



LaboratoireARTEMIS

BP 53x - 38041 - Grenoble cedex



Photo Focale Studio Grenoble

Rapport de recherche

MINIMUM COST MULTIFLOWS IN UNDIRECTED NETWORKS

Alexander V. Karzanov

RR 849-M-

Avril 1991

MINIMUM COST MULTIFLOWS IN UNDIRECTED NETWORKS

Alexander V. Karzanov

Institute for Systems Studies, 9, Prospect 60 let Oktyabrya, 117312 Moscow, USSR;
and
Université Fourier Grenoble 1, BP 53x, 38041 GRENOBLE Cedex

Abstract. The problem that we study in this paper arises as a natural extension of the classical minimum-cost maximum-flow problem, on one hand; and the problem of finding a collection of flows (a multiflow) in an undirected network, such that a flow connecting any two terminals is allowed and the sum of values of these flows is maximum, on the other hand. It is known, due to Cherkasskij and Lovász, that the latter problem has a half-integral solution when the capacity function is integral.

We consider the problem of finding a multiflow (consisting of flows between any pairs of terminals) whose total cost is minimum provided that the sum of values of the partial flows is maximum. We prove that it has a half-integral optimal solution whenever the capacities are integral. Moreover, we describe a pseudo-polynomial algorithm to solve this problem and show the existence of a half-integral optimal dual solution. Finally, we develop a strongly polynomial algorithm for finding half-integral optimal primal and dual solutions.

Some of the above-mentioned results have been published in an earlier paper of the author, and here we give other, shorter proofs, while the other results are now presented for the first time.

1. Introduction

Throughout the paper by a *graph (digraph)* we shall mean a finite undirected (directed) graph without loops and multiple edges. VG is the *vertex-set* and EG is the *edge-set (arc-set)* of a graph (digraph) G . An edge of a graph with end vertices x and y may be denoted by xy . A *path*, or an $s-t$ *path*, in a graph (digraph) G is a sequence $P = (s = x_0, e_1, x_1, \dots, e_k, x_k = t)$ where $x_i \in VG$ and $e_i = x_{i-1}x_i \in EG$ (respectively, $e_i = (x_{i-1}, x_i) \in EG$). The path P is simple if all x_i 's are distinct.

We shall deal with an undirected network $N = (G, T, c, a)$. Here G is a graph and T is a subset of its vertices, called *terminals* in N . An edge $e \in EG$ has a nonnegative *capacity* $c(e)$ and *cost* $a(e)$. Throughout the paper, if another condition is not explicitly stated, we assume that both c and a are *integer-valued*.

Let $\mathcal{P} := \mathcal{P}(G, T)$ denote the set of simple $s-t$ paths in G with $s, t \in T$ and $s \neq t$. A (c -admissible) *multiflow*, or *multicommodity flow*, in N is a mapping $f : \mathcal{P} \rightarrow \mathbf{Q}_+$ satisfying the capacity constraint

$$(1) \quad \zeta^f(e) := \sum (f(P) \mid e \in P \in \mathcal{P}) \leq c(e) \quad \text{for all } e \in EG$$

(\mathbf{Q}_+ is the set of nonnegative rationals). For $s, t \in T$ the restriction f_{st} of f on the set of $s-t$ paths is called a *flow* from s to t . (Thus a flow from s to t is considered to be a function on the simple $s-t$ paths; such a definition is known to be equivalent to the usual notion of a flow as a function on directed edges satisfying conservation conditions,

cf. [FF]; the former definition will be more convenient for our purposes). The *total value* $v(f)$ of f is $\sum(f(P) \mid P \in \mathcal{P})$, and the *total cost* $a(f)$ is $\sum(a(e)\zeta^f(e) \mid e \in EG)$. A multiflow f is called *maximal* if its total value $v(f)$ is as great as possible.

In the present paper we study the problem

- (2) *given* $N = (G, T, c, a)$, *find a maximal multiflow* f *in* N *whose total cost* $a(f)$ *is minimum.*

When $a = 0$, (2) turns into the problem (*) of finding a maximal multiflow (in which flows connecting any two distinct terminals are allowed). It was proved independently by Cherkasskij [Ch] (this paper was submitted for publication in 1973) and Lovász [Lo] that there exists a maximal multiflow f that is half-integral (i.e., the function $2f$ is integer-valued). Moreover, if c is *inner Eulerian*, that is, for any $x \in V - T$ the value $\sum(c(e) \mid e \in EG, e \text{ incident to } x)$ is even, then (*) has an integral solution. Also it was shown in [Ch] that a half-integral (integral if c is inner Eulerian) solution of (*) can be found in strongly polynomial time (another, faster, algorithm of complexity $O(\eta_n \log |T|)$ was developed in [Ka2], where η_n is the running time of a maximal flow procedure for a network with $n := |VG|$ vertices).

When T consists of two terminals, say s and t , (2) turns into another well-known special case. Namely, we get the (undirected) minimum-cost maximum-flow problem: find a maximal flow f_{st} from s to t whose total cost is minimum. A classical result in network flow theory is that the latter problem has an integer solution, see [FF]. This problem is also solvable in strongly polynomial time, due to [Ta1] (see also [GT] for a “pure combinatorial” algorithm).

For $|T| \geq 3$ the problem (2) has, in general, no integral optimal solution even when c is inner Eulerian, as shown by the following simple example:

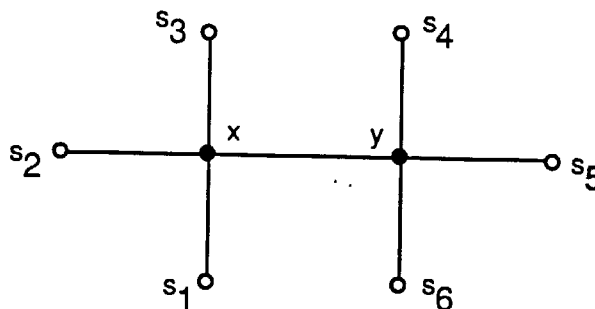


Fig. 1

(Here $T := \{s_1, \dots, s_6\}$; $c(e) = 1$ for all edges e ; $a(xy) = 1$ and $a(e) = 0$ for the other edges.) However, the following is true.

Theorem 1. *Let* $|T| \leq 3$. *Then (2) has a half-integral optimal solution* f , *that is,* $2f$ *is integer-valued.*

The problem in question can be associated with the linear program: given $p \in \mathbf{Q}_+$,

(3) maximize $pv(f) - a(f)$ (subject to (1) and $f \geq 0$).

Clearly (2) is equivalent to (3) when p is a rather large positive number; we shall see later that taking p to be $\tilde{p} := 2a(EG) + 1$ is sufficient. (For $g : S \rightarrow \mathbf{Q}$ and $S' \subseteq S$, $g(S')$ denotes $\sum(g(e) \mid e \in S')$.) Theorem 1 is a direct corollary of the following result.

Theorem 2. For any $p \geq 0$ the program (3) has a half-integral optimal solution.

Theorem 1 (and, in fact, Theorem 2) was stated in [Ka1]; it was a consequence of a pseudo-polynomial algorithm developed there to find a half-integer solution of (2). In the present paper we give a direct proof of Theorems 1 and 2 (in Sections 2 and 4). Furthermore, we present here a pseudo-polynomial algorithm to solve (2) (in Section 3); being, in essence, a modification of the method in [Ka1], this algorithm seems to be simpler in respect to its design as well as the proof of its correctness.

The algorithm is based on ideas of the primal-dual method in linear programming, like the classical algorithm of Ford and Fulkerson [FF] for the min-cost max-flow problem. The program dual to (3) is:

(4) minimize $cl := \sum(c(e)l(e) \mid e \in EG)$ provided that $l \in \mathbf{Q}_+^{EG}$ and $l(P) \geq p - a(P)$ for any $P \in \mathcal{P}$.

(Here for $l' \in \mathbf{Q}_+^{EG}$, $l'(P)$ denotes the “length” $\sum(l'(e) \mid e \in P)$ of a path P .) Further without loss of generality we shall assume that G is connected (otherwise we would consider each component of G separately), and that $c(e) > 0$ for each edge $e \in EG$ (since an edge e with $c(e) = 0$ can be deleted from G). For $l' \in \mathbf{Q}_+^{EG}$ and $x, y \in VG$ let $\text{dist}_{l'}(x, y)$ denote the minimum l' -length $l'(P)$ of an $x - y$ path P in G (the l' -distance between vertices x and y). The system of inequalities in (4) can be rewritten as

(5) $\text{dist}_{a+l}(s, t) \geq p$ for any $s, t \in T, s \neq t$.

The algorithm for (2) consists in an iterative process of solving the program (3) and its dual (4) for a sequence of numbers $p_0 < p_1 < p_2 < \dots$. When p becomes greater than or equal to \tilde{p} , the solution of (2) will be found.

A sketch of our approach to solve (3) and (4) for $p = p_i$ is as follows. The linear programming duality theorem applied to (3)-(4) implies that a (c -admissible) multiflow f and a vector $l \in \mathbf{Q}_+^{EG}$ satisfying (5) are optimal solutions of (3) and (4), respectively, if the following (complementary slackness) conditions hold:

(6) if $P \in \mathcal{P}$ and $f(P) > 0$ then $a(P) + l(P) = p$; in particular, P is an $(a+l)$ -shortest path in G (i.e., a shortest path with respect to the length $a + l$);

(7) if $e \in EG$ and $l(e) > 0$ then e is saturated by f , i.e., $\zeta^f(e) = c(e)$.

The algorithm starts with putting $f := f_0$ to be the zero multifold and $l := l_0$ to be the zero vector on EG ; then (6) and (7) obviously hold for $p := p(a)$. (Here and later, for $\lambda \in \mathbf{Q}_+^{EG}$,

$$(8) \quad p(\lambda) := \min\{\text{dist}_\lambda(s, t) \mid s, t \in T, s \neq t\}.$$

Suppose that after $i - 1$ iterations there have been constructed $f = f_{i-1}$ and $l = l_{i-1}$ which satisfy (6)-(7) for the corresponding p . The i -th iteration consists of two stages:

at the *first stage*, f is transformed into a new multifold f' ($=: f_i$) such that $v(f')$ is maximum provided that (6) and (7) hold for f' , l and p .

at the *second stage*, l is transformed into a new vector l' on EG such that for $\lambda := a + l'$ the number $p' := p(\lambda)$ ($=: p_i$) is as great as possible provided that f' , l' and p' satisfy (6) and (7).

The algorithm finishes when, at the second stage of some iteration, p' can be chosen arbitrarily large. Then (6) and (7) hold for f' , l' and $p' := \tilde{p}$. Hence, the final f' is an optimal solution of (3) (and (2)).

Note that for $|T| = 2$ this approach turns, in fact, into that developed in [FF] for the min-cost max-flow problem. In that case, on each iteration, f' and l' are constructed from f and l by using a maximal flow and a minimal cut in an auxiliary directed network; the arcs of such a network correspond to the *feasible* edges e in G (this means that $l(e) = 0$ and e belongs to a path in \mathcal{P} of $(a + l)$ -length p). However, for $|T| \geq 3$ the method of determining f' and l' turns out to be more sophisticated. The central idea is to form a certain directed network Γ , the so-called *double covering network*. Each edge in G_λ corresponds to a pair of non-adjacent arcs in Γ , where $\lambda := a + l$, and G_λ is the subgraph of G whose edges belong to paths in \mathcal{P} of λ -length p . An important property of Γ is that each path in it which goes from a source to a sink corresponds one-to-one to an $(a + l)$ -shortest path in \mathcal{P} . A new multifold f' (vector l') is constructed from f (respectively, l) by using a maximal flow g (minimal cut) in the feasible-arc subgraph of Γ . Moreover, it follows from the integrality of c that g can be chosen integer-valued. The last property will imply the half-integrality of f' and, in particular, will prove Theorem 2 (and 1).

In fact, the algorithm is applicable to arbitrary “real-valued” c and a ; the total amount of elementary operations (i.e., arithmetical, logical and data transfer ones) is bounded exponentially in $|VG|$. When c (or a) is integer-valued, the algorithm is “pseudo-polynomial” (more precisely, the number of operations in it is estimated as $c(EG)$ (respectively, $a(EG)$) times a polynomial in $|VG|$), like the min-cost max-flow algorithm in [FF].

Now we outline the other results of the present paper. One result concerns the dual problem (4). In Section 5 we prove the following.

Theorem 3. *If a is integer-valued and p is an integer then (4) has a half-integral optimal solution l .*

The second result (Section 5) is that half-integral optimal solutions of (3) (or (2)) and (4) can be determined in strongly polynomial time. To do this, we at first find some (rational) optimal solution l of (4). Since (4) can be rewritten as a linear program whose constraint matrix is of size polynomial in $|VG|$, l can be determined in strongly polynomial time by using a general method of Tardos [Ta2], based on the ellipsoid method [Kh]. Then we show that, knowing l , in order to find a half-integral optimal solution f of (3) it suffices to apply only one iteration of the above-mentioned algorithm. This gives a strongly polynomial combinatorial algorithm for finding such an f . (Notice that even if we state (3) in a compact, “node-edge”, form, general linear programming methods do not guarantee us to obtain a half-integral optimal solution for (3) because, as it can be shown, for any integer $k > 0$ there exist G, T (integral) c and a such that kf is not integral for some optimal basis solution of (3).) Finally, l and f enable us to reformulate (4) in terms of potentials. As a result, we get a linear program with a constraint matrix $A = (a_{ij})$ such that each row i of A has at most two non-zero entries, say a_{ij} and a_{ik} , satisfying $|a_{ij}| + |a_{ik}| = 2$. We explain how to find efficiently a half-integral solution of the arising program by using covering network techniques.

In conclusion of this section let us consider a more general concept of the minimum-cost maximum-multiflow problem. More precisely, let $H = (T, U)$ be a graph, called the *commodity graph*, that indicates the pairs of terminals which are allowed to connect by flows. Denote by $\mathcal{P}(G, H)$ the set of simple $s - t$ paths in G such that $\{s, t\}$ is an edge of H . Consider the problem (2) with the set $\mathcal{P} := \mathcal{P}(G, H)$ rather than $\mathcal{P}(G, T)$. E.g., if $|U| = 2$ we deal with the minimum-cost maximum-two-commodity-flow problem. A natural question arises: given H , what is the minimum integer $k := k(H)$ such that, for any G with $VG \supseteq T$, c (integral) and a , the problem (2) for G, c, a and $\mathcal{P}(G, H)$ has an optimal solution f with kf integral?

So $k(H) = 1$ if $|U| = 1$, and $k(H) = 2$ if H is the complete graph K_m with $m \geq 3$ vertices (by Theorem 1). Theorem 1 can be easily generalized as follows (cf. [Ka3]): if H is a complete m -partite graph and $m \geq 3$ then $k(H) = 2$ (if $m = 2$ then $k(H) = 1$). (H is m -partite if there is a partition $\{T_1, \dots, T_m\}$ of T such that $st \in U$ if and only if $s \in T_i$ and $t \in T_j$ for $i \neq j$.) On the other hand, it was shown in [Ka3] that $k(H) = \infty$ unless H is a complete m -partite graph (in particular, for H consisting of two non-adjacent edges).

2. Geodesics and covering networks

In Sections 2 and 3 we consider a restricted class of cost functions a . Namely, we assume that a is *positive*, that is, $a(e) > 0$ for all $e \in EG$. Such an assumption enables us to simplify the description of the algorithm as well as the proof of Theorems 1 and 2. The case of an arbitrary a will occur in Section 4.

The aim of this section is to introduce the notion of a covering network and to

study relations of such a network to the original one. As a result, we prove Theorems 1 and 2 (under the above assumption).

Let us be given a positive function λ on EG . Put $p := p(\lambda)$ where $p(\lambda)$ is defined as in (8). A path connecting two distinct terminals and having λ -length exactly p is called a *geodesic* for λ , or a λ -*geodesic*. Let G_λ be the subgraph of G whose edges belong to λ -geodesics and vertices belong to λ -geodesics or T .

For brevity, a path $P = (s = x_0, e_1, x_1, \dots, e_k, x_k = t)$ in G may be denoted by $x_0x_1 \dots x_k$ (this leads to no confusion because G has no multiple edges). If $s \in X$ and $t \in Y$, we may say that P is a path from X to Y , or an $X - Y$ path. For $0 \leq i \leq j \leq k$, $P(x_i, x_j)$ is the part $x_i x_{i+1} \dots x_j$ of P from x_i to x_j . The $t - s$ path $x_k x_{k-1} \dots x_0$ opposite to P is denoted by P^{-1} . If $Q = y_0 y_1 \dots y_m$ is a path with $y_0 = t$, $P \cdot Q$ stands for the concatenated path $x_0 x_1 \dots x_k y_1 \dots y_m$.

Consider a vertex $v \in VG_\lambda$. Define the potential $\pi(v) := \pi_\lambda(v)$ of v to be the minimum λ -distance from v to $s \in T$. In particular, $\pi(v) = 0$ if $v \in T$. The set $\{s \in T \mid \text{dist}_\lambda(s, v) = \pi(v)\}$ of terminals closest to v is denoted by $T(v) := T_\lambda(v)$. We say that v is *central* if $|T(v)| \geq 2$. The set of central vertices is denoted by $C := C_\lambda$.

(2.1) *Let v belong to a geodesic P from s to t . Then $\pi(v)$ is the minimum of the lengths $\lambda(P(s, v))$ and $\lambda(P(v, t))$.*

Proof. Let for definiteness $\lambda(P(s, v)) \leq \lambda(P(v, t))$. Then $\pi(v) \leq \lambda(P(s, v)) \leq p/2$ (since $\lambda(P) = p$). Suppose that $p(v) < \lambda(P(s, v))$, and let s' be a terminal such that $\pi(v) = \text{dist}_\lambda(s', v)$. As P is λ -shortest, we have $\lambda(P(s, v)) = \text{dist}_\lambda(s, v)$, therefore, $s \neq s'$. Now $\text{dist}_\lambda(s, s') \leq \text{dist}_\lambda(s, v) + \text{dist}_\lambda(v, s') < p$; a contradiction. •

It follows immediately from (2.1) that $\pi(v) \leq p/2$, and this inequality holds with equality if and only if v is central. For $s \in T$ define $V^s := V_\lambda^s$ to be $\{v \in VG_\lambda \mid \text{dist}_\lambda(s, v) < p/2\}$. Then the sets V^s , $s \in T$, and C are pairwise disjoint and give a partition of VG_λ . (2.2) shows that if $P = v_0 v_1 \dots v_k$ is a geodesic from s to t , then there are q and q' such that $v_0, \dots, v_q \in V^s$, $v_{q'}, \dots, v_k \in V^t$ and either $q' = q + 1$, or $q' = q + 2$ and $v_{q+1} \in C$. Let E^s (respectively, $E^{\{s, t\}}$ where $t \in T - \{s\}$) denote the set of edges in G_λ with one end in V^s and the other in $V^s \cup C$ (respectively, in V^t).

(2.2) (i) *If $uv \in EG_\lambda$ then either $uv \in E^s$ (for some $s \in T$) and $|\pi(v) - \pi(u)| = \lambda(uv)$, or $uv \in E^{\{s, t\}}$ (for some $s, t \in T$) and $\pi(u) + \lambda(uv) + \pi(v) = p$. (ii) Each edge of G_λ belongs to exactly one set among E^s ($s \in T$) and $E^{\{s, t\}}$ ($s, t \in T$) (in particular, none of the edges of G_λ connects two central vertices).*

Proof. Consider a geodesic P containing uv . Let P be an $s - t$ path, and let P meet u earlier than v . Then $\lambda(P(s, u)) + \lambda(uv) + \lambda(P(v, t)) = p$. One may assume that $\lambda(P(s, u)) \leq \lambda(P(v, t))$. Since $\lambda(uv) > 0$, $\lambda(P(s, u)) < p/2$, whence $\lambda(P(s, u)) = \pi(u)$ (by (2.1)) and $u \in V^s$. If $\lambda(P(v, t)) \geq p/2$, we have $\lambda(P(s, v)) \leq p/2$, whence $\lambda(P(s, v) = \pi(v)$, by (2.1). This implies $v \in V^s \cup C$, $uv \in E^s$ and $\pi(v) - \pi(u) = \lambda(uv)$. And if $\lambda(P(v, t)) < p/2$ then $\pi(v) = \lambda(P(v, t))$. In this case we obtain $v \in V^t$,

$uv \in E^{\{s,t\}}$ and $\pi(u) + \lambda(uv) + \pi(v) = p$. Part (ii) easily follows from above arguments.

The following statement, converse, in a sense, to (2.2), describes geodesics in terms of potentials.

(2.3) Let $P = v_0v_1 \dots v_k$ be an $s-t$ path in G_λ with $s, t \in T$ and $s \neq t$. The following are equivalent:

- (i) P is a geodesic;
- (ii) there is q , $0 \leq q < k$, such that $\pi(v_i) - \pi(v_{i-1}) = \lambda(v_{i-1}v_i)$ for $i = 1, \dots, q$ and $\pi(v_i) - \pi(v_{i+1}) = \lambda(v_iv_{i+1})$ for $i = q+1, \dots, k-1$.

Proof. (i)→(ii) follows from (2.2). (ii)→(i). Observe that $\pi(u) = \lambda(P(s, u))$ and $\pi(v) = \lambda(P(v, t))$ where $u := v_q$ and $v := v_{q+1}$. Consider a geodesic Q containing the edge uv ; let Q be an $s'-t'$ path, and let Q meet u earlier than v . Then $\lambda(Q) = \lambda(Q(s', u)) + \lambda(uv) + \lambda(Q(v, t'))$, $\lambda(Q(s', u)) \geq \pi(u)$ and $\lambda(Q(v, t')) \geq \pi(v)$, whence $p = \lambda(Q) \geq \lambda(P)$, and therefore, P is a geodesic. •

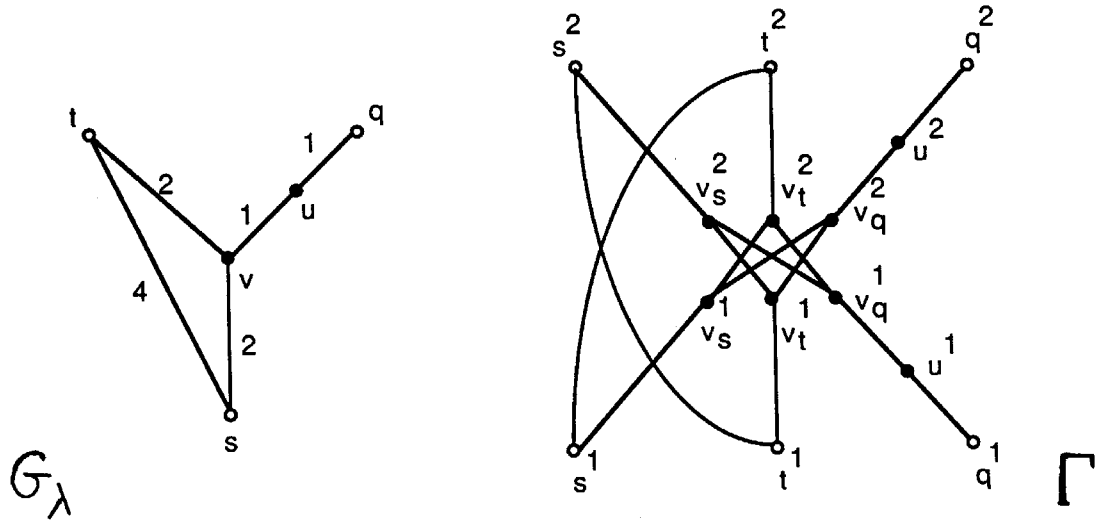


Fig. 2

Now, based on (2.2) and (2.3), we design the covering digraph $\Gamma = \Gamma_\lambda$ for G_λ . Each non-central vertex v of G_λ generates two vertices v^1 and v^2 in Γ . If $v \in VG_\lambda$ is central, it generates $2|T(v)|$ vertices v_s^i , $i = 1, 2$, $s \in T(v)$, in Γ . The arcs of Γ are defined as follows:

- (9) (i) an edge $uv \in E^s$ ($s \in T$) with $\pi(v) - \pi(u) = \lambda(uv)$ induces two arcs (u^1, v^1) and (v^2, u^2) (or (u^1, v_s^1) and (v_s^2, u^2) when v is central) in Γ , each of capacity $c(uv)$;
- (ii) an edge $uv \in E^{\{s,t\}}$ ($s, t \in T$) induces two arcs (u^1, v^2) and (v^1, u^2) in Γ , each of capacity $c(uv)$;
- (iii) a central vertex $v \in C$ induces $2|T(v)|(|T(v)| - 1)$ arcs (v_s^1, v_t^2) in Γ for all

distinct $s, t \in T(v)$, each of capacity ∞ ;

assuming ∞ to be a rather large positive integer. (See Fig. 2 for illustration; here $T = \{s, t, q\}$ and the number on an edge e indicates $\lambda(e)$.)

We shall keep the same notation c for the capacities of arcs in Γ . The subgraph in Γ arising from a central vertex $v \in C$, as well as an arc in it, is called *central*; this subgraph will be denoted by H_v . The set $S := \{s^1 \mid s \in T\}$ is considered to be the set of *sources* of Γ while $S' := \{s^2 \mid s \in T\}$ the set of its *sinks*. Let $P = x_0 x_1 \dots x_k$ be a path in Γ , that is, $(x_{i-1}, x_i) \in E\Gamma$ for $i = 1, \dots, k$ (as above we use brief notation for a path). We say that P is an $S - S'$ path if $x_0 \in S$ and $x_k \in S'$.

The construction of Γ determines the natural mapping τ of $V\Gamma \cup E\Gamma$ onto $VG_\lambda \cup EG_\lambda$ such that $v^i \in V\Gamma$ (or $v_s^i \in V\Gamma$) is mapped by τ to the vertex v , a non-central arc $(x, y) \in E\Gamma$ to the edge $\tau(x)\tau(y)$, and a central arc (v_s^1, v_t^2) to the vertex v .

The mapping τ is naturally extended to paths in Γ and G_λ . Namely, for a path $P = x_0 x_1 \dots x_k$ in Γ , $\tau(P)$ is the path in G_λ induced by the sequence $\tau(x_0), \tau(x_1), \dots, \tau(x_k)$ of vertices (with deleting repeated vertices going in succession).

We also define a mapping $\vartheta : (V\Gamma \cup E\Gamma) \rightarrow (V\Gamma \cup E\Gamma)$ such that a vertex v_i (or v_s^i) is mapped to the other copy v^{3-i} (or v_s^{3-i}), and an arc (x, y) to $(\vartheta(y), \vartheta(x))$. It gives a ("anti"-) symmetry of Γ . For an $s - t$ path $P = x_0 x_1 \dots x_k$ in Γ , $\vartheta(P)$ is the symmetric $\vartheta(t) - \vartheta(s)$ path $\vartheta(x_k)\vartheta(x_{k-1}) \dots \vartheta(x_0)$; obviously, $\tau(P)$ is the path in G opposite to $\tau(\vartheta(P))$.

(2.4) τ determines a one-to-one correspondence between the set of $S - S'$ paths in Γ and the set of λ -geodesics in G .

Proof. In view of (2.2) and (9), for a geodesic P one can directly construct an $S - S'$ path P' in Γ such that $P = \tau(P')$. Conversely, consider an $S - S'$ path $P = x_0 x_1 \dots x_k$ in Γ . It follows from (9) that P contains exactly one arc, say $e = (x_q, x_{q+1})$, such that either u is central, say (v_s^1, v_t^2) , or $\tau(e) \in E^{\{s, t\}}$ for some distinct $s, t \in T$ (let for definiteness $\tau(x_q) \in V^s$). Moreover, $\tau(v_j) \in V^s$ for $j = 1, \dots, q-1$ and $\tau(v_j) \in V^t$ for $j = q+2, \dots, k$. Now (2.3), (9) and the fact that $s \neq t$ imply that $\tau(P)$ is a geodesic. \bullet

One can see that each $S - S'$ path P in Γ is simple and that the paths P and $\vartheta(P)$ are disjoint (since each geodesic in G_λ is a simple path because of the positivity of λ).

We now extend the correspondence of geodesics and $S - S'$ paths to a relationship between certain multiflows in N and $S - S'$ flows in the digraph Γ with the capacity function c as in (9). We say that a multiflow f in N goes along λ -geodesics if

$f(P) > 0$ implies that P is a λ -geodesic.

For a function $g : E\Gamma \rightarrow \mathbf{Q}_+$ and a vertex $x \in V\Gamma$ define

$$(10) \quad \operatorname{div}_g(x) := \sum_{y:(x,y) \in E\Gamma} g(x, y) - \sum_{y:(y,x) \in E\Gamma} g(y, x).$$

g is called a (c -admissible) *flow* from S to S' , or $S - S'$ *flow*, if it satisfies the conservation condition $\operatorname{div}_g(x) = 0$ for all $x \in V\Gamma - (S \cup S')$ as well as the capacity constraint $g(e) \leq c(e)$ for all $e \in E\Gamma$. The *value* $v(g)$ of a flow g is $\sum(\operatorname{div}_g(x) | x \in S)$; g is called maximal if $v(g)$ is as large as possible.

A routine fact is that a flow g as above can be represented as a collection of elementary flows along paths. More precisely, there are $S - S'$ paths P_1, P_2, \dots, P_m ($m \leq |E\Gamma|$) and positive rationals $\alpha_1, \alpha_2, \dots, \alpha_m$ such that:

$$(11) \quad \sum(\alpha_i | e \in P_i) = g(e) \text{ for any } e \in E\Gamma.$$

It follows from (11) that $\sum(\alpha_i | P_i \text{ starts at } s) = \operatorname{div}_g(s)$ for any $s \in S$, whence

$$(12) \quad v(g) = \sum(\alpha_i | i = 1, \dots, m).$$

We say that $\mathcal{D} := \{(P_i, \alpha_i) | i = 1, \dots, m\}$ is a *decomposition* of g . If g is integral then there exists a decomposition with all α_i 's integral; such a decomposition can be found by a trivial procedure of complexity $O(|V\Gamma||E\Gamma|)$ (cf. [FF]). A decomposition \mathcal{D} determines a multiflow $f := f_{\mathcal{D}}$ in N by setting $f(\tau(P_i)) := \alpha_i/2$ and $f(P) := 0$ for the remaining paths P in \mathcal{P} . Indeed, it is easy to see from (11) that for any $e \in E\Gamma$,

$$\zeta^f(\tau(e)) = \frac{1}{2}(g(e) + g(\vartheta(e))) \leq \frac{1}{2}(c(e) + c(\vartheta(e))) = c(\tau(e)),$$

that is, f is c -admissible. Moreover, f goes along geodesics, and, by (12), $2v(f) = v(g)$. A converse property is also true. More precisely, for a (c -admissible) multiflow f in N going along geodesics, define the function $g = g_f$ on $E\Gamma$ such that for $e \in E\Gamma$, $g(e)$ is the sum of values $f(\tau(P))$ over all $S - S'$ paths P in Γ containing e . One can check that g_f is a (c -admissible) $S - S'$ flow in Γ , and $v(g_f) = 2v(f)$. These observations are summarized as follows.

(2.5) (i) If g is an $S - S'$ flow in Γ and $\mathcal{D} = (P_i, \alpha_i)$ is a decomposition of g , then $f_{\mathcal{D}}$ is a multiflow in N going along geodesics, and $v(g) = 2v(f)$. Moreover, if all α_i 's are integral, then $f_{\mathcal{D}}$ is half-integral. (ii) If f is a multiflow in N going along geodesics, then g_f is an $S - S'$ flow in Γ , and $v(g) = 2v(f)$. (iii) If g is an $S - S'$ flow in Γ , \mathcal{D} is a decomposition of g , and $f := f_{\mathcal{D}}$, then $g_f = g$. •

Since c is integral, there exists an integral maximal $S - S'$ flow in Γ . This gives the following corollary of (2.5) (it will not be used later, but is interesting in its own right).

(2.6) If c is integral then there exists a half-integral maximal multiflow in N going along geodesics. •

Now we can prove Theorem 2 (and Theorem 1) in the assumption of positivity of a . Given $p \geq 0$, suppose that f and l are optimal solutions of (3) and (4), respectively.

Let $\lambda := a + l$. By (6), f is a multiflow in N going along λ -geodesics. Consider the $S - S'$ flow $g := g_f$ in $\Gamma := \Gamma_\lambda$. We say that an arc $e \in E\Gamma$ is *feasible* if either e is central or $l(\tau(e)) = 0$; the set of feasible arcs is denoted by A . In view of (7), for each $e \in E\Gamma - A$ the edge $\tau(e)$ in G is saturated by f (since $l(\tau(e)) > 0$), which implies that $g(e) = c(e)$, and hence g is integral on $E\Gamma - A$. Let γ be the restriction of g on A . We say that a function $h : A \rightarrow \mathbf{Q}_+$ is *flow-like* γ if h is c -admissible (i.e., $h(e) \leq c(e)$ for $e \in A$) and

$$(13) \quad \operatorname{div}_h(x) = \operatorname{div}_\gamma(x) \quad \text{for all } x \in V\Gamma - (S \cup S'),$$

(where for a function b on A , div_b is defined as in (10) with respect to A rather than $E\Gamma$).

Since c is integral, there is an *integral* h which is flow-like γ . Then g' , defined by $g'(e) := h(e)$ for $e \in A$ and $g'(e) := g(e)$ for $e \in E\Gamma - A$, is an integral $S - S'$ flow in Γ . Now the multiflow $f' := f_{\mathcal{D}}$, where $\mathcal{D} = \{(P_i, \alpha_i)\}$ is a decomposition of g with all α_i 's integral, gives a half-integral multiflow in N . Since g' can differ from g only on arcs in A , f' satisfies (7) (with l as before). Furthermore, (2.5)(i) implies (6) for f' . Hence, f' is an optimal solution of (3), and the result follows.

3. Algorithm

As before we assume that the cost function a is positive everywhere. Let $f := f_{i-1}$ be a multiflow in N and $l := l_{i-1}$ be a function on EG that have been constructed after $i - 1$ iterations and satisfy (6) and (7) with $p := p_{i-1} := p(\lambda)$, where $\lambda := a + l$.

The first stage of the iteration. As we mentioned in the Introduction, it consists in a transformation of f to a new multiflow f' , and we now describe it. Let G_λ and $\Gamma = \Gamma_\lambda$ be objects defined as in the previous section for given λ ; g be the $S - S'$ flow g_f in Γ induced by f ; and A be the set of feasible arcs in Γ , that is, either e is central or $l(\tau(e)) = 0$. Then $g(e) = g(\vartheta(e)) = c(e)$ for any $e \in E\Gamma - A$, by (7).

Let γ be the restriction of g on A . Determine a function h on A such that h is flow-like γ and maximal, that is, the value $v(h) := \sum(\operatorname{div}_h(x) \mid x \in S)$ is as great as possible. (Such an h can be found by usual flow techniques using a polynomial in $|V\Gamma|$ number of operations, cf., e.g., [ADK].) Form the function g' on $E\Gamma$ that coincides with h on A and with g on $E\Gamma - A$. Find a decomposition \mathcal{D} of g' . Then the multiflow f' resulting on the i -th iteration is just $f_{\mathcal{D}}$. It was explained at the end of Section 2 that f' , l and p satisfy (6) and (7). We now demonstrate other properties of f' , which will be used in what follows.

Let $L = (s = x_0, e_1, x_1, \dots, e_k, x_k = t)$ be a sequence such that $s \in S$, $x_i \in V\Gamma$, $e_i \in A$, and e_i is either (x_{i-1}, x_i) or (x_i, x_{i-1}) . We say that L is an *active path* with respect to a function $h' \in \mathbf{Q}_+^A$ if $h'(e_i) < c(e_i)$ holds for each *direct* arc $e_i = (x_{i-1}, x_i)$ of L , and $h'(e_i) > 0$ holds for each *reverse* arc $e_i = (x_i, x_{i-1})$. Standard arguments in network flow theory show that if h' is flow-like γ then exactly one of the following is true: (i) h' is maximal, or (ii) there is an active path with respect to h' which reaches some sink in S' .

For $Z \subseteq V\Gamma$ and $B \subseteq E\Gamma$ let \bar{Z} denote $V\Gamma - Z$, and $(Z)_B$ denote the set of arcs $(y, z) \in B$ with $y \in Z$ and $z \in \bar{Z}$. Let X be the set of vertices in $V\Gamma$ that are reachable by active paths for h as above. Then $S \subseteq X$. Furthermore, since h is maximal, we have $X \cap S' = \emptyset$,

$$(14) \quad h(e) = c(e) \text{ for } e \in (X)_A, \quad \text{and } h(e) = 0 \text{ for } e \in (\bar{X})_A.$$

Consider the set $Y := \vartheta(X)$ in $V\Gamma$ "symmetric" to X .

(3.1) (i) $h(e) = c(e)$ for $e \in (\bar{Y})_A$, and $h(e) = 0$ for $e \in (Y)_A$. (ii) X and Y are disjoint.

Proof. For the $S-S'$ flow g' as above and a subset $Z \subseteq V\Gamma$ let $\alpha(Z)$ denote $g'((Z)_{E\Gamma}) - g'((\bar{Z})_{E\Gamma})$ (recall that for $\tilde{g} : E \rightarrow \mathbf{Q}$ and $E' \subseteq E$, $\tilde{g}(E')$ is $\sum(\tilde{g}(e) \mid e \in E')$). Clearly $\alpha(Z) = \sum(\text{div}_{g'}(x) \mid x \in Z)$. This implies

$$(15) \quad \alpha(X) = \alpha(\bar{Y}) = \sum(\text{div}_{g'}(s) \mid s \in S) = v(g')$$

(taking into account that $S \subseteq X \subseteq \bar{S}'$ and $S \subseteq \bar{Y} \subseteq \bar{S}'$). Put $B := E\Gamma - A$. By the definition of g' , we have:

$$(16) \quad \alpha(X) = g((X)_B) + h((X)_A) - g((\bar{X})_B) - h((\bar{X})_A); \quad \text{and}$$

$$(17) \quad \alpha(\bar{Y}) = g((\bar{Y})_B) + h((\bar{Y})_A) - g((Y)_B) - h((Y)_A).$$

By symmetry, $(\bar{Y})_C = \vartheta((X)_C)$ and $(Y)_C = \vartheta((\bar{X})_C)$ for any $C \subseteq E\Gamma$. This and the fact that $g(e) = c(e)$ for any $e \in B$ imply $g((X)_B) - g((\bar{X})_B) = g((\bar{Y})_B) - g((Y)_B)$. Then, in view of (15)-(17),

$$(18) \quad h((X)_A) - h((\bar{X})_A) = h((\bar{Y})_A) - h((Y)_A).$$

By (14), $h((X)_A) - h((\bar{X})_A) = c((X)_A)$. Furthermore, $h((\bar{Y})_A) - h((Y)_A) \leq h((\bar{Y})_A) \leq c((\bar{Y})_A)$ (since h is c -admissible). Now (18) and $c((\bar{Y})_A) = c((X)_A)$ implies (i).

To prove (ii), consider the set $Z := X \cap \bar{Y}$. By (i) and (14), each arc in $(Z)_A$ is saturated by h , and $h(e) = 0$ for any $e \in (\bar{Z})_A$. This means that no vertex in \bar{Z} is reachable by an active path. Hence, $X \subseteq Z$, and (ii) is true. •

The second stage of the iteration. Let X and Y be the sets as above. Put $W := V\Gamma - (X \cup Y)$. For subsets $Z, Z' \subseteq V\Gamma$ denote by (Z, Z') the set of arcs (x, y) in Γ with $x \in Z$ and $y \in Z'$. We say that $u \in E\Gamma$ is

Proof. If $e \in EG - EG_\lambda$ then $l(e) = 0$ (since $\zeta^f(e) = 0 < c(e)$, and (7) holds for f and l), whence $l^\varepsilon(e) = 0$ (by (20)). Consider an edge $e \in EG_\lambda$. Let $u \in E\Gamma$ be such that $\tau(u) = e$. If $u \notin A$ then $\zeta^{f'}(e) = \zeta^f(e) = c(e)$. And if $u \in A$ then $l'(e) > 0 = l(e)$ would imply that u is either a (+)-arc or a (++)-arc (by (20)), whence $\zeta^{f'}(e) = c(e)$, in view of (14) and (3.1). •

Let Q be a path in Γ . We say that Q is *regular* if Q has no negative arcs in A . In particular, it follows from (14) and (3.1) that Q is regular if $g'(u) > 0$ for all arcs u of Q . Define $n^+ := n_Q^+$ to be the number of (+)-arcs in Q ; and similarly define the numbers n^{++} , n^- and n^{--} . It is easy to see that if Q is an $S - S'$ path then

$$(24) \quad 2n^+ + n^{++} - 2n^- - n^{--} = 2.$$

(3.3) (i) Let P be a λ -geodesic in G , and let Q be an $S - S'$ path in Γ such that $P = \tau(Q)$. If Q is regular then $l^\varepsilon(P) = l(P) + 2\varepsilon$; otherwise $l^\varepsilon(P) > l(P) + 2\varepsilon$.
(ii) f' , l^ε and p' satisfy (6).

Proof. (i) obviously follows from (20) and (24). To see (ii), consider $P \in \mathcal{P}$ with $f'(P) > 0$. We know that P is a λ -geodesic (as f' , l and p satisfy (6)). Moreover, f' was yielded from a decomposition \mathcal{D} of g' , so Q is a member of \mathcal{D} . Then $g'(u) > 0$ for all arcs u of Q . Hence, Q is regular, and we now obtain from (i) that $\lambda^\varepsilon(P) = a(P) + l(P) + 2\varepsilon = p + 2\varepsilon = p'$. •

We observe that $\tilde{\varepsilon} > 0$. Indeed, three types of paths P in \mathcal{P} are possible. (i) P is a λ -geodesic with $f(P) > 0$. Then $\lambda^\varepsilon(P) = p + 2\varepsilon$ for any $\varepsilon \leq \varepsilon_1 > 0$, by (3.3)(i). (ii) P is a λ -geodesic with $f(P) = 0$. Then $\lambda^\varepsilon(P) \geq p + 2\varepsilon$. (iii) P is not a geodesic for λ , i.e., $\lambda(P) > p$. Since $l^\varepsilon(P)$ is a linear function of ε , and $l^0(P) = l(P)$, there exists a (rather small) $\varepsilon' > 0$ such that $a(P) + l^\varepsilon(P) \geq p + 2\varepsilon$ for any $\varepsilon \leq \varepsilon'$. Note that the set of paths as in (i) is nonempty (since G is connected and c is positive, whence $g' \neq 0$ and $f' \neq 0$). Thus (22) is true for some $\varepsilon > 0$ (and any smaller ε).

Arguments above prompt an efficient procedure to determine $\tilde{\varepsilon}$. Note that if P is a simple path in G then the value $l^\varepsilon(P)$ is expressed as $l(P) + \varepsilon k(P)$ for some integer $k(P)$ such that $|k(P)| < 2|VG|$ (since P has at most $|VG| - 1$ edges, and for any $e \in EG$, $l^\varepsilon(e) = l(e) + \beta\varepsilon$ where $\beta \in \{-2, -1, 0, 1, 2\}$). Put $\varepsilon := \varepsilon_1$, and determine p^ε by solving $|T|$ shortest path problems for G with the length function λ^ε (if $\varepsilon_1 = \infty$ we put ε to be a rather large positive integer). If $p^\varepsilon = p + 2\varepsilon$ then ε_1 is just $\tilde{\varepsilon}$. But if $p^\varepsilon < p + 2\varepsilon$ then choose a path $P_1 \in \mathcal{P}$ such that $\lambda^\varepsilon(P_1) = p^\varepsilon$ (by an argument above, P_1 is not a λ -geodesic); and put ε_2 to be the solution of

$$(25) \quad (\lambda^{\varepsilon_{j+1}}(P_j) :=) \quad a(P_j) + l(P_j) + \varepsilon_{j+1}k(P_j) = p + 2\varepsilon_{j+1}$$

with $j := 1$. Now determine p^ε for $\varepsilon := \varepsilon_2$, and so on. Suppose we have already executed q steps and found rationals $\varepsilon_1 > \varepsilon_2 > \dots > \varepsilon_{q+1} > 0$ and paths

$P_1, P_2, \dots, P_q \in \mathcal{P}$ such that, for $j = 1, \dots, q$, the path P_j and the numbers ε_j and ε_{j+1} satisfy

$$(26) \quad (\lambda^{\varepsilon_j}(P_j) :=) \quad a(P_j) + l(P_j) + \varepsilon_j k(P_j) = p^{\varepsilon_j} < p + 2\varepsilon_j$$

and (25). Comparing (25) for $j := q - 1$ and (26) for $j := q$ we obtain $\lambda^{\varepsilon_q}(P_{q-1}) > \lambda^{\varepsilon_q}(P_q)$. On the other hand, $\lambda^{\varepsilon_{q-1}}(P_{q-1}) \leq \lambda^{\varepsilon_{q-1}}(P_q)$ (as P_{q-1} is a geodesic for $\lambda^{\varepsilon_{q-1}}$). Since $\varepsilon_q < \varepsilon_{q-1}$, we obtain from these inequalities that $k(P_{q-1}) < k(P_q)$. This implies $q < 4|VG|$. Thus $\tilde{\varepsilon}$ will be found in $O(|VG|)$ steps as above.

Finiteness and complexity of the algorithm. The value p strictly increases during the algorithm because of (22) and $\tilde{\varepsilon} > 0$. The algorithm finishes when, on some iteration, $\tilde{\varepsilon}$ becomes ∞ . Then, by (22) and (3.3)(ii), $f', l' := l^\varepsilon$ and $p' := p^\varepsilon$ satisfy (6) and (7) for any $\varepsilon \geq 0$; taking a rather large ε , we obtain optimal solutions of (2) and (4). Now we have, first, to prove that the algorithm terminates in a finite number of iterations and, second, to estimate this amount.

We say that an iteration i is *positive* if the value of the current multiflow increases on this iteration (i.e., $v(f_{i-1}) < v(f_i)$), and *negative* otherwise. We shall prove below that

$$(27) \quad \text{the number of negative iterations going in succession is at most } |VG| - |T|.$$

Assuming that (27) is valid, we complete examination of the algorithm as follows. By (27), if the number of occurrences of positive iterations is M then the total number of iterations in the algorithm is $O(nM)$ where $n := |VG|$. Hence, the algorithm uses $O(MP(n))$ operations, P is a polynomial (since an iteration takes a polynomial in n number of operations). We estimate M for three cases.

(C1) If c is integer-valued, then, as we explained earlier, the value $v(f)$ of any current multiflow f is half-integral. Obviously, $v(f)$ does not exceed $c(EG)$. Therefore, M is bounded by $2c(EG)$, and the complexity of the algorithm is $O(P(n)c(EG))$, P is a polynomial.

(C2) Let a be integer-valued. Consider a positive iteration, and let $L = x_0x_1 \dots x_k$ be an active path entering S' in the corresponding Γ . Denote by L^+ (L^-) the set of non-central direct (reverse) arcs in L . Using the fact that for any (directed) $S - S'$ path Q in Γ the $(a + l)$ -length of $\tau(Q)$ is p , one can see that

$$\sum_{u \in L^+} (a(\tau(u)) + l(\tau(u))) - \sum_{u \in L^-} (a(\tau(u)) + l(\tau(u))) = p.$$

Furthermore, $l(\tau(u)) = 0$ for each $u \in L^+ \cup L^-$, and there are at most two arcs u in L with the same $\tau(u)$. Hence,

$$(28) \quad p = \sum_{u \in L^+} a(\tau(u)) - \sum_{u \in L^-} a(\tau(u)) \leq 2a(EG),$$

whence $M \leq 2a(EG)$ (since a is integral and p increases during the algorithm). Thus the algorithm has complexity $O(P(n)a(EG))$.

(C3) Suppose both c and a are arbitrary (“real-valued”). (28) shows that M does not exceed the number of functions on EG taking their values in $\{-2, -1, 0, 1, 2\}$, and we therefore obtain a corresponding finite upper bound for the amount of iterations, which depends only on n (as an exponential). Of course, such a bound is viewed as too large and inexact.

It remains to prove (27). Consider two subsequent iterations i and $i + 1$. Let l denote the current function on EG at the *beginning* of the i -th iteration; g denote the $S - S'$ flow in $\Gamma := \Gamma_\lambda$, $\lambda := a + l$, *resulting* on the i -th iteration; τ denote τ_λ ; and X denote the set of vertices of Γ reachable by active paths *after* the first stage of the i -th iteration. Analogous objects for the $(i + 1)$ -th iteration will be denoted with primes.

(3.4) Lemma. *Let the $(i + 1)$ -iteration be negative. Then $\tau'(X')$ strictly includes $\tau(X)$.*

This lemma immediately implies (27) (taking into account that $T \subseteq \tau(X)$ because of $S \subseteq X$).

Proof. First of all we show that $\tau(X) \subseteq \tau'(X')$. We use notations A, Y, W for the corresponding objects in Γ as in the description of the algorithm; analogous objects in Γ' are denoted by A', Y', W' .

Note that there is a simple relationship between g and g' . Namely, let f be the multiflow in N resulting on the i -th iteration; then $f = f_{\mathcal{D}}$ where $\mathcal{D} = \{(Q_j, \alpha_j) \mid j = 1, \dots, m\}$ is a decomposition of g . Since the $(i + 1)$ -iteration is negative, there is a decomposition $\mathcal{D}' = \{(Q'_j, \alpha'_j) \mid j = 1, \dots, m'\}$ of g' such that $m' = m$, $\alpha'_j = \alpha_j$, and the path $\tau'(Q'_j)$ coincides with $\tau(Q_j)$. Note also that if $u \in E\Gamma$ is an arc with $g(u) > 0$ then there is a regular $S - S'$ path in Γ which contains u (e.g., a path in \mathcal{D} which traverses u).

Claim 1. *Each arc $u = (x, y) \in E\Gamma$ with $x, y \in X$ belongs to a regular $S - S'$ path in Γ .*

Proof. We show that there are regular paths R_1 from S to x and R_2 from y to S' ; then $R_1 \cdot u \cdot R_2$ gives an $S - S'$ path as required.

Choose an active (with respect to g) path $L = x_0x_1 \dots x_k$ in Γ with $x_k = x$. If L has no reverse arc then we can put $R_1 := L$ (all vertices of L are in X , and therefore, L has no negative arcs). Otherwise, let j be the maximal index such that $u := (x_j, x_{j-1})$ is a reverse arc in R_1 . Then $g(u) > 0$, whence there is a regular $S - S'$ path Q containing u . Put $R_1 := Q(s, x_j) \cdot L(x_j, x)$, where s is the first vertex in Q .

Next, choose a path $L' = y_0y_1 \dots y_m$ in Γ from y to S' . Let y_j be the first vertex of L' not in X . Then $u' := (y_{j-1}, y_j)$ is either a (+)-arc or a (++)-arc, whence $g(u') = c(u') > 0$. Choose a regular $S - S'$ path Q' containing u' . Put $R_2 := L'(x, y_j) \cdot Q'(y_j, t)$,

