Splitting Algorithms for a Class of Monotone Mappings with Application to the Traffic Equilibrium Problem

Masao FUKUSHIMA*

October 1993

* Graduate School of Information Science,
Nara Institute of Science and Technology,
Ikoma, Nara 630-01, Japan

Abstract We apply Douglas-Rachford splitting algorithm to a class of multi-valued equations consisting of the sum of two monotone mappings. Compared with the dual application of the same algorithm, which is known as the alternating direction method of multipliers, the primal application yields algorithms that seem somewhat involved. However the resulting algorithms may be applied effectively to some problems. In particular we show that they are used to derive decomposition algorithms for solving the variational inequality formulation of the traffic equilibrium problem.

Key Words Splitting algorithms, monotone mapping, proximal point algorithm, variational inequality problem, traffic equilibrium.

Abbreviated Title Splitting algorithms for monotone mappings

1. Introduction

The *proximal point algorithm*, first introduced by Martinet [13] and later developed by Rockafellar [17], serves as a fundamental tool for solving the problem of finding a zero of a maximal monotone mapping, which contains as special cases monotone variational inequality and convex optimization problems. The main part of the proximal point algorithm consists of evaluating the resolvent of the given maximal monotone mapping at each iteration.

Splitting algorithms studied by Lions and Mercier [12] and Passty [15] may be considered variants of the proximal point algorithm, which are designed to find a zero of the sum of two maximal monotone mappings by evaluating the resolvents of the two mappings separately at each iteration. Those algorithms are therefore expected to be most effective when the given maximal monotone mapping is decomposed into two in such a way that each of them can be treated conveniently. In particular, by exploiting a special structure of the mappings involved, those algorithms are often amenable to parallel computation [4, 19, 20].

Among the splitting algorithms considered in [12, 15], the one called *Douglas-Rachford splitting algorithm* enjoys the most favorable convergence properties [4, §3.4] (see also [6]). Gabay [9] has considered the problem with particular structure (cf. problem (3.1) stated in §3) and shown that Douglas-Rachford splitting algorithm applied to its *dual* yields the *alternating direction method of multipliers* [10, 11]. The latter method has turned out to be useful to obtain highly parallelizable algorithms for various optimization problems [1, 5, 8].

In this paper we consider the Douglas-Rachford splitting algorithm applied directly to the given problem, rather than to the dual as in Gabay [9]. The resulting algorithm may be regarded as a generalization of the one recently presented by Eckstein and Fukushima [7] in the context of convex programming. In particular, we derive a decomposition algorithm for solving the asymmetric traffic equilibrium problem formulated as a variational inequality problem.

2. Preliminaries

A multi-valued (point-to-set) mapping T from a real Hilbert space X into itself is called *monotone* if

$$\langle z - z', x - x' \rangle \geq 0 \qquad \forall x \in X, x' \in X, z \in T(x), z' \in T(x'),$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. It is said to be *maximal* if there is no monotone mapping whose graph strictly contains the graph of T. The inverse of T defined by $T^{-1}(z) = \{x \in X \mid z \in T(x)\}$ is maximal monotone if and only if T is maximal monotone. Also it is known [2] that T is maximal monotone if and only if, for any $x \in X$ and any scalar $\lambda > 0$, there exists a unique

$z \in X$ such that $x \in (I + \lambda T)(z)$, where I denotes the identity mapping. Thus, for any maximal monotone mapping T, the mapping $J_T^\lambda = (I + \lambda T)^{-1}$, called the *resolvent* of T, is a single-valued mapping defined everywhere on X.

In view of the relation

$$0 \in T(x) \quad \text{if and only if} \quad x = J_T^\lambda(x),$$

the *proximal point algorithm* for solving the multi-valued equation

$$0 \in T(x) \tag{2.1}$$

generates a sequence $\{x^{(k)}\}$ by the recurrence

$$x^{(k+1)} = J_T^\lambda(x^{(k)}).$$

It is known [17] that the sequence $\{x^{(k)}\}$ thus generated converges weakly to a solution of (2.1) from an arbitrary starting point $x^{(0)}$, whenever T is maximal monotone and (2.1) has at least one solution.

In some cases, T can be expressed as the sum of two maximal monotone mappings, say $T = R + S$, in such a way that their resolvents J_R^λ and J_S^λ are much easier to evaluate than the resolvent J_T^λ of the given mapping T. Splitting algorithms [12, 15] are designed to solve

$$0 \in R(x) + S(x), \tag{2.2}$$

using J_R^λ and J_S^λ instead of J_T^λ. In particular, *Douglas-Rachford splitting algorithm* studied by Lions and Mercier [12] generates a sequence $\{z^{(k)}\}$ by the recurrence

$$z^{(k+1)} = [J_R^\lambda \circ (2J_S^\lambda - I) + (I - J_S^\lambda)](z^{(k)}) \tag{2.3}$$

with an arbitrary starting point $z^{(0)} \in X$, where $\circ$ denotes the composition of mappings. For this algorithm, the following convergence result has been established.

Theorem 2.1 [12] *Let R and S be maximal monotone multi-valued mappings from X into itself. If the set of solutions of (2.2) is nonempty, then the sequence $\{z^{(k)}\}$ generated by (2.3) weakly converges to a point $z \in X$ such that $x = J_S^\lambda(z)$ solves (2.2).*

In [6], Douglas-Rachford algorithm has been generalized to allow over- or under-relaxation and inexact evaluation of the resolvents. Although such generalization may be useful in practice, we restrict our attention here to the original algorithm (2.3) for simplicity of presentation. In the next section, we shall apply it to a class of multi-valued equations.

3. Algorithms

Let F and G be maximal monotone multi-valued mappings on real Hilbert spaces X and Y, respectively. Let A be a continuous linear mapping from X into Y and let A^T denote its adjoint. We are interested in the problem of finding a point $x \in X$ such that

$$0 \in F(x) + A^T G(Ax). \tag{3.1}$$

Throughout we shall assume that

$$AA^T \text{ is an isomorphism of } Y. \tag{3.2}$$

We remark that problem (3.1) contains as a special case the convex minimization problem

$$\text{minimize} \quad f(x) + g(Ax), \tag{3.3}$$

where f and g are closed proper convex functions on X and Y, respectively. In fact, a necessary and sufficient optimality condition for (3.3) can be expressed as (3.1) by letting $F = \partial f$ and $G = \partial g$, where ∂ denotes the subdifferential operator, and the subdifferential mapping of any closed proper convex function is maximal monotone [16].

We also remark that problem (3.1) contains the variational inequality problem of finding $y \in W$ such that

$$\langle G(y), y' - y \rangle \geq 0 \quad \forall y' \in W, \tag{3.4}$$

where $W \subset Y$ is the closed convex set defined by

$$W = \{y \in Y \,|\, y = Ax,\ x \in V\}$$

with the set $V \subset X$ being nonempty, closed and convex. (In (3.4), we have tacitly assumed that G is a single-valued mapping. This is for notational convenience and the arguments below can easily be extended to the case where G is multi-valued.) In fact, problem (3.4) is equivalent to the problem of finding $x \in V$ such that

$$\langle G(Ax), Ax' - Ax \rangle \geq 0 \quad \forall x' \in V,$$

namely

$$\langle A^T G(Ax), x' - x \rangle \geq 0 \quad \forall x' \in V.$$

Moreover the last problem may be stated equivalently as

$$-A^T G(Ax) \in \partial \delta(x \,|\, V), \tag{3.5}$$

where

$$\partial\delta(x \,|\, V) = \{z \in X \,|\, \langle z, x' - x\rangle \leq 0 \;\; \forall x' \in V\}$$

is the subdifferential of the indicator function of V at x, which is called the normal cone to V at x [16, p. 215]. Since V is closed and convex, the mapping $\partial\delta(\cdot \,|\, V)$ is maximal monotone. Thus, by letting $F = \partial\delta(\cdot \,|\, V)$, problem (3.5) is reduced to problem (3.1).

There is an alternative way of representing a variational inequality problem in the form (3.1). Let $W \subset Y$ be a given nonempty closed convex set and consider the following variational inequality: Find $x \in X$ such that $Ax \in W$ and

$$\langle F(x), x' - x\rangle \geq 0 \qquad \forall x' \in X \;\; \text{such that} \;\; Ax' \in W.$$

It is not difficult to see that this problem is equivalent to finding $x \in X$ such that

$$\langle F(x), x' - x\rangle + \delta(Ax' \,|\, W) - \delta(Ax \,|\, W) \geq 0 \quad \forall x' \in X,$$

which in turn is equivalently written

$$0 \in F(x) + A^T \partial\delta(Ax \,|\, W). \tag{3.6}$$

Since $\partial\delta(\cdot \,|\, W)$ is maximal monotone, the last problem is contained in (3.1).

It should be noted that assumption (3.2) differs from the assumption that $A^T A$ is an isomorphism, which is commonly made in the dual application of Douglas-Rachford algorithm, i.e., the alternating direction method of multipliers [1, 9]. (The latter algorithm is described in §5 of this paper.) Obviously the two assumptions are incompatible except in the special case where A is square and nonsingular. Thus the primal and the dual approaches have in general different domains of applications, even though they are both designed to solve problems of the form (3.1). (By dual approach, we mean that an algorithm is obtained via the dual problem but it is expressed in terms of the mappings involved in the primal problem. By primal approach, we mean that an algorithm is obtained via the primal problem and it is expressed in terms of the mappings involved in the primal problem.)

Now let us apply Douglas-Rachford splitting algorithm to problem (3.1). To this end, we show that the mapping $A^T \circ G \circ A$ is maximal monotone. This fact itself follows immediately from [9, Proposition 4.1]. Nevertheless we include the proof here, because the analysis will turn out to be helpful in deriving computational procedures later.

Proposition 3.1 *Under assumption (3.2), the mapping $A^T \circ G \circ A$ is maximal monotone.*

Proof. First observe that $A^T \circ G \circ A$ is monotone, since we have for any $x \in X, x' \in X$ and any $z \in G(Ax), z' \in G(Ax')$

$$\langle A^T z - A^T z', x - x'\rangle = \langle z - z', Ax - Ax'\rangle \geq 0,$$

where the inequality holds by the monotonicity of G.

It is known [2] that a monotone mapping is maximal if and only if its resolvent is defined everywhere. To prove the maximality of $A^T \circ G \circ A$, it is therefore sufficient to show that its resolvent $J^{\lambda}_{A^T \circ G \circ A} = (I + \lambda A^T \circ G \circ A)^{-1}$ is defined everywhere on X, i.e., there exists $x \in X$ such that

$$u \in (I + \lambda A^T \circ G \circ A)(x) \tag{3.7}$$

for any given $u \in X$ and any scalar $\lambda > 0$. In the following, we shall show that (3.7) is actually satisfied by

$$x = u - \lambda A^T (G^{-1} + \lambda A A^T)^{-1}(Au). \tag{3.8}$$

Note that $(G^{-1} + \lambda A A^T)^{-1}$ is a single-valued mapping defined everywhere on X, because G^{-1} is maximal monotone and the positive semidefinite linear mapping AA^T is an isomorphism by assumption (3.2). So the right-hand side of (3.8) is well defined for any $u \in X$. Let us rewrite (3.8) as

$$x = u - \lambda A^T v, \tag{3.9}$$

where

$$v = (G^{-1} + \lambda A A^T)^{-1}(Au). \tag{3.10}$$

Since (3.10) is equivalent to

$$Au \in (G^{-1} + \lambda A A^T)(v),$$

we have

$$A(u - \lambda A^T v) \in G^{-1}(v),$$

which together with (3.9) yields

$$v \in G(Ax). \tag{3.11}$$

It then follows immediately from (3.9) and (3.11) that x certainly satisfies (3.7) as claimed. □

In the proof of the above proposition, we have shown that the resolvent of $A^T \circ G \circ A$ can be expressed as

$$J^{\lambda}_{A^T \circ G \circ A}(u) = u - \lambda A^T (G^{-1} + \lambda A A^T)^{-1}(Au). \tag{3.12}$$

However, it appears that the right-hand side of (3.12) is difficult to evaluate in practice, because it explicitly contains the inverse of G. We now give an alternative way of evaluating the resolvent $J^{\lambda}_{A^T \circ G \circ A}(u)$. Let v be given by (3.10) and let

$$w = Au - \lambda A A^T v. \tag{3.13}$$

Then (3.9) and (3.13) imply

$$w = Ax. \tag{3.14}$$

Moreover, since $(AA^T)^{-1}$ exists by assumption (3.2), (3.13) yields

$$v = \frac{1}{\lambda}(AA^T)^{-1}(Au - w). \tag{3.15}$$

By (3.11), (3.14) and (3.15), we obtain

$$0 \in \frac{1}{\lambda}(AA^T)^{-1}(w - Au) + G(w). \tag{3.16}$$

Thus $x = J^{\lambda}_{A^T \circ G \circ A}(u)$ can be computed by the following procedure: Given $u \in X$,

(i) find the unique solution $w \in X$ of (3.16);

(ii) compute $v \in X$ by (3.15);

(iii) compute $x \in X$ by (3.9).

Now we are ready to apply Douglas-Rachford algorithm (2.3) to problem (3.1). Obviously there are two possible choices

$$R = A^T \circ G \circ A \quad \text{and} \quad S = F \tag{3.17}$$

and

$$R = F \quad \text{and} \quad S = A^T \circ G \circ A. \tag{3.18}$$

Taking into account the above procedure (i)-(iii) for evaluating the resolvent $J^{\lambda}_{A^T \circ G \circ A}$, we shall derive two algorithms called Algorithm I and Algorithm II, which correspond to (3.17) and (3.18), respectively.

First consider the case (3.17). Let

$$x^{(k)} = J^{\lambda}_S(z^{(k)}) = J^{\lambda}_F(z^{(k)}).$$

Then we have

$$(I - J^{\lambda}_S)(z^{(k)}) = z^{(k)} - x^{(k)}$$

$$(2J^{\lambda}_S - I)(z^{(k)}) = 2x^{(k)} - z^{(k)}$$

and hence we may write (2.3) as

$$z^{(k+1)} = J^{\lambda}_R(2x^{(k)} - z^{(k)}) + z^{(k)} - x^{(k)}. \tag{3.19}$$

Since $R = A^T \circ G \circ A$ and since $J^{\lambda}_{A^T \circ G \circ A}(u)$ is given by the vector x determined by (3.9), (3.15) and (3.16), we may substitute $u = 2x^{(k)} - z^{(k)}$ in (3.9) to get

$$J^{\lambda}_R(2x^{(k)} - z^{(k)}) = 2x^{(k)} - z^{(k)} - \lambda A^T v^{(k)}, \tag{3.20}$$

where $v^{(k)}$ is given by

$$v^{(k)} = \frac{1}{\lambda}(AA^T)^{-1}(A(2x^{(k)} - z^{(k)}) - w^{(k)})$$

with $w^{(k)}$ being the solution of

$$0 \in \frac{1}{\lambda}(AA^T)^{-1}(w - A(2x^{(k)} - z^{(k)})) + G(w).$$

Finally, by (3.19) and (3.20), we get

$$z^{(k+1)} = x^{(k)} - \lambda A^T v^{(k)}.$$

Summarizing the above arguments, we obtain the following algorithm for solving problem (3.1).

Algorithm I:

Step 1: Choose a constant $\lambda > 0$ and an initial vector $z^{(0)} \in X$. Let $k := 0$.

Step 2: Find the unique solution $x^{(k)} \in X$ of the equation

$$0 \in \frac{1}{\lambda}(x - z^{(k)}) + F(x).$$

Step 3: Put

$$r^{(k)} = A\,(2x^{(k)} - z^{(k)}).$$

Find the unique solution $w^{(k)} \in Y$ of the equation

$$0 \in \frac{1}{\lambda}(AA^T)^{-1}(w - r^{(k)}) + G(w)$$

and put

$$v^{(k)} = \frac{1}{\lambda}(AA^T)^{-1}(r^{(k)} - w^{(k)}).$$

Step 4: Put

$$z^{(k+1)} = x^{(k)} - \lambda A^T v^{(k)}.$$

Set $k := k + 1$ and go to Step 2. $\square$

For the case (3.18), we may derive the following algorithm. The detail of the derivation is omitted because it is similar to the previous algorithm.

Algorithm II:

Step 1: Choose a constant $\lambda > 0$ and an initial vector $z^{(0)} \in X$. Let $k := 0$.

Step 2: Find the unique solution $w^{(k)} \in Y$ of the equation

$$0 \in \frac{1}{\lambda}(AA^T)^{-1}(w - Az^{(k)}) + G(w).$$

and put

$$v^{(k)} = \frac{1}{\lambda}(AA^T)^{-1}(Az^{(k)} - w^{(k)}).$$

Step 3: Put

$$x^{(k)} = z^{(k)} - \lambda A^T v^{(k)}.$$

Step 4: Find the unique solution $y^{(k)} \in X$ of the equation

$$0 \in \frac{1}{\lambda}(y - (2x^{(k)} - z^{(k)})) + F(y)$$

and put

$$z^{(k+1)} = y^{(k)} + \lambda A^T v^{(k)}.$$

Set $k := k + 1$ and go to Step 2. $\square$

Potential difficulties in implementing the above algorithms lie in the fact that they involve the inverse mapping $(AA^T)^{-1}$. Though the evaluation of such a mapping is in general costly, it may not be an obstacle if A has some favorable structure. In the next section, we shall show that this is indeed the case for a variational inequality formulation of the traffic equilibrium problem.

4. Decomposition of Traffic Equilibrium Problems

The problem of determining an equilibrium flow pattern on a transportation network can be formulated as a variational inequality problem [3, 18]. Here we consider the case where the travel cost function on each arc is monotone but possibly asymmetric and the travel demand for each origin-destination (OD) pair is fixed.

Let $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ denote a transportation network with the set $\mathcal{N}$ of nodes and the set $\mathcal{A}$ of arcs. We shall denote the cardinality of the sets $\mathcal{N}$ and $\mathcal{A}$ by $|\mathcal{N}|$ and $|\mathcal{A}|$, respectively. Suppose there are K OD pairs $\alpha = (O_\alpha, D_\alpha)$ with $O_\alpha, D_\alpha \in \mathcal{N}$, and each OD pair α has a *fixed* demand $\Delta_\alpha > 0$. Let $x_\alpha \in R^{|\mathcal{A}|}$ denote an arc flow vector for OD pair α. Then, for each $\alpha = 1, \cdots, K$, the flow vector x_α has to satisfy the conservation of flow and the nonnegativity conditions

$$Ex_\alpha = d_\alpha, \qquad x_\alpha \geq 0,$$

where E is the node-arc incidence matrix of the network $\mathcal{G}$ and $d_\alpha \in R^{|\mathcal{N}|}$ is the travel demand vector whose components are all zero except the O_αth and D_αth components being Δ_α and $-\Delta_\alpha$,

respectively. We denote $x = (x_1, \cdots, x_K) \in R^{K|\mathcal{A}|}$ and call it *feasible* if

$$x_\alpha \in \Phi_\alpha \qquad \alpha = 1, \cdots, K,$$

where Φ_α is the polyhedral convex set in $R^{|\mathcal{A}|}$ defined by

$$\Phi_\alpha = \left\{ x_\alpha \in R^{|\mathcal{A}|} \;\middle|\; Ex_\alpha = d_\alpha, \;\; x_\alpha \geq 0 \right\}. \tag{4.1}$$

We shall denote the set of feasible flows by

$$\Phi = \Phi_1 \times \cdots \times \Phi_K. \tag{4.2}$$

Any feasible flow $x = (x_1, \cdots, x_K)$ determines the vector of *total* flow

$$y = \sum_{\alpha=1}^{K} x_\alpha. \tag{4.3}$$

We also call y feasible if $y \in \Psi$, where Ψ is the polyhedral convex set defined by

$$\Psi = \left\{ y \in R^{|\mathcal{A}|} \;\middle|\; y = \sum_{\alpha=1}^{K} x_\alpha, \;\; x \in \Phi \right\}. \tag{4.4}$$

We suppose that the *cost* of travel on the network depends on the amount of total flow $y \in R^{|\mathcal{A}|}$. Therefore the cost function, which we denote C, is a mapping from $R^{|\mathcal{A}|}$ into itself. We suppose throughout that C is single-valued. This is for notational convenience and the multi-valued case can be treated in an analogous but somewhat bothersome manner. Every component of the vector $C(y)$ represents the cost of travel on an arc when the total flow y on the entire network is given. Note that the travel arc on an arc may depend not only on the flow on that arc but on the flows on other arcs. We assume that the mapping C is maximal monotone.

As shown by Smith [18], a feasible (total) flow vector $y \in \Psi$ is *user-optimal* if and only if

$$\langle C(y) \,,\, y' - y \rangle \geq 0 \qquad \forall y' \in \Psi. \tag{4.5}$$

This is a variational inequality problem on the space of total flow vectors. In view of (4.1), (4.3) and (4.4), we may reformulate problem (4.5) in terms of the vector $x = (x_1, \cdots, x_K)$ as follows: Find $x = (x_1, \cdots, x_K) \in \Phi$ such that

$$\left\langle C\left(\textstyle\sum_{\alpha=1}^{K} x_\alpha\right) \,,\, \sum_{\alpha=1}^{K} x'_\alpha - \sum_{\alpha=1}^{K} x_\alpha \right\rangle \geq 0 \qquad \forall (x'_1, \cdots, x'_K) \in \Phi. \tag{4.6}$$

This problem is defined on the space of flows of individual OD pairs and has the special Cartesian product structure in the sense of [14].

Now define the $|\mathcal{A}| \times K|\mathcal{A}|$ matrix

$$A = [\, I \; I \cdots I \,], \tag{4.7}$$

where I denotes the $|\mathcal{A}| \times |\mathcal{A}|$ identity matrix. Obviously $(AA^T)^{-1}$ exists and is given simply as

$$(AA^T)^{-1} = \frac{1}{K}\, I. \tag{4.8}$$

Since

$$Ax = \sum_{\alpha=1}^{K} x_\alpha,$$

we may rewrite (4.6) as

$$\langle C\,(Ax)\,,\; Ax' - Ax\rangle \geq 0 \qquad \forall x' \in \Phi,$$

namely

$$\langle A^T C\,(Ax)\,,\; x' - x\rangle \geq 0 \qquad \forall x' \in \Phi.$$

As pointed out in §3, the last variational inequality can be written

$$0 \in \partial\delta(x \,|\, \Phi) + A^T C\,(Ax), \tag{4.9}$$

where $\partial\delta(x \,|\, \Phi)$ is the normal cone to Φ at x. Note that, by (4.2), we may write

$$\partial\delta(x \,|\, \Phi) = \partial\delta(x_1 \,|\, \Phi_1) \times \cdots \times \partial\delta(x_K \,|\, \Phi_K).$$

Since problem (4.9) is of the form (3.1), we are able to apply the algorithms described in the previous section, with the identifications $F = \partial\delta(x \,|\, \Phi)$ and $G = C$. In the following we shall restrict ourselves to Algorithm I. The arguments for Algorithm II are similar.

First observe that the problem to be solved in Step 2 of Algorithm I may be written

$$0 \in \lambda(x - z^{(k)}) + \partial\delta(x \,|\, \Phi),$$

which is equivalent to solve the minimization problem

$$\begin{aligned} &\text{minimize} \quad ||x - z^{(k)}||^2 \\ &\text{subject to} \quad x \in \Phi, \end{aligned}$$

which, by (4.2), can be decomposed into K independent smaller problems

$$\begin{aligned} &\text{minimize} \quad ||x_\alpha - z_\alpha^{(k)}||^2 \\ &\text{subject to} \quad x_\alpha \in \Phi_\alpha. \end{aligned} \tag{4.10}$$

Since Φ_α is given by (4.1), problems (4.10) are quadratic minimum cost (single-commodity) flow problems. (We have ignored the positive parameter λ in problems (4.10), since their solutions are not affected by the presence of λ.)

Next, by (4.7) and (4.8), the problem to be solved in Step 3 becomes the equation

$$\frac{1}{\lambda K}(w - r^{(k)}) + G(w) = 0, \tag{4.11}$$

where $r^{(k)}$ is given by

$$r^{(k)} = \sum_{\alpha=1}^{K} (2x_\alpha^{(k)} - z_\alpha^{(k)}). \tag{4.12}$$

The vector $v^{(k)}$ can then be obtained by

$$v^{(k)} = \frac{1}{\lambda K}(r^{(k)} - w^{(k)}). \tag{4.13}$$

Finally Step 4 can be written simply as

$$z_\alpha^{(k+1)} = x_\alpha^{(k)} - \lambda v^{(k)} \qquad \alpha = 1, \cdots, K. \tag{4.14}$$

The above procedure for solving the traffic equilibrium problem can be summarized as follows:

Decomposition Algorithm for the Traffic Equilibrium Problem:

Step 1: Choose a constant $\lambda > 0$ and an initial vector $z^{(0)} = (z_1^{(0)}, \cdots, z_K^{(0)}) \in R^{K|\mathcal{A}|}$. Let $k := 0$.

Step 2: For each OD pair $\alpha = 1, \cdots, K$, solve the quadratic minimum cost single-commodity flow problem (4.10) to obtain $x_\alpha^{(k)} \in R^{|\mathcal{A}|}$.

Step 3: Compute $r^{(k)} \in R^{|\mathcal{A}|}$ by (4.12) and solve the nonlinear equation (4.11) to obtain $w^{(k)} \in R^{|\mathcal{A}|}$. Compute $v^{(k)} \in R^{|\mathcal{A}|}$ by (4.13).

Step 4: For each $\alpha = 1, \cdots, K$, compute $z_\alpha^{(k+1)} \in R^{|\mathcal{A}|}$ by (4.14). Set $k := k + 1$ and go to Step 2. □

By Theorem 2.1, we may deduce that the sequence $\{x^{(k)}\}$ generated by the above algorithm converges to a solution of the traffic equilibrium problem (4.6). Note that the computations in Steps 2 and 4 can be done in parallel for all $\alpha = 1, \cdots, K$.

5. Concluding Remarks

We have applied Douglas-Rachford splitting algorithm to the particular class of multi-valued equations (3.1). To conclude the paper, we compare the algorithms presented here with the alternating direction method of multipliers [9] for solving (3.1). (In fact, Gabay [9] considered the latter method to solve the variational inequality problem (3.6), which is a slightly restricted version of (3.1).) Associated with problem (3.1) is the following "dual" problem

$$0 \in G^{-1}(y) - AF^{-1}(-A^T y). \tag{5.1}$$

For the special case of the minimization problem (3.3), problem (5.1) reduces to its Fenchel dual

$$\text{minimize} \quad g^*(y) + f^*(-A^T y),$$

where f^* and g^* are the conjugate functions of f and g, respectively. Since (5.1) belongs to the class of problems (2.2), Douglas-Rachford algorithm (2.3) may be applied to solve problem (5.1). In particular, for the identifications $R = (-A) \circ F^{-1} \circ (-A^T)$ and $S = G^{-1}$, the resulting algorithm can be stated as follows:

Alternating direction method of multipliers:

Step 1: Choose a constant $\lambda > 0$ and $y^{(0)}, p^{(0)} \in Y$. Let $k := 0$.

Step 2: Find a solution $x^{(k+1)}$ of the equation

$$0 \in F(x) + A^T \left[p^{(k)} + \lambda (Ax - y^{(k)}) \right].$$

$$\left[x^{(k+1)} = \arg\min_{x \in X} \left\{ f(x) + \langle p^{(k)}, Ax \rangle + \frac{\lambda}{2} ||Ax - y^{(k)}||^2 \right\} \right]$$

Step 3: Find a solution $y^{(k+1)} \in Y$ of the equation

$$0 \in G(y) - p^{(k)} + \lambda \left(y - Ax^{(k+1)} \right).$$

$$\left[y^{(k+1)} = \arg\min_{y \in Y} \left\{ g(y) - \langle p^{(k)}, y \rangle + \frac{\lambda}{2} ||y - Ax^{(k+1)}||^2 \right\} \right]$$

Step 4: Put

$$p^{(k+1)} = p^{(k)} + \lambda \left(Ax^{(k+1)} - y^{(k+1)} \right).$$

Set $k := k + 1$ and go to Step 2. $\square$

The formulas in the square brackets correspond to the minimization problem (3.3). Notice that the above algorithm is solely represented in terms of the mappings F and G involved in the original ("primal") problem (3.1), rather than the mappings F^{-1} and G^{-1} involved in (5.1). In order to guarantee the maximal monotonicity of the mapping $(-A) \circ F^{-1} \circ (-A^T)$, Gabay [9] assumes either that F is strongly monotone or that $A^T A$ is an isomorphism (cf. assumption (3.2) imposed on Algorithms I and II in §3.)

Unlike the alternating direction method of multipliers, Algorithms I and II described in §3 requires the inversion of the mapping AA^T. This potential difficulty is probably a reason why the primal application of Douglas-Rachford splitting algorithm to the class of problems (3.1) has not been considered so seriously as the dual application. There are, however, some problems that can be formulated in such a way that AA^T is easy to invert. In fact, we have shown in §4 that the presence of $(AA^T)^{-1}$ does not cause any difficulty for the variational inequality problem derived from a traffic equilibrium problem. Finally we remark that it is possible to apply any of the splitting algorithms given in [12, 15] to problem (3.1) in a manner similar to Douglas-Rachford algorithm.

References

[1] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods* (Prentice-Hall, Englewood Cliffs, N.J., 1989).

[2] H. Brezis, *Opérateurs Maximaux Monotones* (North-Holland, Amsterdam, 1973).

[3] S. Dafermos, "Traffic equilibrium and variational inequalities," *Transportation Science* 14 (1980) 42-54.

[4] J. Eckstein, "Splitting methods for monotone operators with applications to parallel optimization," PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1989.

[5] J. Eckstein, "The alternating step method for monotropic programming on the Connection Machine CM-2," *ORSA Journal on Computing* 5 (1993) 84-96.

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

[7] J. Eckstein and M. Fukushima, "Some reformulations and applications of the alternating direction method of multipliers," in: W.W. Hager, D.W. Hearn and P.M. Pardalos, eds., *Large Scale Optimization: State of the Art* (Kluwer, Boston, to appear).

[8] M. Fukushima, "Application of the alternating direction method of multipliers to separable convex programming problems," *Computational Optimization and Applications* 1 (1992) 93-111.

[9] D. Gabay, "Applications of the method of multipliers to variational inequalities," in: M. Fortin and R. Glowinski, eds., *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems* (North-Holland, Amsterdam, 1983) pp. 299-331.

[10] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximations," *Computers and Mathematics with Applications* 2 (1976) 17-40.

[11] R. Glowinski and A. Marocco, "Sur l'approximation par éléments finis d'ordre un et la résolution par penalisation-dualite d'une classe de problèmes de Dirichlet non linéaires," *Revue Française d'Automatique, Informatique et Recherche Opérationnelle* 9 (1975) 41-76.

[12] P.L. Lions and B. Mercier, "Splitting algorithms for the sum of two nonlinear operators," *SIAM Journal on Numerical Analysis* 16 (1979) 964-979.

[13] B. Martinet, "Régularisation d'inéquations variationelles par approximations successives," *Revue Française d'Automatique, Informatique et Recherche Opérationnelle* 4 (1970) 154-159.

[14] J.-S. Pang, "Asymmetric variational inequality problems over product sets: Applications and iterative methods," *Mathematical Programming* 31 (1985) 206-219.

[15] G.B. Passty, "Ergodic convergence to a zero of the sum of monotone operators in Hilbert space," *Journal of Mathematical Analysis and Applications* 72 (1979) 383-390.

[16] R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, N.J., 1970).

[17] R.T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization* 14 (1976) 877-898.

[18] M.J. Smith, "The existence, uniqueness and stability of traffic equilibria," *Transportation Research* 13B (1979) 295-304.

[19] P. Tseng, "Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming," *Mathematical Programming* 48 (1990) 249-264.

[20] P. Tseng, "Application of a splitting algorithm to decomposition in convex programming and variational inequalities," *SIAM Journal on Control and Optimization* 29 (1991) 119-138.