G_P(v)$ parametrized by $g$. If the constraint $G_P(v,g) = 0$ is penalized with parameter $1/\alpha_\ell$, then (30) can be approximated by:

$$Min_{(v,g)\in\mathbf{\Omega}} \quad \psi_\ell(v,g) = G_P(v,g) + \alpha_\ell H(v,g) + \frac{\alpha_\ell\mu_\ell}{2} \, \|g - g^\ell\|_2^{\,2} \tag{31}$$

where $\alpha_\ell \cdot \mu_\ell \to 0$. The optimality conditions for problem (31) can be stated by means of the Fermat rule as:

$$0 \in \bar{\partial} G_P(v,g) + \alpha_\ell \nabla H(v,g) + \alpha_\ell\mu_\ell(g - g^\ell) + N_{\mathbf{\Omega}}(v,g) = \mathbf{T}(v,g) \tag{32}$$

The calculation of the Clarke's subgradient of the primal gap function is done in Codina and Montero [9] and it is shown there that:

$$\bar{\partial} G_P(\tilde{v},\tilde{g}) = \begin{pmatrix} c(\tilde{v}) + \left(\dfrac{\partial c}{\partial v}\right)_{\tilde{v}}^{\top} (w - \tilde{v}) \\ \hline t(\tilde{g}) \end{pmatrix} \tag{33}$$

where $w \in argmin\{\, G_P(\tilde{v}) = -Min_{\,w\in\mathcal{V}(\tilde{g})}\, c(\tilde{v})^\top (w - \tilde{v}) \,\}$.

The generalized equation (32) can be difficult to solve and we will consider now the case in which the travel costs $c(v)$ are strongly monotone with modulus $m_c$. Now it is easy to see that

$$\begin{pmatrix} c_\alpha(v^*(g),g) \\ \hline -t_{\alpha,\mu}(v^*(g),g) \end{pmatrix} + N_{\mathbf{\Omega}}(v^*(g),g) \in \mathbf{T}(v^*(g),g) \tag{34}$$

where $c_\alpha(v,g) = c(v) + \alpha\nabla_v H(v,g)$ and $t_{\alpha,\mu}(v,g) = t(g) - \mu\alpha(g - g^\ell) - \alpha\nabla_g H(v,g)$.

A good approximation for the inclusion (32) can be the variational inequality:

$$c_\alpha(v,g)^\top(v'-v) - t_{\alpha,\mu}(v,g)^\top(g'-g) \geq 0, \ \ \forall\,(v',g') \in \mathbf{\Omega} \tag{35}$$

presenting the structure of an elastic demand traffic assignment problem that can be converted to a fixed demand traffic assignment problem by means of the Gartner's transformation [15].

The cost approximation algorithms in Patriksson [22] can be used for the case of the gap function $G_P(v,g)$ and $H = H(v)$. At iteration $k$, of the cost approximation algorithm, the resulting variational inequality subproblem at point $g^k$ is:

$$c_\alpha(v)^\top(w-v) + (-t(g^k) + \mu\alpha(g-g^\ell))^\top(g'-g) \geq 0, \ \ \forall(w,g') \in \mathbf{\Omega} \tag{36}$$

where $t(g_\ell)$ are the equilibrium O-D travel times resulting from the assignment of the O-D matrix $g^k$ with costs $c(v)$. It must be remarked that the variatinal inequality (36) does not present the required condicitions for convergence of the algorithms cited in Section 3 and it is here where the use of the proximal point algorithm can be advantageous in order to obtain a solution of (36).

# 6. Conclusions and further work

This paper reviews firstly the conditions for convergence of the more used algorithms to solve variational inequalities with special incidence in the traffic assignment problem and summarizes the recently new conditions developed by Pennanen [24] for the proximal point algorithm which do not require strong monotonicity. The application of the proximal point algorithm to a problem of bilevel programming that formulates the adjustment of origin-destination trip matrices is also shown. A set of numerical experiments will be required in order to evaluate practically and in comparison with other methods, the performance of the proximal point algorithm to the problems examined in this paper.

# 7. Acknowledgments

This research has been supported under Spanish I+D research project BFM2002-04548-C03-01.

## References

[1] Aashtiani H.Z., Magnanti T.L. (1981). *Equilibria on a congested transportation network.* SIAM J. Algebraic and Discrete Methods 2, pp 213-226.

[2] Clarke, F. H. (1983). *Optimization and Nonsmooth Analisys*, John Wiley Inter-Science.

[3] Codina, E. (2000). "Lipschitz Continuity of O-D Travel Times for the Asymmetric Traffic Assignment Problem", *Working Paper*, DR/2000-21, Departament d'Estadística i Investigació Operativa. Universitat Politècnica de Catalunya.

[4] Codina, E. (2000). "A Test for the Approximation of Gradients of the Upper Level Function in Demand Adjustment Problems", *Working Paper*, DR/2000-23. Departament d'Estadística i Investigació Operativa. Universitat Politècnica de Catalunya.

[5] Codina, E. and Barceló, J. (2000). "Adjustment of O-D Trip Matrices from Traffic Counts: an Algorithmic Approach Based on Conjugate Directions", *Proceedings of the* $8^{th}$ *EWGT.*, Rome, Italy, July 2000, pp. 427-432. (forthcoming in European Journal of Operational Research)

[6] Codina, E. (2001). "Consistency of an approximation to the upper level objective function gradients in the o-d demand adjustment problem". *Working Paper*, DR/2001-16. Departament d'Estadística i Investigació Operativa. Universitat Politècnica de Catalunya.

[7] Codina E., Montero L. (2001) *"A Method to Approximate the Steepest Descent Direction of the O-D Matrix Adjustment Problem"*. Proceedings of the TRISTAN IV Conference at Açores, Portugal, June 2001. pp. 537-542. (forthcoming in Annals of Operations Research)

[8] Codina, E., García, Ricardo and Marín Angel. (2001). "New Algorithmic Alternatives for the O-D Matrix Adjustment Problem on Traffic Networks", *Proceedings of the* $9^{th}$ *EWGT.*, Bari, Italy, June 2001, pp. 421-425. (submitted for publication to European Journal of Operational Research)

[9] Codina, E., Montero L. (2003). "Evaluation of subgradients for the primal gap function of variational inequalities and an application to the demand adjustment problem", EURO-INFORMS joint International Meeting. Istanbul, July 2003.

[10] Dafermos S. C. (1982) *Relaxation algorithms for the general asymmetric traffic equilibrium problem.* Transportation Science 16 (2), 231-240.

[11] Dafermos S., Nagurney A. (1984): Sensitivity analysis for the asymmetric network equilibrium problem. Math. Prog. 28, 174-184.

[12] Dontchev, A. L. and Rockafellar, R. T. (1996). "Characterizations of Strong Regularity for Variational Inequalities over Polyhedral Convex Sets", *SIAM Journal of Optimization*, 12 No. 1, 1087-1105.

[13] Dontchev A. L. and Rockafellar, R. T. (2001). "Ample Parametrization of Variational Inclusions". *SIAM Journal of Optimization*, 12 No. 1 170-187.

[14] Eckstein J. and Bertsekas D.P. (1992) *On the Douglas-Rachford splitting method and the proximal poin algorithm for maximal monotone operators*, Mathematical Programming 55 293-318.

[15] Gartner, N.H. (1980) *Optimal traffic assignments with elastic demands: a review. Part II: algorithmic approaches*, Transpn. Sci. 14 (1980) 192-208.

[16] Hearn D. W. (1982) *The gap function of a convex program.* Oper. Res. Lett. 1, 67-71.

[17] Hearn D.W., Lawphongpanich S., Ventura J.A. (1987) *" Restricted Simplicial Decomposition: Computation and Extensions."* Mathematical Programming Study Vol 31, pp 99-118.

[18] Heydecker B.G. (1986) *On the definition of traffic equilibrium.* Transportation Research 20 B pp 435-440.

[19] Lawphongpanich S., Hearn D. W. (1984) *Simplicial decomposition of the asymmetric traffic assignment problem.* Transp. Res 18B, 123-133.

[20] Martinet B., (1970): Regularisation d'inequations variationelles par approximations successives *"Revue Française d'Informatique et de Recherche Operationelle"* 4(R-3) pp. 154-158.

[21] Pang J. S., Yu C. S. (1984) *Linearized simplicial decomposition methods for computing traffic equilibria on networks.* Networks 14(3), 427-438.

[22] Patriksson, Michael (1999). Nonlinear Programming and variational inequality problems. A unified Approach. Kluwer Academic Publishers.

[23] Patriksson, Michael (1993) *"Partial linearization Methods in Nonlinear Programming ".* Journal of Optimization Theory and Applications, 78, pp. 227-246.

[24] Pennanen, T.(2002): Local convergence of the proxinal point algorithm and multiplier methods without monotonicity *Mathematics of Operations Research*, Vol. 27 No. 1, pp 170-191.

[25] Rockafellar, R. Tyrrell (1976) *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. 14 pp. 877-898.

[26] Rockafellar, R. Tyrrell and R. Wets, Roger J-B., (1998) Variational Analysis, Springer Verlag.

[27] Smith M. J. (1979) *Existence, uniqueness and stability of traffic equilibria.* Transp. Res. 13B(4), 295-304.

[28] Smith M. J. (1983) *An algorithm for solving asymmetric equilibrium problems with continuous cost-flow function.* Transp. Res. 17B (5), 365-372.

[29] Wardrop J. G. (1952) *Some theoretical aspects of road traffic research.* Proc. Inst. Civ. Eng. Part II, 1 (2), 325-378.